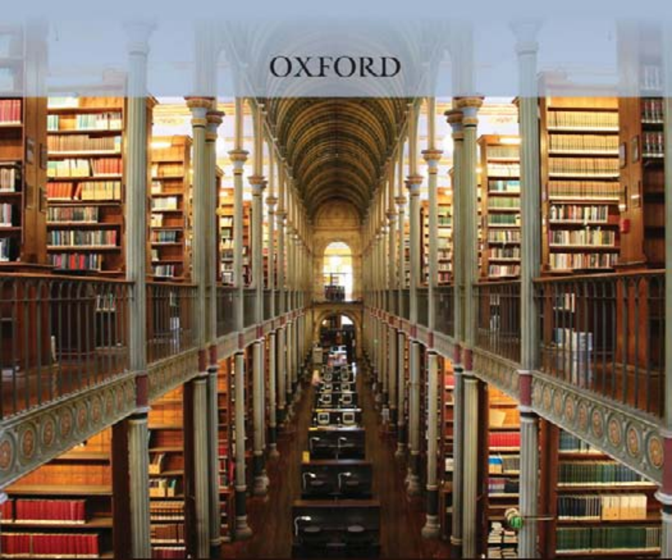
OXFORD



Words and the Mind

How Words Capture Human Experience

Barbara C. Malt
Phillip Wolff

Words and the Mind

This page intentionally left blank

Words and the Mind

How Words Capture Human Experience

Edited by
Barbara C. Malt and Phillip Wolff

New York Oxford
OXFORD UNIVERSITY PRESS
2010

Oxford University Press

Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Library of Congress Cataloging-in-Publication Data
Words and the mind : How words capture human experience / Barbara C. Malt
and Phillip Wolff, editors.
p. cm.
ISBN 978-0-19-531112-9 (cloth : alk.paper)
1. Psycholinguistics. I. Malt, Barbara C. II. Wolff, Phillip M.
P37.W68 2010
401′.9—dc22
2009016072

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

TABLE OF CONTENTS

CONTRIBUTORS

Jürgen Bohnemeyer
Linguistics Department
University at Buffalo
Buffalo, NY

Eve V. Clark
Department of Linguistics
Stanford University
Stanford, CA

Silvia Gennari
Department of Psychology
University of York
York, United Kingdom

Aubrey L. Gilbert
Department of Psychology
University of California, Berkeley
Berkeley, CA

Cliff Goddard
School of Behavioural, Cognitive and Social Sciences
University of New England
Armidale NSW
Australia

Peter Gordon
Biobehavioral Sciences Department
Teachers College, Columbia University
New York, NY

J. Richard Hanley
Department of Psychology
University of Essex
Colchester, United Kingdom

Kathy Hirsh-Pasek
Department of Psychology
Temple University
Ambler, PA

Mutsumi Imai
Keio University at Shonan-Fujisawa
Fujisawa
Kanagawa
Japan

Richard B. Ivry
Department of Psychology
University of California, Berkeley
Berkeley, CA

Ga-hyun Jeon
Department of Psychology
University of Illinois
Champaign, IL

Paul Kay
Department of Linguistics
University of California, Berkeley
Berkeley, CA

David Kemmerer
Department of Speech, Language, and Hearing Sciences
Purdue University
West Lafayette, IN

Bianca Klettke
School of Psychology
Deakin University
Burwood, Australia

Yu Li
Department of Russian and East Asian Languages and Cultures
Emory University
Atlanta, GA

John A. Lucy
Department of Comparative Human Development
University of Chicago
Chicago, IL

Weiyi Ma
School of Education
University of Delaware
Newark, DE

Asifa Majid
Max Planck Institute for Psycholinguistics
Nijmegen, The Netherlands

Barbara C. Malt
Department of Psychology
Lehigh University
Bethlehem, PA

Roberta Michnick Golinkoff
School of Education
University of Delaware
Newark, DE

Julia Parish-Morris
Department of Psychology
Temple University
Ambler, PA

Shannon M. Pruden
Department of Psychology
University of Chicago
Chicago, IL

Terry Regier
Department of Linguistics, and Cognitive Science Program
University of California, Berkeley
Berkeley, CA

Debi Roberson
Department of Psychology
University of Essex
Colchester, United Kingdom

Henrik Saalbach
ETH Zurich
Zurich, Switzerland

Ann Senghas
Department of Psychology
Barnard College of Columbia University
New York, NY

Phillip Wolff
Department of Psychology
Emory University
Atlanta, GA

Words and the Mind

This page intentionally left blank

THE LANGUAGE–THOUGHT INTERFACE

An Introduction

Phillip Wolff and Barbara C. Malt

The world presents a dazzlingly rich array of sights and sounds, actions, and events to its observers. The cognitive processes that allow humans to make sense of this rich sensory input and that guide their interactions with the world are, in a number of respects, shared with other higher mammals. But only humans have the added capacity of language, allowing them to selectively capture some of this richness in words and thereby receive and transmit information about the world through a symbolic system. This symbolic system not only facilitates communication with the outside world but may also provide tools for the mental manipulation of information (e.g., Gentner, 2003).

Although language may be crucial to human cognition, the basic units of cognition are clearly not words. For instance, people can have thoughts that are difficult to express, and they understand expressions that are ambiguous in ways that their thoughts are not. When people see a sign in a restaurant that says "Please wait for the hostess to be seated," they do not puzzle over whether they should wait for the hostess to seat herself or whether she will guide them to their seat. Furthermore, if words were the units of thought, new words could not be coined, and no one would ever say "that's not what I meant to say" (Pinker, 1994). These logical arguments and more indicate that there must be a medium of thought that is independent of language (Fodor, 1975). This book is about how this medium of thought is coordinated with the knowledge of words. That is, it addresses the *language–thought interface*, with a focus on the portion of language that constitutes the lexicon.

WHY UNDERSTANDING THE LANGUAGE–THOUGHT INTERFACE IS IMPORTANT FOR COGNITIVE SCIENCE

The language–thought interface is crucial to characterizing fully the human cognitive architecture and its operations at the most basic level. Through language, people can communicate visual, auditory, and haptic experiences, as well as feelings, beliefs, and theories. The language system must be connected to all of these systems of the brain. In addition, understanding the messages received from other people entails an interface between language and various systems in the brain going in the other direction. Some of the most fundamental questions about the cognitive architecture concern how information flows among these different areas and how information from one level or system becomes integrated with the output of another. The nature of these connections also bears on research in many specific domains of cognitive science. For instance, classic debates about the relation between external stimuli and perception and memory turn on ideas about how information from the various systems are brought to bear on the processing of the input. More recently, a key goal of models of language production has been to explain how people move from thoughts and feelings to the selection of words. Conversely, models of sentence processing aim to identify in detail how linguistic input is unpacked to create meaningful interpretations of the input.

Understanding how the systems interact will also reveal much about the contents of the mind. From the ancient Greeks to Hume and Kant in the eighteenth century to modern cognitive scientists, the question of where knowledge comes from has been debated. At one extreme is the idea that knowledge is acquired through the senses and is built from experience in the world; at the other extreme is the notion that knowledge could be heavily innate. Possibilities in between also have traction. Some knowledge might be inherent in the developing mind but can be realized only through input from the world. And even if knowledge is acquired through experience in the world, there must be some form of filtering or focusing of attention, because not all information encountered is encoded. In either of these cases, language may play a role: It might point out certain ways of interpreting experiences in the world or it might serve as a releasing factor that allows such knowledge to emerge and be connected to other parts of the conceptual system. Assuming that at least some knowledge is transmitted from generation to generation directly through language, it must also be recognized that languages are not entirely neutral with respect to the information they carry. As discussed in detail in the chapters of this book, every language reflects a certain perspective on the world through its inventory of words and encoding strategies. Because of these cross-linguistic differences, a message sent through one language will likely differ to some degree in meaning from the "same" message sent through another language. These differences could play a subtle but significant role in what the speakers of different languages learn. Thus, answers to what conceptual representations are like, how they are acquired, and what information they contain are intimately tied to understanding how linguistic and nonlinguistic systems are related.

FROM TRADITION TO A NEW TRAJECTORY IN CONSIDERING THE LANGUAGE–THOUGHT INTERFACE

The distinction between language and thought has long been recognized in many domains of investigation. In the domain of the lexicon, one approach to considering how knowledge of words relates to general knowledge about the world has been to think of the connection as akin to the relationship between a dictionary and an encyclopedia (Clark & Clark, 1977; Evans, Bergen, & Zinken, 2007). The mental lexicon is analogous to the dictionary: It specifies the pronunciation, syntactic characteristics, and meaning of a word, with meaning conceived of as limited in content (such as a set of defining features; e.g., Katz & Fodor, 1963). General world knowledge is analogous to the encyclopedia: It contains much broader knowledge about the world and links the limited content of word meanings to more elaborate associated knowledge. For instance, the meaning of the word *bachelor* might be merely "adult, unmarried male," but the encyclopedia adds that bachelors often like fast cars and parties. An alternative is to think of lexical knowledge as more like the encyclopedia itself. In this sort of view, word meanings encompass the broader knowledge of the world and may not differ in content from chunks of the encyclopedia (e.g., Murphy, 2002). Thus, the meaning of *bachelor* might include some notions of what typical bachelors like to do. A different version of an encyclopedic approach suggests that word meanings are not prepackaged (e.g., Clark, 1983; Evans et al., 2007); rather, words serve as prompts to construct meaning using general encyclopedic knowledge. The construction starts with elements of meaning conventionally associated with the word and factors in surrounding words, grammatical units, and nonlinguistic context to arrive at a more fully specified interpretation. In this view, there is no principled distinction either between semantics and pragmatics or between lexical semantics and encyclopedic knowledge. Although what counts as prestored lexical knowledge differs across these versions, all three share an assumption that encyclopedic knowledge comes packaged into coherent chunks and word meanings are closely aligned with these chunks.

But in the past decade, methodological, empirical, and theoretical advances have deepened the understanding of the nature of word meanings. Two elements of these advances in particular seem to demand new ways of thinking about the nature of the interface of these meanings with conceptual representations.

1. *Word meanings sometimes work against the correlational structure of the world.* One problem with the idea that word meanings are simply portions of, or pointers to, coherent chunks of encyclopedic knowledge is that word meanings sometimes appear to work against the way knowledge is organized in memory. According to many current approaches to memory organization, especially connectionist approaches, the general conceptual system is built up through statistical associations of experiences (Burgess & Graham, 1999; McClelland, 1994). If people were asked to organize the meanings of words into categories, they might group them in a way similar to how groceries are organized in a supermarket. For example, in grouping the meanings associated with verbs, people might separate cooking verbs (*fry, sauté, boil, bake, brown, wipe, clean, scrub,* etc.) from sports verbs (*tackle, defeat, hit, kick, knock down, run, swim, jump,* etc.) and gardening verbs (*grow, plant, plow, rake, shovel, spray, transplant,* etc.). The members of these groups tend to cooccur in situations that are psychologically salient and perceived as integrated events. However, this principle often does not seem to apply in the case of word meanings. Much work in lexical semantics has shown that certain components of meaning ("structural components," as we will discuss) tend to appear in the meaning of a wide range of words, leading to categories of word meaning that cross-cut the categories that might be expected to emerge on the basis of cooccurrence. For example, such components allow for classes of verbs entailing causation (e.g., *boil, bake, brown, defeat, knock down, grow, plow*), contact (e.g., *wipe, scrub, tackle, hit, kick, rake, shovel*), and manner (e.g., *fry, sauté, run, swim, jump, spray*), among others. These classes mix together cooking verbs with sports and gardening verbs. In short, as noted by Pinker (1989), word meanings in the lexicon imply categories that often do not seem to correspond to the kinds of groupings that people find cognitively useful or intuitive for storing their general knowledge of the world.

2. *Diversity in word meanings across languages is pervasive.* The difference between the general conceptual system and word meaning is further suggested by the extent to which word meanings vary across languages. In one sense, it has long been evident that languages differ in their inventories of word meanings. Because the development of vocabulary depends in part on the physical and cultural environments of a language community, languages tend to vary in how many distinctions within a domain are encoded in words. For instance, industrialized societies tend to have larger vocabularies describing color than traditional societies (e.g., Kay, Berlin, Maffi, & Merrifield, 1997). But a recent explosion of cross-linguistic research on word meaning has made evident much deeper diversity across languages in the meanings associated with words of a domain. The universalist idea that all languages make essentially the same distinctions, give or take granularity, is simply not correct. Languages differ markedly in how they partition by name many domains including color, space, body parts, motion, emotion, mental states, causality, and ordinary household containers. These differences in language are greater than can readily be motivated by differences in experience of the physical or cultural environment. For instance, language communities can differ in whether their verbs encode the manner or path of movement (e.g., Slobin, 1996; Talmy, 1985) even when their physical and cultural environments are rather similar

(as for English and Spanish speakers along the Texas/Mexico border). The linguistic diversity may simply reflect the dual facts that word meanings are highly selective in what elements of experience they encode and that because of this selectivity, there are many possible ways to map between words and the world. In the face of this diversity, short of accepting a priori that every linguistic difference is matched by a substantial difference in thought between language communities, the possibility that the pervasive cross-linguistic variability reveals a relatively loose fit between language and the underlying conceptual system must be taken seriously.

These two observations suggest that the relation of nonlinguistic content and word meaning is likely to be more complex than the traditional approaches have assumed. They help make clear why word meaning and encyclopedic knowledge need to be separated, and why, once this separation is acknowledged, the correct characterization of the interface between the two kinds of representation is likely to be nontrivial. They also make clear why the nature of the interface can be understood only by looking across languages. The kinds of mapping principles, structures, or processes that are postulated must accommodate not just one language, but the full extent of diversity that exists.

ADVANCES MEET OPPORTUNITY

For several years in the early 2000s, the two editors of this book met at conferences to chat about our shared interests in how languages encode meaning in words and to consider the implications of these new advances for the language–thought interface. In the course of our conversations, it struck us that it was time to bring researchers working on these and related fronts together for a larger discussion on words, thoughts, and their relation. In 2004, we began to pursue funding opportunities for holding a workshop, and in 2005 we were awarded a grant from the National Science Foundation. The workshop was held on June 6–7, 2005, at Lehigh University in Bethlehem, Pennsylvania. Fifteen speakers participated, representing cognitive and developmental psychology, linguistics, and anthropology, and covering work grounded in studies of languages from around the world. The workshop was attended by graduate students and faculty members from universities in five countries, representing fields that included education, communication disorders, and computer science as well as fields associated with the speakers. The enthusiasm with which our workshop announcement was met confirmed the timeliness of such a gathering, as did the liveliness of the meeting itself. Here, we present chapters contributed by 13 of the original workshop speakers, along with one focusing on insights from neuroscience to enrich the mix.

LESSONS FROM THE CHAPTERS

The chapters in this book offer a number of important lessons for thinking about aspects of the language–thought interface. Rather than provide a blow-by-blow account of each chapter, we highlight key themes that emerge across them.

Linguistic Diversity Occurs across Many Domains and in Many Forms

As we previously noted, the recent explosion of cross-linguistic investigation has revealed far more diversity in the content of word meanings across languages than was previously suspected. The chapters in this volume showcase the range of semantic domains in which diversity occurs—from the perceptual (color: Regier, Kay, Gilbert, & Ivry; Roberson & Hanley) to the abstract (causality: Wolff, Jeon, Klettke, & Li; mental states: Goddard; number: Gordon) to the very concrete (body parts: Majid; containers: Malt, Gennari, & Imai; toys: Clark) to terms for motion, direction, and spatial relations (Bohnemeyer; Kemmerer; Parish-Morris, Pruden, Ma, Hirsh-Pasek, & Golinkoff; Malt et al.), and to grammatical devices such as numeral classifiers (Imai & Saalbach; Lucy). They also make clear that diversity occurs across both open-class words including nouns and verbs and closed-class items such as classifiers and terms for spatial relations. Given the extent of documented diversity, it seems safe to project that there may be few or no domains of human experience in which the vocabulary words covering the domain map cleanly onto one another across languages. Furthermore, the chapters illustrate the variety of relations that can exist between the meanings encoded across languages for a domain—from cases in which languages may differ in granularity but otherwise are drawing similar distinctions (as may happen for some locomotion terms: Malt et al., or some body part terms: Majid), to cases in which terms vary in their boundaries but share lexical category centers (as argued by Regier et al. for color, though cf. Roberson & Hanley), in which there is more substantial cross-cutting of membership (e.g., container terms: Malt et al.), and in which the dimensions encoded from a domain are orthogonal (in some verbs of motion, as discussed by Bohnemeyer; Kemmerer; Parish-Morris et al; and Senghas; see also Slobin, 1996; Talmy, 1985).

As we previously suggested given the high degree of variation in how human experience is encoded into words, the differences in word meaning are likely to be greater than differences at the conceptual level. For instance, if all humans perceive certain dimensions of events involving animate agents such as their manner of movement and their path, then they must differ in how this perception comes to be mapped onto words because of the fact that some languages generally express manner in their verbs whereas others more commonly express path in the verbs (see chapters by Bohnemeyer; Kemmerer; Parish-Morris et al.; and Senghas). This possibility is explicitly evaluated in several of our chapters. If this possibility is right, then at the most basic level these chapters document why any general characterization of the human cognitive architecture that assumes a straightforward and universal mapping from conceptual representations to word meanings (albeit realized via different word forms) must be wrong.

Ways of Describing Word Meaning

To construct more accurate ideas about the mapping, it will be necessary to have good ways of describing the meanings that do exist. In the past, word meaning was regularly rendered in terms of other words, which raised concerns about the potential for circularity. Given the high degree of diversity in word meanings across languages, though, it is apparent that this method is also treacherous in another way. As discussed in Goddard's chapter, it may infuse the definitions with assumptions inherent in the meanings of one language, imposing them on elements of meaning from words in other languages in ways that are misleading at best. The chapters demonstrate several major improvements in ways of expressing meaning that can lead to a better understanding of the nature of similarities and differences across languages.

Chapters by Bohnemeyer, Kemmerer, Parish-Morris et al., Wolff et al., and Senghas all draw on recent advances in lexical semantics that distinguish between the *structural* and the *idiosyncratic* parts of meaning (Levin & Rappaport Hovav, 2009). The structural part specifies components that are significant to the grammar of a language and that comprise part of the meaning of a wide range of words. In the case of verbs, structural components include notions such as CAUSE, MANNER, CONTACT, ACT ON, CHANGE, and PATH (Jackendoff, 1990; Levin & Rappaport Hovav, 2009; Pinker, 1989; Talmy, 1985, among others). The idiosyncratic components distinguish among words with similar structural components. For example, many verbs of motion encode either a manner of motion (e. g., *walk, run, skip*) or a path of motion (e.g., *pass, arrive, enter*). Verbs encoding the manner of motion are assumed to specify the structural component of MANNER, and verbs encoding PATH are assumed to specify a structural component such as FROM or TO (Jackendoff, 1990). Among these two word classes, idiosyncratic components of meaning then indicate different manners or different paths (Jackendoff, 1990; Levin & Rappaport Hovav, 1992; Slobin, 1996; Talmy, 1985).

Similarly, words for spatial relations can be divided into a structural part, which specifies the abstract geometry of a spatial relation, and the more idiosyncratic part, which distinguishes spatial terms having the same underlying geometric characteristics (e.g., *over* and *above*) (Talmy, 2000; Landau & Jackendoff, 1993). By adopting this structural perspective, Bohnemeyer, Kemmerer, Parish-Morris et al., Senghas, and Wolff et al. are able to identify elements of meaning that may be appreciated nonlinguistically across speakers of all languages and then examine the varying ways in which they are encoded in words across different languages.

Other chapters illustrate other ways of describing elements of meaning that avoid reliance on the terms available in English or any other individual language. One approach is to limit the number of words that can be used in such definitions. Goddard provides an overview of the Natural Semantic Metalanguage approach to describing word meaning (e.g., Wierzbicka, 1996) in which limited numbers of semantic "primes" or primitives (irreducible elements of meaning thought to be universal) are deduced through experimentation with reductive paraphrase. He presents a detailed illustration, using terms of emotion and cognition, of how this approach can be used to reveal subtle differences in meaning among related words of different languages. Another approach is to express word meanings in terms of objective features of the world. The analysis of word meaning is thereby grounded in descriptions independent of the study of language or concepts. Majid uses the physical segmentation of the human body and Malt et al. use the biomechanics of human locomotion to provide a basis for understanding some shared tendencies across languages in the meanings encoded in these domains. Wolff et al. draw on the physics of force generation to help explicate the meanings encoded in verbs of causation across languages. Regier et al. and Roberson and Hanley evaluate color terminology against the background of the psychophysical understanding of color perception, and Regier et al. further add the use of Monte Carlo simulation to evaluate the extent to which color terminology across languages may be constrained by color perception. A final approach is to take advantage of new methodologies in neuroscience: Kemmerer's chapter introduces data from neuroimaging studies that demonstrate how the meanings activated by words engage regions of the brain overlapping those involved in the actual experience of their referents. Although his chapter is, of necessity, limited to data for English words, this methodology holds great promise for future cross-linguistic comparisons, as he notes.

Can Words Tell Us about Conceptual Representations?

Words as Pretenders The study of word meanings has sometimes been taken up not for its own sake, but as a means of illuminating the nature of thought itself. If the mapping from words to conceptual representations is neither simple nor universal, though, the view of language as a window into the mind (Chomsky, 1972), as applied to the lexicon (e.g., Lakoff, 1987; Pinker, 2007), is called into question. We have suggested that word meanings may be much sparser and more arbitrary than the experiences they encode for speakers (even though this property can be hard to recognize, as the properties that are not specified in the meaning of a word can be filled in by the general conceptual system when experiences are conveyed through language). Furthermore, because languages are handed down from generation to generation of speakers, some aspects of word meaning at any given time may reflect past influences rather than thought patterns of current speakers (as argued in the chapter by Malt et al.). Given these observations, the words of any single language, although convenient, cannot be counted on to reveal what any shared elements of human thought might be. As Bowerman (1996: 160) remarked in a discussion about the learning of spatial words, "I find it sobering that the 'non-linguistic spatial concepts' often hypothesized to underlie spatial prepositions—e.g., 'containment' and 'support'—lend themselves much more readily to shaping into the spatial categories of English than, say, of Tzeltal. In other words, ideas about plausible 'primitives' in the language of thought may themselves be conditioned by the language we have learned." Goddard's chapter elaborates on this issue for terms for mental states (emotions and cognitions) and demonstrates in some detail the potential fallacies involved in deriving a set of basic human emotions from an analysis of English terminology.

Words as Snapshots Although words of any single language may not provide a direct route into the mind, word meanings can be likened to photographs of three-dimensional objects from a single angle. By examining the meaning of words in multiple languages, it is possible to fill in parts of the shared, underlying understanding of a domain that were obscured from view because of the particular perspective of a given language. Many of the chapters, by looking at the encoding of domains across languages, provide this sort of three-dimensional perspective on what notions are commonly (if not universally) recognized across speakers of many different languages. For instance, Wolff et al. discuss underlying notions of causality; Majid demonstrates shared tendencies in the conceptualization of body parts, Regier et al. argue for shared aspects of color perception (though cf. Roberson and Hanley), Goddard does so for mental state words, Bohnemeyer, Malt et al., and Senghas all describe basic elements of motion events such as path and manner of motion that are encoded across languages, and Kemmerer and Parish-Morris et al. add consideration of elements of spatial relations to those of motion. Indeed, Senghas shows how deaf children exposed to only a rudimentary sign language as input have, over time, elaborated their language to encode some of the same elements found in spoken languages. The analyses in these chapters suggest that despite the striking cross-linguistic variation in what elements of thought are encoded into words in different languages, the diversity is not a reflection of free variation but rather is constrained by some shared tendencies in how speakers of different languages think about the world. Thus, looking across words of different languages may indeed provide insights into some important contents of the conceptual substrate.

These snapshots together might also reveal something about the derivation of conceptual content. It is tempting to suggest that commonalities reveal what is innate, and non-shared components of word meaning indicate what content is built up from experience in the world. But as discussed in the chapter by Malt et al., forces other than the existence of innate concepts may, in some combination, contribute to the existence of shared elements of word meaning. These forces include shared basic cognitive and perceptual capacities that might create special sensitivities to some distinctions among experiences; shared cultural needs, goals, and experiences; and shared exposure to salient discontinuities among entities that the world presents to the observer. Regier et al. suggest that the universal structure of perceptual color space makes some color naming systems preferable to others, implicating a direct impact of shared basic perceptual capacities. Roberson and Hanley, in contrast, make the case for similarities in color terminology stemming more from shared cultural needs, goals, and experiences. Parish-Morris et al. indicate that infants can discriminate certain spatial relations or components of motion events by about 5–7 months of age, but they do not form categories that include multiple instances of the relations or event components until somewhat later, implying shared perceptual and cognitive capacities that require maturation and perhaps sufficient experience with input to build the more complex content. The chapters by Majid and by Malt et al. provide examples of domains (body parts and locomotion) in which semantic commonalities seem to reflect the salient structure inherent in a stimulus domain, with the structure salient enough to observers across diverse cultures and languages to be frequently encoded in words. The discussion of force dynamics by Wolff et al. likewise implies a shared discrimination among distinct types of causal events that emerges from the laws of physics. In short, all three sources—pan-human sensory and cognitive mechanisms, needs, goals, and experience, plus the structure the world presents—may contribute to shared tendencies in word meaning. This sort of three-dimensional consideration of lexical encoding provides clues about the origins of some shared elements of nonlinguistic representations and so places constraints on theorizing.

The Perspective from Development If there were a simple mapping from words to concepts, possibilities for what the acquisition of both is like would be relatively constrained: Acquisition of one half of the mapping (a word or a concept) would bootstrap acquisition of the other, or else they would tune each other in some interactive fashion. But if adult speakers map from a substantially shared conceptual substrate to word meanings that are shared to a much lesser degree, the problem space of acquisition becomes more complex. Parish-Morris et al. contrast two possibilities for the relatively late mastery of "relational" words such as verbs and prepositions: First, children lack the conceptual foundations for acquiring the word meanings, and second, children have trouble establishing an appropriate mapping from conceptual elements to words. They show that preverbal infants can, in fact, form abstract representations of categories including spatial relations and forms of motion, giving the infant the necessary conceptual foundation for grasping the information to be encoded in words. The learner's problem appears to lie, instead, in constructing the right mapping from this knowledge onto word meanings, which sometimes requires suppressing the correlational structure of the world. This scenario raises the question of how a young child can accomplish such a feat. Clark's chapter indicates what the answer looks like: It can be accomplished by "offers" of words to the language learner that do not just present novel lexical items but that highlight specific contrasts in the semantic domain, helping the learner determine which features are bundled together to carve up that domain in the particular language being learned. The child faces a challenge in coordinating burgeoning knowledge of words and of the world, but adult input can provide scaffolding that allows the challenge to be met.

The Perspective from Neuroscience Neuroscience has begun to contribute insights about the nature of cognition in a variety of fields, and the language–thought interface is no exception. Chapters by Kemmerer, Regier et al., and Roberson and Hanley all provide information about the nature of the interface that has been revealed by methodologies in neuroscience. As pointed out by Regier et al. and Roberson and Hanley, judgments concerning the similarity of colors may involve both verbal and visual codes, and it is likely that the left hemisphere will be the locus of activation of a verbal code. Regier et al. provide evidence that color discrimination is faster for colors with different names (e.g., *blue* versus *green*) only when stimuli are presented in a lateralized fashion such that they are first processed in the left hemisphere. Roberson and Hanley discuss their own results and the results of others, including functional magnetic resonance imaging (fMRI) evidence, that are compatible with the idea that left hemisphere brain regions associated with language processing are actively associated with postperceptual processing of color. Consistent with the findings of Regier et al., Roberson and Hanley conclude that categorical perception effects for color do not reflect superior discrimination, per se, of colors when they cross a lexical boundary. Instead these effects may reflect the fact that decisions about color are hampered when perceptual and verbal codes conflict. Such findings indicate that describing the language–thought interface requires understanding not just the content relation of one representation to another, but also how activation of both types of information proceeds through the system to produce a behavioral output. The functional architecture of the brain, by determining what information is sent where and under what timing, will matter in what outputs are observed. Taking this fact into account can help illuminate what it means for language to influence thought.

In a different vein, Kemmerer's chapter shows how neuroimaging studies are uncovering intriguing similarities and differences in the neural networks associated with the naming of objects, colors, shapes, spatial relations, and motions, and the nonlinguistic processing of these stimuli. In particular, extracting meaning from words appears to activate regions of the brain that overlap or lie adjacent to the same regions that are engaged during the visual processing of their referents. These findings raise questions about how cross-linguistic variation in the encoding of thought in words may influence the development of brain structures. For instance, Kemmerer suggests that cross-linguistic differences in ways of talking about locations in space, which require sensitivity to different elements of the physical environment, might result in the differential development of certain neural systems. Although such speculation remains to be tested, if correct, it would provide evidence of a form of influence of language on nonlinguistic representations that is quite different from that described by Regier et al. and Roberson and Hanley. Thus, neuroscience is providing new types of evidence about processing of language and non-linguistic content that will help flesh out possible views of their interconnections.

The Answer to the Whorfian Question Is Not Just Yes or No In its most extreme form, the Whorfian hypothesis that language shapes thought implies that the mental lexicon and conceptual representations have a one-to-one relationship. But there are few researchers (including Whorf) who would argue for this extreme form (as Gordon's chapter points out). As we previously noted, this possibility seems untenable in light of current evidence about the nature of the lexicon and its relation to thought. As further discussed in the chapter by Malt et al., there are forces other than language that will contribute to how attention is allocated to aspects of experience, and it is unlikely that language blinds people to dimensions of experience not encoded in their own language. Thus, differences in word meanings are likely to be greater than differences at the conceptual level, ceteris paribus, and chapters by Imai and Saalbach and by Malt et al. provide data supporting this contention. Nevertheless, language may in some way influence its speakers' sensitivity to certain dimensions and have some influence on the way continuous dimensions of experience are partitioned. Chapters by Goddard, Gordon, Roberson and Hanley, Imai and Saalbach, Lucy, and Regier et al. all consider such possibilities. Although the specific conclusions differ, the authors are in agreement that research on this hypothesis needs to move beyond simple "Yes" or "No" answers. Imai and Saalbach, Roberson and Hanley, and Regier et al. indicate that for certain kinds of nonlinguistic tasks, differences in language do not lead to differences in performance, whereas for other nonlinguistic tasks, differences in language do seem to result in differences in performance. They raise some specific possibilities about how linguistic and nonlinguistic information may combine to produce different outcomes in different circumstances. It may also be possible, as argued in Gordon's chapter, that language is essential to building certain kinds of concepts. The Pirahã, a hunter–gatherer tribe in Lowland Amazonia, lack labels for exact quantities, and Gordon provides evidence that the Pirahã are unable to encode exact cardinalities below 10 (much less higher ones). He suggests that Pirahã performance cannot be explained simply by the tribe's environment or culture and thus implicates their lack of exact number words. The results from Gordon's studies are striking and, along with the studies discussed in the other chapters, suggest possibilities concerning how language may contribute to nonlinguistic thought. The discussions in these chapters make clear that explorations of when and how language might influence thought can help shed light on the nature of the interface more broadly. Working from the other direction, a better understanding of the nature of the interface will help clarify where the potential lies for language to influence nonlinguistic thought and performance.

The "Language" Part of the Language–Thought Interface Is Ultimately Not about Words Alone Although we have been talking about the interface from the perspective of words and their mapping onto knowledge of the world, Lucy's chapter makes a compelling argument that the interface is not just between the general conceptual system and individual word forms. It is between the general conceptual system and the meanings associated with various linguistic units, which include not only content words but grammatical roles, inflections, closed class terms, and so on. As Lucy points out, it is sometimes difficult even to determine what should count as a single lexical item in a language. For instance, English has a nondecomposable word, *boy*, that conveys the notion of male child, but in Yucatec, the same notion is conveyed by a compound consisting of two morphemes, one for male and one for child, each of which can stand alone in other contexts. In Spanish, it is conveyed with a gender affix attached to a stem, yielding *muchacho* (versus *muchacha* for a female child); neither affix nor stem can stand alone. Furthermore, Lucy notes, the meaning of a single lexical item derives in part from what it contrasts with in the semantic field in which it is embedded, and even what parts of related semantic fields it encompasses. Many of the other chapters implicitly illustrate this point by showing that the closest corresponding word meanings across languages only partially overlap in denotation or referential range, and Senghas' illustrations of how a signed language conveys elements of meaning using motion through space as well as hand shape raise additional complexities. The chapters by Majid and Malt et al. also make related points explicitly, and it is clear that the issues are ones with which researchers need to more actively grapple in analyzing cross-linguistic data. From this perspective, the title of this book, which stresses words, does not fully capture the nature of the interface.

The chapter by Wolff et al. illustrates the influence of syntax on semantics in detail. It notes that languages vary in the types of causal agents that can appear in the subject position in sentences of different languages. For instance, in English it is fine to say *The knife cut the bread* but in German, Korean, and (perhaps) Russian, it is not; knives cannot serve as the subject of a sentence with a causal verb such as *cut*. Taken at face value, this observation suggests that there must be some way in which the meaning of the words for knives or for cutting differs among the languages. But Wolff et al. go on to show that the kinds of entities that can appear as grammatical subjects in causal statements can be predicted by whether a language codes for grammatical relations through morphology or through word order. Languages that have a relatively fixed word order allow a greater range of entities as causal subjects; those that use case marking and hence have freer word order have a more restricted range. Ultimately, this pattern may derive from pragmatic needs: If it is preferable to position given information before new information (Clark & Clark, 1977), then fixed word-order languages such as English may need to allow greater flexibility in what can appear in the subject position because subjects must occur before objects. In view of this demonstration, it is clear that understanding the nature of cross-linguistic differences in how thought is mapped onto words requires considering the grammatical (and pragmatic) context in which the words appear.

Levels of Representation Once overly simplistic views of the language–thought interface are set aside, it becomes important to consider what kinds of representations may be involved in a more realistic account of the interface. The analysis of Parish-Morris et al. of the early conceptual foundations of word learning suggests thinking of the conceptual level of representation as containing primitives such as path, manner, support, and containment (e.g., Talmy, 2000) that are packaged into words in various combinations in different languages. Goddard's analysis of shared semantic primitives underlying cross-linguistically variable word meanings makes a similar suggestion, and Malt et al. are also sympathetic to this type of approach. Whereas Goddard suggests that the acquisition of language-specific word meanings may then create correspondingly language-specific concepts, Parish-Morris et al. and Malt et al. are less inclined toward this suggestion; in their accounts, conceptual elements are packaged together at a linguistic level. Bohnemeyer's chapter introduces a more complex set of representational distinctions, drawing on Jackendoff's (e.g., 1983, 1990) analysis in which reasoning and transfer of information between different systems are divided between a language-independent, noniconic, level of representation, termed Conceptual Structure, and another module of higher cognition, termed Spatial Structure, that encodes geometric properties in terms of image schemas. In light of evidence from English and Yucatec, Bohnemeyer concludes that cognitive representations of motion are comparable across languages at the level of Spatial Structure but not at the level of Conceptual Structure.

There are undoubtedly many other ways of thinking about what the architecture of the interface may look like. For instance, connectionist models provide an explicit account of architecture that is not represented in our chapters. Notably, though, connectionist models, such as that of Rogers and McClelland (2004), generally assume that lexical knowledge consists of associations of names with concepts that are acquired by building up connections in the network through interactions with the world. If words often package information in a way different from that given by statistical cooccurrence of experience in the world, as we have suggested, then it remains to be seen whether this sort of architecture can be adjusted to better capture the relations between lexical knowledge and conceptual representation.

TOWARD A MODEL OF THE LANGUAGE–THOUGHT INTERFACE

The insights provided in these chapters suggest the shape that a better specified model of the language–thought interface needs to take. The model needs to be able to discriminate and group experiences in the world before the onset of language and then begin to map words onto its grasp of the world using perceptual, social, and linguistic cues. It needs to be able to use such cues to accommodate a wide range of mappings of the nonlinguistic content to language depending on the specific language environment, and to take into account that the words are not learned in isolation but are part of a larger system, in which the meaning of each word derives in part from the role it plays in the system. It needs to embed an explanation of why the content packaged into the words is to some extent constrained across languages, and why the content still can vary and in what ways—specifying what the free parameters are and what fixes them for specific languages. It needs to account for how different systems handling nonlinguistic information interface with the mental lexicon, both when language is produced and when it is comprehended. It needs to be consistent with observations from neuroscience and take advantage of the observation that the interface may be multilayered; the interface does not have to be conceived of simply as a set of concepts that in a simple fashion is linked to a set of words. The model needs to explain when language will affect performance on nonlinguistic tasks and when not. In doing so, it can take into account the fact that such effects may be hemisphere-dependent and may depend on the speed and flow of information across elements of the system, not necessarily reflecting permanent changes to the conceptual representations.

The field is a long way from having anything close to what the model will ultimately need to look like, but the path toward it is becoming clearer. Following the workshop on which this book is based, an attendee e-mailed us from back home in Canada saying that she awoke with her head buzzing with new questions and ideas. We hope that this book will do the same for our readers.

CHAPTER ARRANGEMENT

Because the chapters address so many crosscutting themes, we have not attempted to place them into discrete sections in the book.

Instead, we have arranged them in a sequence that feels natural to us, though many other arrangements would have also been possible. We begin with foundational considerations of how words emerge from nonword representations and what the range of possible word-to-world mappings are. Next come chapters that discuss cross-linguistic universals and variation within one or a small set of related domains. Following them are chapters that focus more heavily on whether cross-linguistic variation has implications for thought itself. Several chapters then take up the developmental implications of this variation, ending with one that additionally critiques elements of the standard approach to understanding cross-linguistic variation. Finally, the book closes with a consideration of the neuropsychological underpinnings of word representation.

ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under Grant No. 0446538. We thank the National Science Foundation for funding the conference leading to this book and supporting the preparation of the book. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Contributions to the conference were also provided by Lehigh University's Psychology Department, College of Arts and Sciences, and Office of the Vice Provost for Research. Preparation of the book was also supported in part by an award from the University Research Committee of Emory University. We thank Carol Sabo-Berrian from Lehigh's Psychology Department for serving as administrative coordinator of the entire conference and Teri Loew for additional assistance. We also thank our editor at Oxford University Press, Catharine Carlin, for her enthusiasm for this project and her assistance in preparing the book.

The two editors contributed equally to the preparation of the book. The order of names on the volume and on this introduction is arbitrary.

References

Bowerman, M. (1996). The origins of children's spatial semantic categories: Cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145–176). Cambridge: Cambridge University Press.

Burgess, N., & Graham, H. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551–581.

Chomsky, N. (1972). *Language and mind* (Enlarged edition). New York: Harcourt Brace Jovanovich, Inc.

Clark, H. H. (1983). Making sense of nonce sense. In J. B. Flores D'Arcais & R. J. Jarvella (Eds.), *The process of language understanding* (pp. 297–331). London: Wiley.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt, Brace Jovanovich.

Evans, V., Bergen, B. K., & Zinken, J. (2007). The cognitive linguistics enterprise: An overview. In V. Evans, B. K. Bergen, & J. Zinken (Eds.), *The cognitive linguistics reader* (pp. 3–26). London: Equinox Publishing Co.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

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

Jackendoff, R. S. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Katz, J.J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39*, 170–210.

Kay, P., Berlin, B., Maffi, L., & Merrifield, W. (1997). Color naming across languages. In C. L. Hardin & L. Maffi (Eds.), *Color categories in language and thought* (pp. 21–56). Cambridge: Cambridge University Press.

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

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16*, 217–238.

Levin, B., & Rappaport Hovav, M. (1992). The lexical semantic of verbs of motion: The perspective from unaccusativity. In I. Roca (Ed.), *Semantic structure: Its role in grammar* (pp. 247–269). Berlin: Mouton de Gruyter.

Levin, B., & Rappaport Hovav, M. (2009). Lexical conceptual structure. In K. von Heusinger, C. Maienborn & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning*. Berlin: Mouton de Gruyter.

McClelland, J. L. (1994). The organization of memory: A parallel distributed processing perspective. *Revue Neurologique (Paris), 150, 8–9,* 570–579.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Pinker, S. (1994). *The language instinct*. New York: William Morrow and Company, Inc.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Penguin.

Rogers, T. T., and McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach.* Cambridge, MA: MIT Press.

Slobin, D. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195–220). Oxford: Clarendon Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Grammatical categories and the lexicon: Language typology and syntactic description* (pp. 57–149). Cambridge: Cambridge University Press.

Talmy, L. (2000). *Towards a cognitive semantics II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford: Oxford University Press.

1 REINVENTING THE WORD

Ann Senghas

Where do words come from? As we consider the thousands of words in our language today, we see many sources: modern words came from older words, words that came from other languages. But what if we wanted to start at the very beginning, to create a new language without predecessors to draw from? This chapter follows the beginning stages of a new language as some of its first words are being created. We ask whether words are the first part of the language to emerge, or whether they, in turn, come from something else. If so, what might that something else be?

After all, words do seem to be the natural place to begin in the task of building a new language. Once you have words, you could work on conventions for combining them into sentences. You could also use them as raw materials for other linguistic elements. In short, it seems that words are what you need to get the other parts of a language off the ground.

Linguists have traced the history of many words, and found such a path in their past. Through changes called grammaticalization, content words are retooled over time to serve as function words (Hopper & Traugott, 1993; Slobin, 2002; Traugott & Heine, 1991). For example, the word *will* in English used be an ordinary verb meaning "to want." It can still have that meaning today, but more often it is linked to another verb to mark future tense. So, in modern English, *Marie will run* does not mean that Marie wants to run; it means only that the running will take place in the future. Further along in time, function words can grammaticalize and become bound morphemes. For example, the verb ending *–ed* started out as the word *did*, and gradually reduced to less than a syllable that marks past tense (Swinton, 1882). Forms can become more and more reduced, and more and more abstract, until they disappear entirely from the language. Then the process begins again, with the language taking up other words to serve functions that were lost. Over centuries, as a language is passed down from one generation to the next, its words are constantly reshaped. Is this same reshaping process at work back before that first step, retooling nonwords into basic words?

To answer this question, we must consider the nature of nonwords. If we wanted to represent an event, and had no words available, we might turn to the nonlinguistic world for the means to build our expression. Accordingly, we could adopt the structure of the event itself into the structure of the expression. Aspects of the event that happen simultaneously would be represented simultaneously in the expression. Objects that are near each other physically in the world would be represented near each other in the expression. This would be an analog, iconic representation.

Of course, humans often package information in exactly this way. Nonlinguistic representations such as maps, paintings, and acted-out imitations of behaviors derive their structure iconically, directly from their referents. Patterns in the representation correspond, part-for-part, to patterns in the thing represented. In this way, half of a city map represents half of the city, and the initial moment of acting out a behavior represents the initial moment of the behavior. However, such representations are quite unlike language, which is linear, discrete, and combinatorial. The patterns in language are not the patterns in the world it represents. There is no part of New York City that corresponds to the word "York."

Even though such iconic, analog representations do not look like language, in this chapter I propose that they can be a starting point for language. Words do not come first. Rather, language begins with expressions, expressions that can initially draw their structure from the world. Mapping directly from their experiences, people first generate continuous, nondiscrete expressions to represent continuous, nondiscrete events. As these expressions are taken up by generations of new learners, they are separated into discrete elements that can be reassembled to form new, fully linguistic expressions. In this way, even though the original expressions may not be language-like, the words that derive from them are. In other words, we first make a not quite bona fide chicken and we extract a very real egg.

THE CASE OF NICARAGUAN SIGN LANGUAGE

A language emerging today allows us to observe this process as it is underway. Over the past 30 years, a sign language has been developing within the newly formed deaf community in Managua, Nicaragua (Kegl & Iwata, 1989). Because Nicaraguan Sign Language (or NSL) is so young, its originators are around today, able to show us its earliest forms. These deaf Nicaraguans provide us with the rare opportunity to follow a language as it is created from its raw beginnings.

NSL uses the visual-manual channel as its physical modality. This is the modality that is also used for gesture, and it can easily support an iconic, analog representation of the world. Indeed, when people are speaking, they often accompany their speech with gestures that iconically represent continuous, analog aspects of their message, like the weaving back and forth in someone's path of movement (Kita & Özyürek, 2003; McNeill, 1992). With a visual-manual channel, NSL has the raw materials to get the process started.

These two conditions—NSL's recent emergence and the use of the visual-manual modality—make the language ideal for observing iconic expressions and the creation of new words. They make a rare combination; neither condition holds in mature spoken languages.

Although this is an unusual situation, it is nevertheless a natural one. By natural, I mean that the language has arisen naturally within a population of people who are using it to communicate, and passing it on as they use it. The children who learn the language are learning it though the natural means of immersion and everyday use, not intentional construction and instruction. Nobody sat down and decided what NSL should be. Nor did anyone map out its grammar and teach it to deaf children, through lessons in a classroom. NSL came into existence naturally, using natural learning devices. It therefore gives us a special glimpse into the natural mechanisms with which humans learn and pass on language.

Indeed, the history of NSL is not unusual for a sign language.[1] Sign languages have emerged many times in the past, all over the world, when a critical mass of deaf people gathered in one place. This has occurred in close-knit communities with a high prevalence of deafness, or, in many places, with the formation of a new school or residence for deaf children (Groce, 1985; Johnson, 1991; Padden & Humphries, 1988; Sandler, Meir, Padden, & Aronoff, 2005).

For deaf Nicaraguans, the critical event was the opening of a new center for special education in 1977. Before that time, deaf children and adults rarely had the opportunity to meet (Polich, 1998; Senghas, 1997; R. J. Senghas, 2003). Societal attitudes kept most deaf individuals at home, and the few schools and clinics available served small numbers of children for short periods, with no contact outside school hours. Thus, deaf Nicaraguan children had minimal contact with each other, and no contact with deaf individuals older than themselves. It is quite clear, when you meet deaf adults who grew up before the 1970s, that the conditions then were insufficient for a sign language to develop. Very few of those in their late 40s and older socialize with other deaf people, and they do not have a sign language, even today.

From 1977 onward, rapidly expanding programs in special education brought deaf children and adolescents together in unprecedented numbers, and peer contact increased dramatically. Classes in deaf education, separate from other students, were offered from preschool through grade six. Initially, approximately 50 deaf students enrolled, increasing to 100 by 1979 (Polich, 2005). Although language instruction aimed to teach students to lip-read and to speak Spanish (with little success), the children spontaneously used gestures to communicate with each other, and they did so with enthusiasm whenever they had the chance—on the buses, during recess, even behind the teachers' backs in class. Soon a new language was born, a language they could use every day. In 1981, a vocational center for adolescents with disabilities opened, and many of the deaf teens who enrolled were graduates of the primary school. By 1983, the two centers served more than 400 deaf students (Polich, 2005).

In 1986, some parents and teachers set up a social club for deaf adolescents and adults. By 1990, this club had become a National Association of Deaf Nicaraguans, and was deaf-run. The Association actively promotes the use of sign language, and provides regular contact between deaf adults, adolescents, and children. These include social gatherings, vocational classes, deaf-coached athletic teams, and training for deaf adults to work as teaching assistants in the primary school classrooms.

Within the memory of today's adults, the lives of deaf Nicaraguans have changed dramatically, from essentially no deaf contact to a rich social world populated mostly by deaf people who sign. They visit each others' homes daily, celebrate major holidays together, and some have married each other and started families together. The social community continues to provide a fertile ground for the sign language. Every year, for the past 30 years, new waves of children have entered the school and the community, and learned to sign by socializing with slightly older peers. With this constant influx of learners, Nicaraguan Sign Language thrives. Today there are more than 1000 signers who have used it as their primary language throughout their lives.

SUCCESSIVE COHORTS OF SIGNERS

The version of NSL that individuals use depends on their age (Kegl, Senghas, & Coppola, 1999; Senghas, 1995). Those who arrived in the earliest years used a new, fledgling system of signing, whereas those who arrived more recently encountered a richer, more developed language. This changing linguistic environment has led to an inverted language community, in which the most sophisticated signers are the youngest ones. If we line everyone up, from oldest to youngest, we have a successive sequence of signing systems, revealing the historical progression of the language over time. Like archeologists, we can chart out the progression of NSL by sampling the signing of different age "strata," or cohorts, over its three decades.

Following this logic, in the studies described in this chapter, deaf participants are grouped according to the year that they joined the deaf community and learned to sign. For convenience, we have divided this continuum evenly into three cohorts. The first cohort includes those who arrived in the late-1970s and early-1980s; the second, those who arrived in the mid- to late-1980s; and the third, those who arrived since 1990. When the data were collected, the first-cohort signers were adults in their 20s, the second-cohort signers were adolescents, and the third-cohort signers were children. All of them learned to sign by the age of 6, and all of them have used NSL as their primary, daily language since then.

FROM ANALOG TO DIGITAL

In the present chapter, we will take a new look at two studies that have shown how NSL developed as it was passed down through these cohorts. What we consider here is how they specifically represent the transformation from continuous, holistic representations (un-wordlike) to discrete elements (word-like), in two very different domains. The first study considers the expressions used to describe motion events, like rolling down a hill, to see how the different aspects of the event are described (Senghas, Kita, & Özyürek, 2004; Senghas, Özyürek, & Kita, 2002, 2005).

The second study considers how signing space is used to indicate the participants in an event (A. Senghas, 2003; Senghas, Coppola, Newport, & Supalla, 1997).

STUDY 1: DESCRIBING MOTION EVENTS

Whenever an object moves from one location to another, it must move along some path. This can be a straight, direct line, a weaving, zigzag route, or even an arc through the air. It also must have some manner of movement: balls can bounce, people can walk, run, or even roll, and birds can fly. These two aspects of motion events occur simultaneously and are experienced holistically. When we roll down a hill, we experience the rolling and the descent all together; neither occurs before the other. If we wanted to represent such an event iconically, such as acting it out or demonstrating with a model, we would represent the manner and the path simultaneously. In this way, the expression would faithfully match the event in the world.

However, as has been shown in previous cross-linguistic research, languages do not work this way. Instead, they typically separate the expression of complex motion into separate elements that encode the manner and the path of the motion, and combine these elements according to the rules of the particular language (Talmy, 1985).

For example, English produces a verb to express manner (*rolling*) and a satellite element to express path (*down*), and assembles them tightly into the sequence *rolling down*. Japanese expresses manner and path elements with two separate verbs, as in *korogot-te ochiru* (rolling descends) (Kita & Özyürek, 2003). It was not known whether Nicaraguan signers would express motion iconically, in an analog manner faithful to the physical motion, or with the separate elements typical of developed languages.

To find out, we asked signers from each of the three cohorts to describe a collection of motion events, such as a cat climbing up a drainpipe or rolling down a hill. We also asked nondeaf Nicaraguans to describe the same events in Spanish, so that we could observe their co-speech gestures. Each participant watched an animated video cartoon that included these events, and narrated its story to a peer. Their narratives were videotaped, and the expressions that described the motion events were analyzed with respect to how the different aspects of the motion were integrated. Specifically, we determined whether information about manner and path was expressed simultaneously, as a single hand movement, or sequentially, as a string of manner-only and path-only elements. Examples of these two types of expression are shown in Figure 1.1.

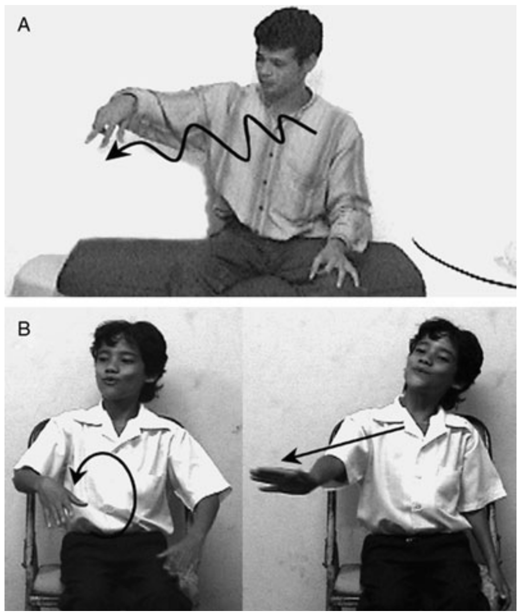


FIGURE 1.1. Examples of motion event expressions from participants' narratives. (A) Manner and path expressed simultaneously. This example shows a Spanish speaker describing a rolling event in Spanish; the gesture shown here naturally accompanies his speech. Here manner (wiggling) and path (trajectory to the speaker's right) are expressed together in a single holistic movement. (B) Manner and path expressed sequentially. This example shows a third-cohort signer describing the same rolling event in Nicaraguan Sign Language. Here manner (circling) and path (trajectory to the signer's right) are expressed in two separate signs, assembled into a sequence.

(From Senghas, Kita, & Özyürek, 2004.)

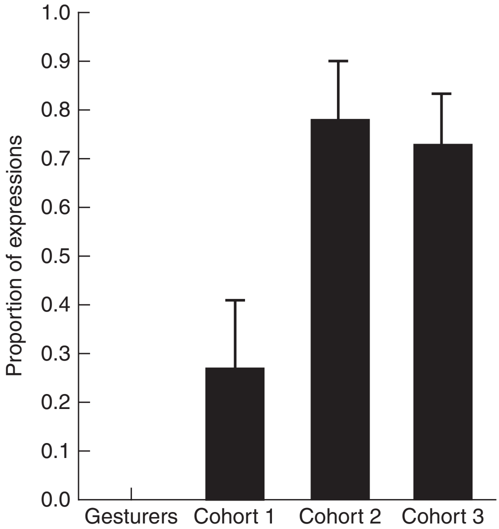


FIGURE 1.2. The proportion of expressions with manner and path in which they are produced sequentially as manner-only and path-only elements. Bars indicate mean proportions for individuals in each of the four groups; error bars indicate standard error. Such sequences are never observed in the co-speech gestures. First-cohort signers sometimes produce them; second- and third-cohort signers include them in most of their expressions.

(From Senghas, Kita, & Özyürek, 2004.)

Echoing the structure of the event, the gestures that accompanied the Spanish-speakers' narratives were continuous and analog in nature. Whenever the manner and path of the events were gestured, they occurred together in a single, holistic movement. As we move from the gesturers to the signers, we see a shift across the cohorts (see Fig. 1.2). The first-cohort signers, similar to the hearing gesturers, predominantly used holistic expressions, but also sometimes produced sequences of manner-only and path-only signs. By the second and third cohorts, this new, segmented type of expression had become the preferred one.

Although gesture offers the possibility of producing holistic expressions, children creating a language from gestures were inclined to treat manner and path as more basic, separable elements. Consequently, the temporal structure of NSL does not imitate the temporal structure of the events it represents. Instead, discreteness has been imposed on the temporal domain: Time, which is continuous, has been carved up into sequential, discrete, discontinuous moments. In the transformed expressions, learners are not using time to represent time. They are using time to order elements, and the order of elements is determined by other things in the language, not by the order in which things happened in the world.

STUDY 2: SPATIAL SIGNING: INDICATING THE PARTICIPANTS OF EVENTS

In the case of the previous motion event descriptions, there was a transformation in the temporal domain, from simultaneous to segmented and sequential. Sign languages have another domain that they can carve up—the spatial domain.

If you are familiar with only spoken languages, conceptualizing a spatial grammar can be challenging. Developed sign languages around the world actively use the three-dimensional space in front of the signer. Many signs are produced in their neutral form in a central location in front of the signer. By altering the direction of a sign's movement, a signer can alter its meaning. Thus, a sign produced with a movement away from the body has a different meaning than a sign that moves toward the body. A sign that moves from left to right has a different meaning than a sign that moves from right to left. The signing space serves a variety of grammatical functions, specific to different sign languages, including locative and temporal marking, indication of perspective, and anaphoric reference. In many sign languages, it is central to the system of verb-argument agreement.

The verb agreement use is a good example of a discrete use of the signing space. Here, the space includes some finite number of locations that are available to be incorporated into an utterance. By moving a sign toward or away from a specific location, the signer adds a spatial morpheme to the sign. For example, in American Sign Language (ASL), nouns are marked as definite and specific by being indexed to a particular location in front of the signer; verbs then "agree" with their noun arguments by taking on these same locations. An agreeing verb might begin at the location assigned to its subject, and move to the location assigned to its object (Adamo, Acuna, Cabrera, & Lattapiat, 1999; Clibbens, 1998; Fischer & Gough, 1978; Klima & Bellugi, 1979; Meier, 1987; Meir, 1998; Padden, 1983; Supalla, 1982). [For an alternative to this discrete analysis of such constructions, see Liddell (2000).]

Other uses of the signing space are more iconic and continuous. One such use is locative marking, used in descriptions of physical spatial relations. For example, to describe objects arranged in a circle, the signs representing the objects would be produced in a circular arrangement in the signing space (Emmorey, 1996; Klima & Bellugi, 1979). Although much of the structure of such utterances are drawn from the structure of the world, even this use of space is not automatic or unambiguous (Emmorey & Reilly, 1995). Spatial signs can be interpreted only relative to other signs. A single movement may be simultaneously to the north, toward the door, to the right of the signer, or to an indexed location. The interlocutor must be able to identify which interpretation of the movement is intended, including whether the signing space at that moment represents an analog continuum or a collection of discrete locations. Thus, for spatial signing to be effective, the language must include devices that convey how space is being used in a particular utterance.

The emergence of spatial devices offers another opportunity to explore the relationship between continuous and discrete uses of space. The more analog, locative use of space offers a possible precursor to the more discrete use of space for indicating argument relations previously discussed. Many of the movements used in indicating arguments appear to be metaphorically linked to physical spatial relations (Taub, 2001). For example, the ASL sign GIVE includes a movement toward a location associated with the recipient.

Although deaf Nicaraguans in the 1970s were never exposed to a developed grammar that included spatial devices, it is likely that a creative use of space to express relations between concepts was present in the gestural environment. Hearing people exhibit fundamental aspects of a spatial system when they use their hands to describe objects in unusual positions, or moving along various paths (Singleton, Goldin-Meadow, & McNeill, 1995). The rudimentary "home sign" systems that emerge in families with a deaf child also can make use of the signing space, such as producing signs in particular locations to represent relationships among people, objects, and actions (Coppola, 2002; Goldin-Meadow & Mylander, 1990). Thus, some spatial resources were certainly available, and were good candidate raw materials for deaf Nicaraguans to develop a more complex system for representing grammatical relations.

In various studies we have explored whether spatial modifications to signs were used to indicate the participants in an event (Kegl et al., 1999; A. Senghas, 2003; Senghas & Coppola, 2001; Senghas et al., 1997; Senghas, Senghas, & Pyers, 2005). That is, is the signing space used in a way that shows who did what to whom? For example, in a sentence that describes a man giving something to a woman, do signers use space to link the referents *man* and *woman* to the roles of giving and receiving?

One production study specifically compared the first and second cohorts' use of space in this way (Senghas et al., 1997). Participants were shown videos depicting brief events, such as a man giving a cup to a woman, a woman giving a cup to a man, a man taking a ball from a woman, a woman looking at a man, etc. They are then asked to produce a simple sentence in NSL describing each event.

We expected signs to move toward and away from the body in ways that metaphorically matched the events, for example, GIVE moving away from the body and TAKE moving toward the body. But we were interested in whether the signing space was differentiated further than this. Does signing to the side differ in meaning from producing a sign neutrally? And even more specifically, once you have a discourse context with certain referents already established, would there be a difference in meaning between signing to the left and signing to the right? Do such spatial movements distinguish referents from each other and link them to verbs, tagged with their respective roles?

Our analyses revealed that only second-cohort signers produce movements toward nonneutral locations in a pattern that corresponds to the role of a referent. For example, the production of a movement to the right or left corresponded to the recipient of a giving event, or the object of a looking event. Examples of this use of spatial modulation are shown in Figure 1.3. Here we see the Nicaraguan verbs SEE and PAY, produced with respect to neutral and nonneutral locations. In the nonneutral cases given in the right panel of the figure, the verbs' shared direction would indicate their link to a shared object. When produced by second-cohort signers, this sequence would mean that a single person was both seen and paid.

The older, first-cohort signers were not constrained in this way. They could produce signs in either direction, regardless of the role of a referent. For them, the construction in the right panel of Figure 1.3 would be equivalent to the left panel. Neither would be specified with regard to whether a single person was seen and paid, or one person seen and another paid. In this way, first-cohort expressions were less restricted in meaning than second-cohort expressions.

The same generational difference was observed even more acutely in a comprehension task comparing members of the first two cohorts (see Fig. 1.4) (A. Senghas, 2003). Signers were presented with video clips of signed sentences that had been collected in the production task, and asked to point at pictures to indicate what these sentences could mean. First-cohort signers gave an unrestricted interpretation; they did not limit their interpretation of the object of the verb (such as the recipient) based on the direction of the sign's movement. Accordingly, the same movement could mean a man giving to a woman or a woman giving to a man. (These first-cohort signers identified such roles using word order, rather than differentiating the signing space.) In contrast, second-cohort signers used the spatial aspects of a verb to restrict their interpretations of referents' roles. Right and left were not equivalent, and the interpretation of roles was restricted based on the direction of movement. This pattern of responses suggests that the differentiated use of space emerged in the mid-1980s, among those who were still children at the time.

With such a substantial change in such a short period, coexisting members of the signing community produce and understand signing differently from each other. Indeed, many of the utterances produced by the older signers would be ungrammatical to a younger signer. Midway through the comprehension task, participants were asked explicitly whether the direction of the movement in a verb, to the left or right, affected their interpretation. First-cohort participants all responded no: a verb could be signed to the left or right without changing the meaning of the sentence. In contrast, second-cohort participants all responded yes: the direction in which the verb was produced did make a difference.

FIGURE 1.3. The left panel shows the signs SEE and PAY produced neutrally. The right panel shows the signs SEE and PAY produced to the left.

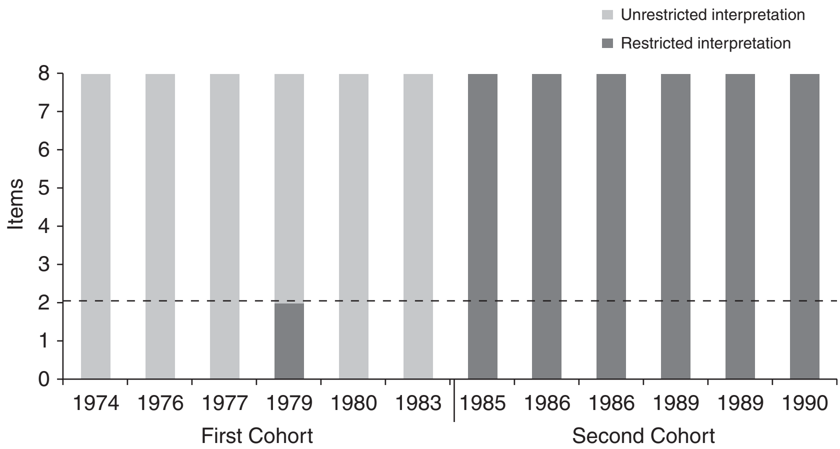


FIGURE 1.4. Interpretations accepted by first- and second-cohort signers for sentences that include a spatially modulated verb, ordered by the date of each participant's entry into the community. Chance level (excluding distracters, which were never selected) is indicated with a broken line. This measure shows a qualitative difference between pre- and post-1984 versions of Nicaraguan Sign Language. First-cohort participants, who entered the community prior to 1984, did not restrict their interpretation of sentences based on whether the verb was produced with a movement to the right or left. Second-cohort participants, who entered after 1984, gave the same sentences a narrower, more specific interpretation, always restricting the object of the verb based on this directional movement.
(Adapted from Senghas, 2003).

Among these younger signers, how much it mattered depended on the signer's age. For example, when asked about a specific sentence that was signed with a directional verb, a 16-year-old and a 15-year-old participant both acknowledged that others would accept a more unrestricted, general interpretation, but that they preferred the more restricted one. A 12.5-year-old participant responded that older signers will often sign first in one direction and then later in another, since they "can't always remember" in which direction they previously produced a sign; that is, they lose track of their referents. The youngest participant, a 12-year-old signer, insisted that the restricted interpretation was the only acceptable option. When presented with an unrestricted interpretation, he rejected it soundly. "If you wanted to say that, you would sign it here on the other side," he replied, rolling his eyes at the obviousness of it. And yet, such sentences are produced today, with that unrestricted meaning, by signers only 10 years older than him. Younger signers reject the very language that served as their input when they first learned to sign.

Because it did not appear in first-cohort signing, contrastive use of locations in the signing space does not appear to have been brought over directly from analog, gestural representations. Instead, there was a period in between during which meaningful spatial contrasts were stripped away from signed expressions. Thus, the first stage of development beyond gestural expressions involved expressions that were (laterally) spatially neutral. Although movement away from the body contrasted with movement toward the body, movement to the left did not contrast with movement to the right. This was the first cycle of differentiation and separation, and in the process, some of the spatial form found in gesture was lost, along with its corresponding meaning.

The first cohort produced this spatially neutral version of NSL for its first 6 or 7 years. Then, as a new, second cohort of children took up the language, another cycle of differentiation occurred, this time a differentiation within the signing space. Movements to different sides of the body became contrastive. Locations on different sides of the body thus became available to be mapped onto the various participants of events, in a way that can be motivated by their relative locations in the world. In a twist of nature, this second iteration of the language thus added some iconically expressed elements, making it somewhat more like the spatial events being represented, and less like the spatially neutral stage of the language that preceded it. Nevertheless, this newest version of NSL does not represent a return to the holistic, mimetic expressions of its origins. Instead, the continuous signing space has been carved up into spatial locations that can be treated as linguistic elements separate from the movement of any one particular sign. These spatial elements can be recombined with other verbs and nouns to link them together. A continuous, homogeneous signing space has been reanalyzed into recombinable components that indicate who is doing what to whom.

PUTTING IT ALL TOGETHER

The linguistic territories involved in these two examples of language change could not be more different. One involves the expression of motion and the other involves the participants of events. Yet over a few decades they have undergone similar changes. With motion expressions, signing that started out as holistic and continuous was temporally differentiated into manner and path elements. That is, expressions were carved up to create words. With participants of events, space that was homogeneous and neutral was spatially differentiated into contrastive locations, bound to the endpoint of a sign. That is, the ends of expressions were carved off to create bound morphemes.

With these changes, the sign language has become less like gesture. On the surface, signing still has many characteristics in common with the gestures that accompany speech. The movements of the hands and body in NSL are undoubtedly derived from everyday gesture. But these closer analyses have revealed a qualitative difference in the fundamental structure that underlies signing and gesturing. Although sign, like gesture, offers the possibility of producing holistic, continuous expressions, the creators of NSL did not take full advantage of that possibility. They did not faithfully reproduce the gestural expressions being used around them. Instead, they broke gestural expressions apart to form elemental components. These elements could then be reassembled to form new complex constructions. As second and third cohorts learned the language in the mid-1980s and 1990s, the new forms and expressions took hold, and the more gestural expressions faded from use. In the end, the structure of NSL no longer imitates the continuous structure of the world it represents, the way gesture does. Competing constructions have supplanted faithful representations.

We cannot know with certainty what factors led to these changes, but we can make some educated guesses about the most significant ones. First, the language had its origins in gesture, a medium that can be directly shaped into analog, iconic representations. With this rich medium, a certain amount of communication could be successful even before many conventions had been established. Individuals could create meaningful representations on the fly, adding a variety of expressions to the common repertoire, ready to be taken up by others and reshaped as they are reused. Second, there was a constant flow of individuals eager to do that reshaping and reusing. Crucially, these individuals included children, followed by yet more children, taking the language up, and then passing it on in adolescence and adulthood. Evidently, the differentiation we observe was the natural outcome when children were presented with continuous, noncombinatorial representations.

Of course, child learners do not typically change a language this much. They do not typically reshape gesture either, usually allowing it to retain its analog, continuous nature generation after generation. Why would Nicaraguan signers do any reshaping, when a serviceable gestural model was available to them? Perhaps some other characteristics emerging in Nicaraguan Sign Language made the elemental, combinatorial constructions preferred. On the other hand, perhaps such constructions were preferred because the children doing the reanalysis were engaged in the task of learning a native language.

It has long been noted that early child language looks quite unlike its model. Of course, children's very first utterances are holistic words spoken in isolation, such as *more, up,* and *bye-bye* (Brown, 1973). These expressions are not combinatorial; they cannot be broken down into smaller components of meaning. But children do not stick to such unanalyzed forms for long. To break into the combinatorial grammar of their language, they must analyze an apparently continuous stream of language into more basic elements. This process is evident in errors that children make when they occasionally produce such elements where they do not belong. For example, children commonly stick the regular past-tense ending *–ed* where adults would use an irregular past-tense verb—they don't need to hear another child make the same mistake to create novel forms such as *breaked* and *goed* (Chomsky, 1959; Lenneberg, 1964; Pinker, 1999). This process of analysis is also evident when children create new complex forms by taking basic elements and appending them to nonsense words, as in *pilked* and *wugs* (Berko, 1958; Marcus et al., 1992). These nonsense words are not found in the speech of adults, but their final elements are, available to be extracted by an adept learning mechanism.

Children learning sign language must actively break down the spatial as well as the temporal nature of the constructions they see. The spatial analysis takes some time, and is acquired later than temporal features such as word order (Newport & Meier, 1985; Supalla, 1982). You can see, in children's productions, how they pull out the elements embedded in the complex signal. For example, children learning American Sign Language are exposed to spatially complex signs, such as a verb that moves through the signing space from one specific location to another. However, in their own production, they initially break expressions down into sequential strings of basic elements, such as a verb in neutral space followed by a point to a location. Only later do they begin to combine these elements into a single complex movement (Meier, 1987), eventually converging on an adult-like form.

This is where NSL differs from the typical case of language learning. There was no adult signing with which to converge; instead, children converged with members of their own cohort, as a group. Because the language is so young, and because its creators were children, NSL may still prominently include the consequences of the learning strategies available during early language acquisition. That is, there are two complementary sources of the characteristics of Nicaraguan signing. On the one hand, it is very close to its gestural roots. On the other hand, it has recently passed through several iterations of native child language learning. Each generation of learners left its signature on the language as it altered its structure.

By examining this signature, we can reverse-engineer some of the natural inclinations of child learners. Evident here is the inclination to analyze a linguistic signal as discrete and combinatorial, even if it is originally continuous or holistic. Though clay goes in, bricks come out, bricks that can then be assembled into completely new expressions.

If we can imagine this learning process being repeated over many generations, across many learning mechanisms, we see how universal characteristics in languages might arise. For universal characteristics to be the result of learning, they need not correspond, one-for-one, to learning mechanisms. They need only be the result of the repeated application and interaction of many mechanisms. For example, de Saussure (1916) argued that the relationship between a word and its meaning is always arbitrary. Where does this arbitrariness come from, if people are capable of learning nonarbitrary, iconic symbols? I would suggest that arbitrariness is the product of sending words repeatedly through the child-learning mill. The products of human language-learning mechanisms must be made up of basic elements (a trait we have explored at length here), and they must be hierarchically organized, and pronounceable, and distinguishable from other symbols (traits we have not explored here)—and such traits compete directly with analog representation, iconicity, indeed, with any links between form and meaning. With generations of reanalysis and differentiation, iconicity is typically lost, and an arbitrary word remains. The "universal" characteristic of arbitrariness is thus an artifact of acquisition.

So this is where Nicaraguan Sign Language got its words. It is no chicken-and-egg puzzle. In the beginning, there was the expression, continuous and analog and holistic. The world gave this expression its rich, un-language-like form. Then the expression begat the word, pure and simple.

ACKNOWLEDGMENTS

Research described in this paper was supported by the Language and Cognition Group at the Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, NIH grant R01-DC00167 to Elissa Newport, and NIH/NIDCD grant R01-DC05407 to A.S. The content is solely the responsibility of the author and does not necessarily reflect the official views of the funding institutions.

Note

1. It is also not atypical for Nicaragua to have its own sign language. Readers may be surprised to learn that "sign language" is not a single, universal system, any more than "spoken language" is. Sign languages are languages in their own right, and vary from one geographic region to another, just as spoken languages do. They are not alternate encodings of spoken languages (unlike systems such as Braille and Morse Code) and will differ even between places that share a spoken language. For example, even though British Sign Language, American Sign Language, and Australian Sign Language (Auslan) are all used in regions in which English is spoken, they are mutually incomprehensible, separate languages, each with its own grammar and lexicon.

References

Adamo, D., Acuna, X., Cabrera, I., & Lattapiat, P. (1999). Verbos espaciales locativos en la lengua de señas chilena. *Revista de lingüística teórica y aplicada, 37*, 7–21.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150–177.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Chomsky, N. (1959). A review of B. F. Skinner's "Verbal Behavior." *Language, 35*(1), 26–58.

Clibbens, J. (1998). The development of spatial reference in British Sign Language. *Deafness and Education, 22*(3), 18–23.

Coppola, M. (2002). *The emergence of grammatical categories in homesign: Evidence from family-based gesture systems in Nicaragua.* Unpublished Ph.D. dissertation, University of Rochester, Rochester, NY.

de Saussure, F. (1916). *Course in general linguistics* (W. Baskin, Trans., 1966 ed.). New York: McGraw-Hill Book Company.

Emmorey, K. (1996). The confluence of space and language in signed languages. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 171–209). Cambridge, MA: MIT Press.

Emmorey, K., & Reilly, J. (1995). *Language, gesture, and space.* Hillsdale, NJ: Erlbaum.

Fischer, S., & Gough, B. (1978). Verbs in American Sign Language. *Sign Language Studies, 18*, 14–48.

Goldin-Meadow, S., & Mylander, C. (1990). Beyond the input given: The child's role in the acquisition of language. *Language, 66*(2), 323–355.

Groce, N. E. (1985). *Everyone here spoke sign language: Hereditary deafness on Martha's Vineyard.* Cambridge, MA: Harvard University Press.

Hopper, P. J., & Traugott, E. C. (1993). *Grammaticalization*. Cambridge, UK: Cambridge University Press.

Johnson, R. E. (1991). Sign language, culture and community in a traditional Yucatec Maya village. *Sign Language Studies, 73*, 461–474.

Kegl, J., & Iwata, G. (1989). Lenguaje de Signos Nicaragüense: A pidgin sheds light on the "creole?" ASL. In R. Carlson, S. DeLancey, S. Gilden, D. Payne, & A. Saxena (Eds.), *Fourth Annual Meeting of the Pacific Linguistics Conference* (pp. 266–294), Eugene, OR.

Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony, and development* (pp. 179–237). Cambridge, MA: MIT Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*(1), 16–32.

Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Lenneberg, E. H. (1964). The capacity for language acquisition. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall.

Liddell, S. K. (2000). Indicating verbs and pronouns: Pointing away from agreement. In K. Emmorey & H. Lane (Eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima* (pp. 303–320). Mahwah, NJ: Erlbaum.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago Press.

Meier, R. P. (1987). Elicited imitation of verb agreement in American Sign Language—Iconically or morphologically determined? *Journal of Memory and Language, 26*(3), 362–376.

Meir, I. (1998). Syntactic-semantic interaction in Israeli Sign Language verbs: The case of backwards verbs. *Sign Language & Linguistics, 1*(1), 3–37.

Newport, E. L., & Meier, R. P. (1985). The acquisition of American Sign Language. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 1: The data* (pp. 881–938). Hillsdale, NJ: Lawrence Erlbaum Associates.

Padden, C. (1983). *Interaction of morphology and syntax in American Sign Language.* Unpublished Ph.D. dissertation, University of California, San Diego, CA.

Padden, C., & Humphries, T. (1988). *Deaf in America.* Cambridge, MA: Harvard University Press.

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Polich, L. (1998). *Social agency and deaf communities: A Nicaraguan case study.* Unpublished Ph.D. dissertation, University of Texas at Austin.

Polich, L. (2005). *The emergence of the deaf community in Nicaragua: "With sign language you can learn so much."* Washington, DC: Gallaudet University Press.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *PNAS, 102*(7), 2661–2665.

Senghas, A. (1995). *Children's contribution to the birth of Nicaraguan Sign Language.* Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development, 18*(4), 511–531.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science, 12*(4), 323–328.

Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. *Proceedings of the 21st Boston University Conference on Language Development*, 550–561.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging Sign Language in Nicaragua. *Science, 305*(5691), 1779–1782.

Senghas, A., Özyürek, A., & Kita, S. (2002). Encoding motion events in an emerging sign language: From Nicaraguan gestures to Nicaraguan signs. In A. Baker, B. van den Bogaerde, & O. Crasborn (Eds.), *Cross-linguistic perspectives in sign language research. Selected papers from TISLR 2000.* Hamburg: Signum Press.

Senghas, A., Özyürek, A., & Kita, S. (2005). Language emergence in vitro or in vivo? Response to comment on "Children creating core properties of language: Evidence from an emerging sign language in Nicaragua." *Science, 309* (5731), 56.

Senghas, R. J. (1997). *An "unspeakable, unwriteable" language: Deaf identity, language & personhood among the first cohorts of Nicaraguan signers.* Unpublished Ph.D. dissertation, University of Rochester, Rochester, NY.

Senghas, R. J. (2003). New ways to be deaf in Nicaragua: Changes in language, personhood, and community. In L. Monaghan, K. Nakamura, C. Schmaling, & G. H. Turner (Eds.), *Many ways to be deaf: International, linguistic, and sociocultural variation* (pp. 260–282). Washington, DC: Gallaudet University Press.

Senghas, R. J., Senghas, A., & Pyers, J. E. (2005). The emergence of Nicaraguan Sign Language: Questions of development, acquisition, and evolution. In J. Langer, C. Milbrath, & S. T. Parker (Eds.), *Biology and Knowledge revisited: From neurogenesis to psychogenesis* (pp. 287–306). Mahweh, NJ: Lawrence Erlbaum.

Singleton, J. L., Goldin-Meadow, S., & McNeill, D. (1995). The cataclysmic break between gesticulation and sign: Evidence against a unified continuum of gestural communication. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 287–311). Hillsdale, NJ: Erlbaum.

Slobin, D. (2002). From ontogenesis to phylogenesis: What can child language tell us about language acquisition? In J. Langer, S. T. Parker, & C. Milbrath (Eds.), *Biology and knowledge revisited: From neurogenesis to psychogenesis*. Mahweh, NJ: Lawrence Erlbaum.

Supalla, T. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language.* Unpublished Ph.D. dissertation, University of California, San Diego, San Diego, CA.

Swinton, W. (1882). *A grammar containing the etymology and syntax of the English language*. New York: Harper & Brothers.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Grammatical categories and the lexicon, Vol. III* (pp. 57–149). Cambridge, UK: Cambridge University Press.

Taub, S. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge, UK: Cambridge University Press.

Traugott, E. C., & Heine, B. (1991). *Approaches to grammaticalization*. Philadelphia: J. Benjamins.

2 LEXICALIZATION PATTERNS AND THE WORLD-TO-WORDS MAPPING

Barbara C. Malt, Silvia Gennari, & Mutsumi Imai

Knowledge of words and knowledge of the world must somehow be linked. Words evoke knowledge about the world, and thoughts about the world are conveyed through words. The precise nature of this linkage is far from known, however. Our goal in this chapter is to shed light on the connection. In the first part of the chapter, we consider arguments for a tight mapping between words and conceptual representations and discuss reasons why these arguments are not entirely convincing. We also briefly consider and dismiss the extreme alternative that there is only the loosest relation between words and conceptual representations of the world. In the second part of the chapter, we turn our attention to a third alternative that we call a "constrained but flexible" mapping. In this section we review data indicating that in at least some domains, the mapping between words and conceptual representations is not tight. We consider the ways in which the mappings may nevertheless be constrained, and we discuss where flexibility is possible despite the constraints. We present data from two studies on the naming of human locomotion that test the ideas about both where the mapping is constrained and where flexibility may enter the picture. In a final section, we discuss implications of the "constrained but flexible" idea for several associated issues.

THE TIGHT AND LOOSE MAPPING POSSIBILITIES

Tight Mapping

A widely accepted view of the language-thought linkage is that words map closely onto coherent packets of nonlinguistic knowledge constituting concepts. Under such a view, word meanings consist of concepts (e.g., Bloom, 2000; Markman, 1989; Murphy, 2002), and prelinguistic concepts may provide the basis for word learning (e.g., Carey, 2001; Clark, 1983, 2004; Landau & Jackendoff, 1993; Nelson, 1974). There are several different scenarios under which such a tight mapping could come about.

Universal Concepts Lead to Universal Word Meanings One simple scenario by which words and conceptual knowledge could be tightly linked is if human thought is grounded in a large stock of universal concepts. Such universal concepts might come about via any of several mechanisms (or a combination of them): Humans might all be driven by a set of shared needs, goals, feelings, etc., across cultures. Their presumably pan-human basic cognitive and perceptual capacities might segment the world for them in comparable ways by creating special sensitivities to some distinctions. And the world might present itself to the human observed packaged in chunks so salient that given shared perceptual and cognitive capacities, all humans recognize these chunks. These sources of uniformity of thought across cultures could even lead, over the course of evolutionary history, to the large stock of shared concepts being innate (Fodor, 1975). Regardless of origin, if humans across all cultures have largely the same concepts, then their languages may all develop words with parallel meanings to express these notions. This possibility is consistent with (although not required by) the idea that prelinguistic concepts provide the basis for early word learning. If infants universally share certain prelinguistic notions about the world, the word-learning process might help shape meanings that are shared across languages. It is also consistent with evidence in some domains for shared tendencies in patterns of naming across languages. Shared elements of naming patterns have been found in domains including color (Kay, Berlin, Maffi, & Merrifield, 1997), body parts (Majid, Enfield, & van Staden, 2006), cutting and breaking actions (Majid, Bowerman, van Staden, & Boster, 2007), and mental states (Goddard, this volume) and these commonalities occur to a greater extent than would be expected by chance (Kay & Regier, 2003).

But it is by now evident that cross-culturally shared concepts leading to shared meanings can account for only a limited portion of world-to-word mappings at best. A striking finding of recent cross-linguistic research has been pervasive differences in how speakers of different languages talk about the world. Languages show many differences as well as commonalities in how they divide up domains by name including color (see Roberson & Hanley, this volume), space (e.g., Bowerman, 1996; Levinson, 1997), motion (Talmy, 1985; Slobin, 1996a,b), emotion, kin relations, and mental states (Goddard, this volume; Wierzbicka, 1992), causality (Wolff, Jeon, Klettke, & Li, this volume), and even ordinary household containers (e.g., Malt, Sloman, Gennari, Shi, & Wang, 1999). Detailed illustrations of this sort of diversity are provided in many of the chapters of this book. This well-documented and widespread cross-linguistic diversity indicates that the situation must be more complex than implied by strong universals of human thought coupled with a direct causal link from thought to word meanings.

Culturally Variable Concepts Lead to Culturally Variable Word Meanings An alternative scenario yielding a tight mapping between words and conceptual representations, yet taking into account the widespread cross-linguistic diversity, would be to posit that members of different cultures form at least partially different concepts, leading them to develop differences in the meanings attached to words. Thus, concepts are the causal agent driving the associated languages to adopt different patterns of naming.

Under this solution, we need to ask what would have led to the different concepts that drive the linguistic differences. An obvious answer is different cultural needs, interests, ecologies, and experiences. Although the classic example of Eskimos having a highly developed set of lexical distinctions for types of snow may be apocryphal (Pullum, 1991), other examples along such lines are more grounded in reality. For instance, members of nonindustrialized cultures may experience and need to discriminate among relatively few color variations. And, in fact, members of nonindustrialized cultures do tend to have fewer color terms than members of industrialized cultures (Kay et al., 1997). Similarly, wine experts have elaborate vocabularies for wine (Lehrer, 1983). Thus it is not implausible to assume that the importance of a domain to members of a culture will tend to influence their lexicalized distinctions.

However, this answer fails to provide an explanation for many observed cross-linguistic lexical differences. Many differences that have been the focus of recent interest are not readily related to specific cultural needs, interests, ecologies, or experiences. What common aspects of culture would lead English, Russian, and Chinese speakers all to tend to encode manner of motion in their verbs, and what different but shared ones would lead Spanish, Greek, and Japanese speakers all to tend to encode path (or path and ground) of motion instead (e.g., Talmy, 1983, 1985; Slobin, 1996a)? What differing cultural needs, interests, ecologies, or experiences would lead English speakers to make a lexical distinction between spatial relations called *in* and those called *on*, Dutch speakers to further split the relations encoded in English *on* into *op* and *aan*, and Spanish speakers to label all of those by a single word, *en* (Bowerman, 1996)? Furthermore, native speakers of the same language in diverse parts of the world—such as English speakers in the United States, Australia, and India, and Spanish speakers in Mexico, Spain, New York City, and Argentina—all follow these conventions of their shared language despite substantially differing cultures and daily experiences. What underlying cultural commonality would cause them to do so while others having no less cultural similarity diverge (e.g., Dutch citizens vs. Anglo Americans or English-speaking vs. Spanish-speaking Americans)? It could be proposed that these linguistic features were shaped by shared cultural conditions among early speakers of each language. The current distribution of shared patterns may be the result of diffusion of a language to other parts of the world: Emigrants maintain language patterns but their culture changes, and the language may be adopted by other cultures in the new location. Under this solution, we move away from the core idea of a tight connection between words and concepts by suggesting that the tight connection may at best exist only historically and not synchronically.

Cross-Linguistically Variable Word Meanings Lead to Culturally Variable Concepts Another possible scenario for a tight connection between words and concepts makes language the causal agent. Under this scenario, by acquiring and using the naming patterns of their native language, speakers of different languages come to have concepts shaped by those patterns. If different languages have different meanings associated with words, then speakers of those languages will have correspondingly different concepts, producing the tight linkage. This possibility constitutes one version of the Whorfian hypothesis (Whorf, 1956) that language shapes thought (see, e.g., Gumperz & Levinson, 1996), and it is consistent with the developmental perspective on word learning that the learning process guides concept formation (e.g., Choi & Bowerman, 1991; Imai, Gentner, & Uchida, 1994; Imai & Mazuka, 2003; Waxman & Gelman, 1986). Even if prelinguistic infants appreciate global distinctions such as that between objects and substances, provided they do not have more fine-grained concepts already sorted out and ready to receive labels, the language they are exposed to might shape the concepts subsequently acquired (e.g., Imai & Mazuka, 2003; Soja, Carey, & Spelke, 1991).

This scenario raises the question of what would have led to the divergent patterns of naming in the first place. A tempting answer is that it would be some sort of differences in the concepts held by speakers of different languages. To avoid circularity, it is then necessary to postulate conceptual differences existing for reasons independent of language. The likely source would be cultural differences as just discussed. If independent reasons exist, though, it is not clear why language should be invoked as a causal agent in any conceptual differences across speakers of different languages or any word-concept correspondences across languages.

An alternative way of thinking about the origin of cross-linguistic differences allows the possibility that languages shape concepts while avoiding this trap. This version entails that word meanings in a given language are shaped by forces independent of the conceptual representations of its own users. Specifically, at the birth of a language speakers would develop some words, presumably at least in part to express notions of importance to them. However, some elements of arbitrariness in the early development will exist, such as whether the language encodes path or manner in its verbs, simply because a finite set of words can capture only a fraction of the richness of human experience. As the language continues to evolve, a variety of external forces such as contact with other languages could subsequently alter the set of words available in a domain and the meanings associated with each word. (We will elaborate on this idea in the second section of the chapter.) The words passed down to the next generation would thus have significant elements of arbitrariness relative to the particular concepts that spawned the earliest bits of the language. The process of word learning would then itself shape the concepts of the next generation, maintaining a tight mapping between words and concepts in each generation, even as the language continues to evolve. Furthermore, because each language will evolve on its own path, and because convergence in communication systems requires contact between agents using the system (e.g., Barr, 2005), communities of language users that are not in close contact with each other will develop different concepts tied to their own language.

One problem with this scenario is that if we allow for continued drift in word meanings over time, there must be some set of words at a given moment that does not map closely onto the concepts of the current speakers. That is, the tight mapping cannot hold for all word–concept pairs. There is no obvious way to identify the cases in which we should expect the close mapping to hold and in which not. Still, it seems undeniable that as children learn each word, they must learn what dimensions are relevant to its use and how the values on those dimensions contrast with other words in the semantic field (e.g., Clark, this volume). It also seems inevitable that using a word requires accessing that knowledge and attending to the relevant dimensions. Will this directing of attention alter nonlinguistic representations such that there often, if not always, develops a tight mapping between a concept and a word meaning? We suggest that there are several reasons why word learning and use will not necessarily lead to a tight mapping.

First, what is not captured in a given language in its lexical contrasts may be salient for nonlinguistic reasons. Humans develop an understanding of their world not just to talk about it but to move around in it physically and accomplish goals from obtaining food and water to reproducing to creating art and making scientific discoveries. If languages encoded only distinctions key to successful navigation within a culture, then these patterns might be expected to strongly guide and map onto the conceptual distinctions that humans would develop. But if there are significant elements of arbitrariness in the development of naming patterns—the assumption necessary for this version of tight mapping to have force—these patterns may at times be poor guideposts for developing an understanding of the world. English speakers do not lexically distinguish male from female cousins, as speakers of some other languages do, but they still need to choose gender-appropriate birthday gifts for their cousins. Speakers of any language would interact with a jar of peanut butter, a container of yogurt, and a Coke bottle in much the same way in order to extract their contents, and they would drink from cups, mugs, and glasses similarly. The differing patterns of linguistic groupings for these objects across languages (Kronenfeld, Armstrong, & Wilmoth, 1985; Malt et al., 1999) seem to provide no useful information to guide these interactions. To the extent that members of the different cultures actually do need to learn slightly different ways of accomplishing their goals—perhaps because they have different implements for scooping peanut butter or yogurt—these differences concern the current status of the speaker's world. Because the linguistic arbitrariness, by definition, is not determined by current cultural conditions, the cultural differences are not likely to be reflected in any useful way in the naming patterns learned.

Second, the linguistic categories of a language can sensitize an observer to the existence of contrasts, but they do not by themselves reveal what distinctions are being labeled. Appreciating the distinctions requires learning about the entities themselves. Languages make many distinctions that their speakers do not appreciate nonlinguistically. For instance, adult Americans are familiar with words such as *elm*, *maple*, *sycamore*, and *beech*, and *sparrow*, *finch*, *dove*, and *jay*, but most cannot explain the differences among their referents (nor link the names to appropriate referents) (Dougherty, 1978; Wolff, Medin, & Pankratz, 1999). Wolff et al. (1999) found that despite a high frequency of use of genus-level tree terms in the early 1800s, English speakers' knowledge of trees declined sharply in the next 100 years. Subsequent generations appeared to pay less attention to distinctions among trees even though the language they inherited from previous generations made available a rich set of distinctions. Interactions with the world, not only the pattern of word use in the language spoken, give rise to individuals' knowledge of trees and birds.

Third, it is often an oversimplification to suggest that because some conceptual distinction is not encoded in a parallel way across languages, the linguistic attention drawn to it must vary. For instance, although Chinese does not have a subjunctive mood to express counterfactuals, it does have other ways of expressing states counter to reality (Au, 1983). Likewise, although a Spanish speaker may have only one verb, *saltar*, to encode the differing motions labeled *hop* and *jump* in English, Spanish speakers can readily disambiguate by specifying additionally *en un pie* ["with one foot"] or *en dos piernas* ["with two legs"] or the like. In such cases it can still be argued that although both languages can express the notion, it is easier or more often done in one language than the other, which may lead to greater salience or more habitual use of the concepts. In other cases, though, this argument is less persuasive. For instance, as we demonstrate later, Japanese has only a small number of single-word verbs for manners of locomotion, whereas English has a larger set (*stroll, saunter, stomp, march, hop, jump,* etc.). Does that mean that Japanese speakers are less likely to notice differences among the gaits used for human locomotion? In fact, Japanese uses other devices including "mimetics" to express manners of moving. Mimetics are expressions that have a nonarbitrary relation to their referent; their sound in some way gives clues to the nature of the referent. Japanese infants learn verbs consisting of mimetics faster than verbs with a purely arbitrary phonological relation to referents (Imai, Kita, Nagumo, & Okada, 2008). Thus, even though the expression may seem to be more complex, it cannot be concluded that the expression is more awkward for native speakers to learn or use, nor that speaking the language will necessarily result in less salience of the element of experience. In addition, one element of a language may discourage attention to some aspect of experience while a different element encourages it. Lucy (this volume) suggests that English-speaking children will be sensitized to the shape of objects (relative to Yucatek children) because English nouns are extended on the basis of shape, whereas Carroll and Casagrande (1958) suggested that English-speaking children will be insensitive to the shape of objects (relative to Navaho children) because English has no morphemes encoding object shape attached to verbs. Taking into account both aspects of English, it is unclear what level of attention English-speaking children should pay to object shape. Each child is, of course, subject to the potential influence of every dimension of their lexical and grammatical systems, and the impact on cognition any element could have in isolation may be mitigated by the impact of other elements of the language system.

Finally, it must logically be true that any useful conceptual distinction that *is* directly reflected in language had to have been noticed by humans before the words labeling that distinction came about. The existence of a lexical distinction cannot be a prerequisite for appreciating a distinction in the world. The fact that nonhuman animals can make many discriminations in their world likewise indicates that language cannot be a prerequisite to appreciating many distinctions. In both cases, attention to distinctions is most likely shaped by the utility or consequences of making them, and sensitivity to some discriminations that have had value over evolutionary history may even be hard-wired into their brains. Given that human language may have emerged as recently as 30,000 years ago (Crystal, 1987) and the genus *Homo* is believed to have diverged from its relatives approximately 2 million years ago, such sensitivities may be entirely independent of language.

Although these arguments make the case that language need not inevitably shape conceptual representation, there is evidence favoring the possibility that it does so in some cases. It is beyond the scope of this article to review the extensive body of literature evaluating this possibility. Some of this research is reviewed in this volume; see also Gumperz and Levinson (1996) and Gentner and Goldin-Meadow (2003) among others. In some domains in which effects of language on nonlinguistic representations have been found, the exact interpretation of the effects remains up in the air. For instance, the presence and strength of language effects on color perception seem to vary depending on the hemisphere to which the stimulus is presented and the speed at which responding occurs (suggesting that the effects may require engaging the linguistic system at the time of stimulus processing; see Regier et al., this volume; Roberson & Hanley, this volume), and some languages have shown gender effects while others have not (Vigliocco, Vinson, Paganelli, & Dworzynski, 2005). Regardless, there may be cases in which language does shape concepts in an enduring fashion, with the effect not dependent on engaging language at the moment of stimulus processing. Emotions, for instance, are abstract, and the direct experience of them is (by definition) a feeling rather than a reflective thought. The interpretation of emotional reactions may be heavily guided by parent–child discussions of them during the early years of development (Laible & Song, 2006), and so the particular distinctions among emotions that an adult notices may be shaped by the terms that his or her language offers for framing those interpretations. Nevertheless, other evidence suggests that such enduring effects do not occur in all domains. In the next section, we will discuss some evidence for the observation that, in accordance with our arguments, words do not inevitably create a tight mapping between themselves and conceptual representations.

Conclusions about the Tight Mapping Possibility We have considered three possible scenarios in which there would be a tight mapping between words and concepts. One posits universal concepts producing universal word meanings. This one can be confidently discarded on the basis of data showing that word meanings are far from consistently shared across languages. A second scenario posits culturally diverse concepts leading to diversity of word meaning. Although this scenario may account for some cases of linguistic diversity, it does not seem to give a useful account of a substantial body of observed cases. Only one possible version of tight mapping makes sense from the perspective of allowing patterns of word meaning and word use to vary in ways not directly predicted by cultural conditions of current speakers—the version in which languages vary for reasons independent of current cultural conditions and then shape the concepts of their speakers. However, there are a number of reasons why languages might vary in such ways and yet not shape concepts, or shape them only under some circumstances, and so it is not a foregone conclusion that this version is right.

Before moving on, we note that there is one sense of the term *concept* under which words must map directly onto concepts, and cross-linguistic differences must imply corresponding differences in concepts. For speakers of English, Spanish, and French to have differing patterns of applying the words *bottle*, *botella*, and *bouteille* to objects (and so on), the knowledge that the speakers have about the meanings or uses of these words must differ in some respects. In everyday talk, one might say that the speakers differ in what their concepts of bottles are (or that the American concept of a bottle differs from the Spanish concept of a botella and the French concept of a bouteille). This terminology is also sometimes adopted in research literature (e.g., Pavlenko & Jarvis, 2001).

But this sense of *concept* entails that all differences between languages automatically yield differences in concepts. If taken as the relevant sense, there would be no need for any debate or empirical evaluation of the relation of words to conceptual representations. Indeed, if taken as the relevant sense, it would be impossible to empirically evaluate this relation, since the conclusion has been drawn before any data are collected. It is the possibility that the representations engaged in nonlinguistic cognitive processes are not inevitably tied to linguistic differences that makes questions about the relation of language to thought interesting, important, and susceptible to scientific investigation. We therefore take the cross-linguistic differences in the knowledge associated with words in different languages as a priori evidence only for differences in *linguistic* concepts or *word meanings*, making no assumption that these linguistic concepts also constitute the stock of general purpose mental representations engaged in a nonlinguistic understanding of the world.

Loose Mapping

At the other extreme from tight mapping lies the logical possibility of a very loose mapping in which there is little relation between how people experience the world and how their language encodes it. For instance, suppose people see an important similarity among dogs, wolves, and coyotes on the one hand and cats, lions, and tigers on the other, but their language only has words for grouping animals by size and ferocity. The vocabulary places domestic cats, rabbits, and small dogs in one labeled group, large dogs and goats and sheep in another, and lions, tigers, and bears in another, making the linguistic distinctions available to them arbitrary with respect to the salient conceptual distinctions. This relation might come about if language originated with early humans under conditions quite dramatically different from those that currently prevail—culturally and possibly also in terms of perceptual and cognitive capacities—and individual languages failed to evolve as internal and external conditions did.

This scenario is unlikely on two grounds: First, languages do evolve and can reflect changing cultural conditions (as shown, for instance, by vocabulary that emerges with new technologies) even though some significant elements may be arbitrary with respect to current cultural conditions. Second, in many respects, languages seem well-suited to human experience and to conveying the ideas that humans want to convey. After all, language did evolve in order to serve communication needs, and so any language is likely to have devices reasonably well-suited to serving those needs, even if it also has some arbitrariness. For instance, humans perceive differences between objects and events, and languages commonly have ways of lexically distinguishing objects from actions. Kin relations are important in most human cultures, and languages tend to have words to distinguish among varieties of kin (mother vs. father, child vs. parent, etc.; Greenberg, 1966).

We cannot completely rule out the possibility that such correspondences come about from the other causal direction as discussed earlier—namely, that it is language that has shaped human thinking to see these distinctions as the important ones. To the extent that some of these distinctions are shared across languages, it seems more likely that the causal direction is from thought to language. Regardless of the source of the correspondence, though, it seems that the relation between language and the way humans experience the world is not completely awkward, ill-fitting, and arbitrary, and so we can set aside the extreme loose-fit possibility.

THE CONSTRAINED BUT FLEXIBLE MAPPING POSSIBILITY

Our preceding discussion suggests that the extreme loose mapping possibility, in which an arbitrary relation is pervasive and there are few or no constraints on the relation between language and thought, can be discarded. It also suggests that the tight mapping possibility in the form of a causal influence from universal thought to universal language can be discarded, based on overwhelming evidence for linguistic diversity. The tight mapping possibility in the form of a causal influence from culturally variable thought to linguistically variable language has some plausibility, but it seems inadequate as a full explanation of the relation between words and thought. The tight mapping possibility in the form of a causal influence from language to thought is also viable, but, as we have argued, for a variety of reasons it is not inevitably the correct description of the relation. In this section, we first review evidence suggesting that at least in some domains, there is some dissociation between how people think about a domain and how they label it. In other words, the mapping between words and understanding of the world is not always very tight. We then ask, if the mapping is nevertheless constrained in some way that creates some shared tendencies across languages, what do the constraints consist of? Finally, we ask, given constraints, why is there also diversity in how words relate to the world? Where and why does it emerge? Answers to these questions will help illuminate the true nature of the mapping between knowledge of the world and knowledge of words, which we argue is loose enough to allow for considerable flexibility in the relation between them but nevertheless constrained in significant ways.

Dissociations between Experience and the Expression of Experience

If there is consistently a tight mapping between words and concepts, there should be a correspondence between how people talk about a given domain and measures of their nonlinguistic understanding of the domain. Malt et al. (1999) evaluated the relationship between the two for 60 ordinary household storage containers, for speakers of American English, Argentinean Spanish, and Mandarin Chinese. They examined the perceived similarity among the objects as a measure of how people thought about the objects and the relations among them. They also assessed which objects in the set were called by the same name for each language, to determine which objects are linguistically grouped together. There was a surprising degree of divergence in naming patterns. For instance, English speakers put most objects into one of three categories of roughly equal size (*bottle*, *jar*, and *container*) whereas Spanish speakers called 28 of them by a single name (*frasco* or its diminutive, *frasquito*) but used an additional 14 names to partition the rest. Chinese speakers preferred the same name for 40 of the objects, and used only four additional names for the remaining 20 objects. The differences across languages consisted of more than just minor boundary variations around shared prototypes. The categories of the different languages were not always formed around the same prototypes, and in some cases the category memberships across the languages cross-cut each other substantially (Malt, Sloman, & Gennari, 2003a).

In contrast to the cross-linguistic differences in naming, similarity judgments were remarkably consistent across speakers of the three languages. The correlations of similarity matrices between each language group (English–Spanish, Spanish–Chinese, and Chinese–English) were all above 0.90, and analyses using the Cultural Consensus Model (Romney, Weller, & Batchelder, 1986) to assess agreement in naming versus similarity using a common measure confirmed that between-group differences were significantly larger for naming than for similarity. Furthermore, to the extent that there were differences in perceived similarity, these were not systematically related in any detectable way to the differences in naming patterns. For this domain, then, it seems that knowledge of words and knowledge of the world are less tightly linked than the tight mapping possibility posits. Somehow, languages can come to have different patterns of encoding the objects in words even though individual speakers of the languages may perceive and understand their properties in much the same way.

Other data also show a similar outcome for common objects. Ameel et al. (2005) replicated Malt et al.'s (1999) findings by comparing Belgians who speak Dutch with Belgians who speak French. This replication shows that when different languages are spoken by people sharing essentially the same culture, their patterns of naming can still diverge, even though their perception of the similarities among the objects is in close correspondence. Kronenfeld et al. (1985) examined similarity and naming judgments for a smaller set of drinking vessels for speakers of English and Hebrew, and likewise found shared perceived similarity but substantially differing groupings by name.

Studies of other domains suggest that dissociations are not limited to the object domain. Munnich, Landau, and Dosher (2001) noted that English makes an obligatory distinction between situations involving support (labeled *on*) and ones not involving support (receiving labels such as *above* or *in front of*), whereas this distinction is optional in Japanese and Korean. They found comparable memory for the spatial locations of objects despite the differing naming patterns. Similar outcomes have been found for simple motion events such as a person moving into or out of a room. English speakers tend to encode the manner of motion in the main verb of a clause and the path of motion in adverbial phrases (e.g., *She is walking/running/limping out of the room*). Spanish speakers often encode path of motion in the main verb (the equivalent of *She is exiting the room*), and less commonly they mention manner in an adverbial phrase (Talmy, 1983, 1985; Slobin, 1996a). Gennari, Sloman, Malt, and Fitch (2002) found that despite the expected differences in descriptions of action film clips, speakers of English and Spanish had similar confusions between clips on an old–new memory task. Papafragou, Massey, and Gleitman (2002) obtained similar results comparing English to Greek, a language that follows the Spanish pattern. Papafragou, Massey, and Gleitman (2006) further found that Greek speakers increased mention of manner for events in which the manner could not readily be inferred, suggesting that they were monitoring manner information even when the typology of their language did not compel them to express it. Papafragou, Hulbert, and Trueswell (2008) also found that when speakers of English and Greek inspected short video motion events with the instruction to remember them, they showed indistinguishable patterns of eye movements during the event. Once the movement stopped, speakers of English actually paid more attention to path than manner but speakers of Greek did not differ in attention to the two elements. This outcome suggests that English speakers may have been encoding into memory the element that their language did not readily capture in an internal linguistic summary. It seems that speakers of different languages experience the elements of simple motion events independently of their linguistic likelihood of encoding manner.

So, cross-linguistic variability in naming in the face of shared nonlinguistic responses occurs for several different domains. Consistent with our earlier argument, they suggest that cultural differences are not necessarily the source of the disparate word meanings and naming patterns that speakers of different languages have. There is no obvious link between broad cultural differences among our Chinese, American, and Argentinean participants, or among the Americans, Israelis, and Japanese of Kronenfeld et al. (1985), and the groups' varied linguistic partitioning of object stimuli. Furthermore, the French- and Dutch-speaking Belgians of Ameel et al. (2005) show that people may share largely the same culture but maintain differences in their naming patterns. And if members of the cultures see the similarity among the entities in much the same way, then there is little basis for postulating specific differences in the cultural construals of those entities that could lead to differences in how they are partitioned by name. In addition, and importantly, considering the data from the other causal direction, the results indicate that the words of a language do not necessarily create a tight link to the way people perceive or understand entities in a domain. The particular pattern of naming that speakers use does not, at least in the domains studied, fix their perception of the similarity among the entities in the domain.[1] The data argue against the more viable versions of the tight mapping possibility as an across-the-board account of how knowledge of words relates to knowledge of the world.

Constraints

So, it seems that language does not inevitably create a tight mapping to conceptual understanding of objects, and conceptual understandings are not inevitably the source of the specific configuration of lexical categories in a language. Yet, as we have discussed, it also seems that the meanings associated with words of languages are in some way reasonably well shaped to convey the ideas that humans want to convey, and some shared tendencies in naming patterns across languages have been identified. There must be some kind of causal, although imperfect, relation between conceptual representations and the development of patterns of naming. How might whatever correspondences do exist come about? As we suggested, there may be some culture-specific correspondences that arise from the need to communicate certain distinctions of particular importance to a culture, and there may be a contribution of cross-culturally shared needs, goals, feelings, and so on. Further, pan-human basic cognitive and perceptual capacities might segment the world for speakers of all languages in comparable ways by creating special sensitivities to some distinctions, and the world might also present itself to the human observer packaged in chunks so salient that given shared perceptual and cognitive capacities, all humans recognize these chunks. However, there has been little past evaluation of these potential sources for shared tendencies. The best developed is perhaps Kay and McDaniel's (1978) argument for the contribution of pan-human perceptual capacities to universally perceived nonlinguistic color categories (though an alternative explanation has since been proposed; Regier, Kay, & Khetarpal, 2007). In the research we now discuss, we focus on asking whether there is also a contribution from structure in the world and whether certain shared cross-cultural communication needs might complement structure in creating similarities in the mappings across languages.

Locomotion on a Treadmill As we noted earlier, direct experience in the world is a source of much knowledge. Some of that experience may present itself to the observer in ways that make certain distinctions particularly salient. Anthropologists Hunn (1977) and Berlin (1992) made such an argument for structural constraints on the labeling of plants and animals across cultures. Drawing on analyses from biology, they suggested that at the level of the biological genus, properties of plants and animals occur in clusters, and there are distinct gaps between clusters. Thus canines share certain sets of characteristic features, felines share others, equines share others, and so on, and there are few or no animals in between these clusters that have properties coming from two or more of the clusters (see also Rosch, 1978). People will perceive these property clusters, and their labeled distinctions will tend to map onto the clusters. As a result, people from many cultures in disparate parts of the world, and speaking different languages, will tend to label the same distinctions among plants and animals.

We further investigated the possibility of structural constraints on naming by examining lexicalization of part of the domain of human locomotion (walking versus running). For plants and animals, people living in different places speaking different languages are exposed to different members of the domain, which allowed Berlin and Hunn only indirect assessments of the consistency of labeled groupings across cultural groups by comparing each culture's groupings to biological taxonomies. In contrast, people in different parts of the world, speaking different languages, will be exposed to many of the same gaits. Although locomotion is not immune to cultural variation (Devine, 1985), human bodies are capable of a number of different basic gaits regardless of culture (such as those labeled in English as *walk*, *run*, and *hop;* e.g., Alexander, 2002), and so speakers of different languages should see or engage in many of them universally. In particular, for the current study, what is crucial is that people of all cultures will have been exposed to both walking and running, and, since these are the primary gaits used in daily life, they should all find these gaits salient.

For locomotion, as for plants and animals, an independent analysis of the domain structure exists. This structural description is given by the biomechanical analysis of human gaits (e.g., Alexander, 2002; Bennett, 1992). Some salient parts of the domain are described as highly structured, with strong clusters of exemplars having sharp discontinuities between them. In particular, in walking, the legs act like a pendulum around a fulcrum point and one foot is always in contact with the ground. Running has an impact-and-recoil motion, and there is a point in each stride in which neither foot is in contact with the ground. Due to the dynamics of motion, there are abrupt transitions from one gait to another rather than gradual shifts through intermediate versions. Studies of English speakers observing locomotion on a treadmill demonstrate that the abrupt discontinuity in biomechanical properties of walking versus running is reflected in English naming, which shows uniform responses of *walk* up to a certain speed and then uniform responses of *run* (Diedrich & Warren, 1995).

Taking into account this pan-human exposure to two salient gaits, and the structural distinction between the gaits (which may be apparent in an individual's own motor experience as well as in the perceptual experience of observing gaits executed by others), we can ask whether these discontinuities in the locomotion stimulus space are consistently drawn in the naming of gaits across languages.

We studied speakers of English, Spanish, and Japanese. [See Malt, Gennari, Imai, Ameel, Tsuda, and Majid (2008) for a report of the study that also includes Dutch data added after the preparation of this chapter.] English and Spanish are both in the Indo-European family, but English is largely Germanic and Spanish is in the Romance branch. Modern English does have considerable Romance influence in its vocabulary, but its manner verb lexicon is generally Germanic in origin. Japanese is most often classified as belonging to the Altaic family (Crystal, 1987). As relatively unrelated languages, any similarities in naming patterns across these three languages are unlikely to be due to shared linguistic histories.

Furthermore, English is a language that is characterized as a manner verb language, in which verbs frequently express manner, whereas Spanish more commonly uses verbs that express path of motion (Talmy, 1983, 1985; Slobin, 1996a). In Japanese, verbs tend to express path or path plus ground/trajectory (Muehleisen & Imai, 1997). Languages that more often encode path in the main verb do have some manner verbs in their vocabulary, however. This trio of languages allows us to investigate the extent to which the manner verbs that do exist in such languages encode the same distinctions lexicalized in a language in which mention of manner within the main verb is more common. In this portion of the locomotion domain, experienced with high frequency and in cross-culturally similar ways, if the biomechanical discontinuity is salient, then all languages may develop manner verbs marking the same distinction between walking and running regardless of verb typology.[2]

We filmed a student locomoting on a treadmill that varied systematically in speed and slope. There were three slope levels: flat, a slight slope, and a steeper slope. Within each slope, we started at the slowest speed possible and increased it by one treadmill unit at a time until it became too difficult for the student to remain on the treadmill. This process resulted in nine clips on the flat surface, eight on the slight slope, and seven on the steeper slope, for a total of 24 clips. The clips were embedded in a web page in random order, each followed by "What is the woman doing? She is. . ." or its translation into Spanish or Japanese as appropriate. Participants watched each clip and answered the question by typing a word or phrase into a response box on the web page. A portion of the English language version of the experiment is illustrated in Figure 2.1. Participants were native, largely monolingual speakers of their language (recruited in the United States for English speakers, Argentina for Spanish speakers, and Japan for Japanese speakers).

Because our interest for the current purpose is in thinking about the relation between nonlinguistic experience of the world and the meanings captured in individual words, we focus here on the head verb of each response produced. This focus is not to say that speakers of these languages are unable to, or unlikely to, differentiate the gaits linguistically in other ways when they do not have a unique verb for a manner of motion. Modifiers may be attached to verbs, or other descriptive phrases may occur within a sentence to distinguish among motions. On-going analyses are examining these other expressions of locomotion.

We tabulated the frequency of the verbs produced to each clip by speakers of each language and then focused on the use of verbs that were the dominant (i.e., most frequent)

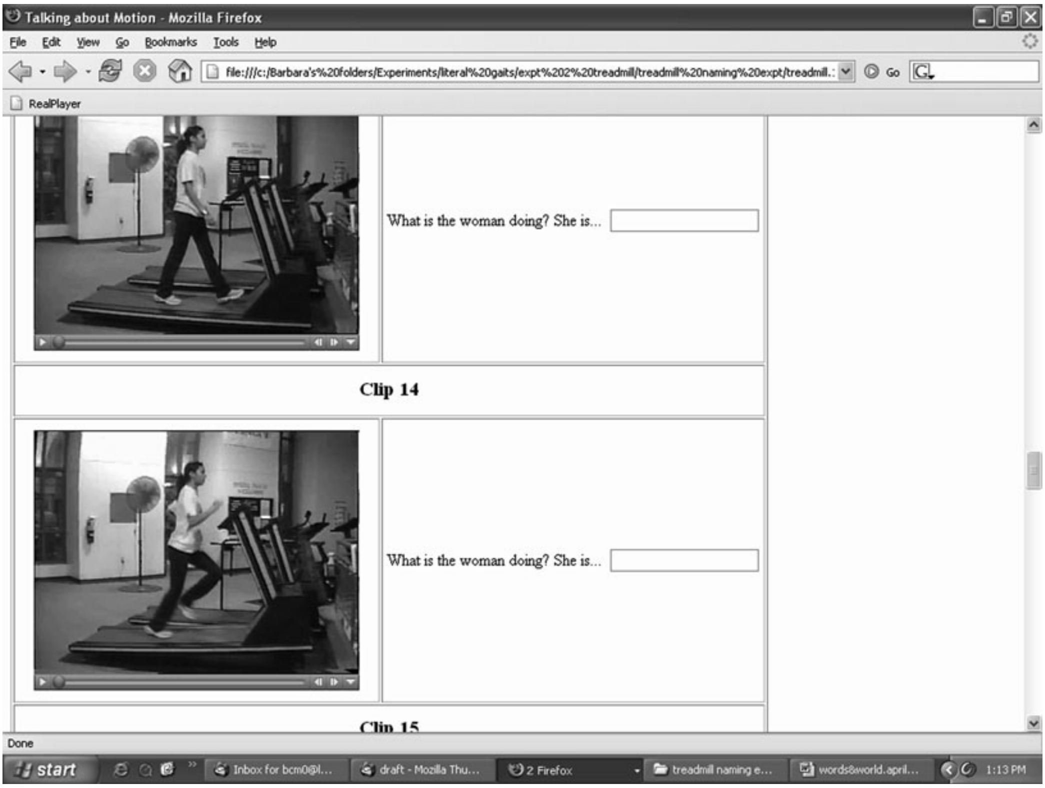


FIGURE 2.1. Sample of web page presenting treadmill clips to participants. The top clip is on the flat surface at speed 2; the second clip is on the slight slope at speed 7.

response for at least one clip. If speakers use terms in a way that observes the structural discontinuity, we would expect that verbs will be applied to clips in a categorical, not graded, fashion, and each language should have terms used in a complementary distribution that is paralleled by the other languages.

What we found is shown in Figures 2.2 and 2.3. Verb distributions were not graded. For each language, speakers switched from one set of names to another in an all-or-none fashion. Furthermore, speakers of all three languages made their transitions from one set of terms to another at exactly the same points in the stimulus continuum; these points correspond to the biomechanical discontinuities in the movements produced. And Japanese and Spanish users made this distinction with unique, single-word manner verbs just as did English speakers, despite the fact that they speak languages that do not, overall, encode manner in the verb of a sentence nearly as often as English does. This result provides strong evidence that there can be mappings from the world to words that are shared based on a shared perception of structure in the world.

Interestingly, as the figures make clear, there are nevertheless some differences between the languages. English and Spanish speakers had a term limited to slow running exemplars (*jogging* and *trotando*, respectively). The distribution of these terms relative to *running* and *corriendo* is graded; use of *jogging* and *trotando* gradually drops off as *running* and *corriendo* increase over the speed manipulation. English speakers also sometimes used *sprinting* for the fastest gaits. The graded nature of the trade-off along the speed dimension here reinforces the conclusion that

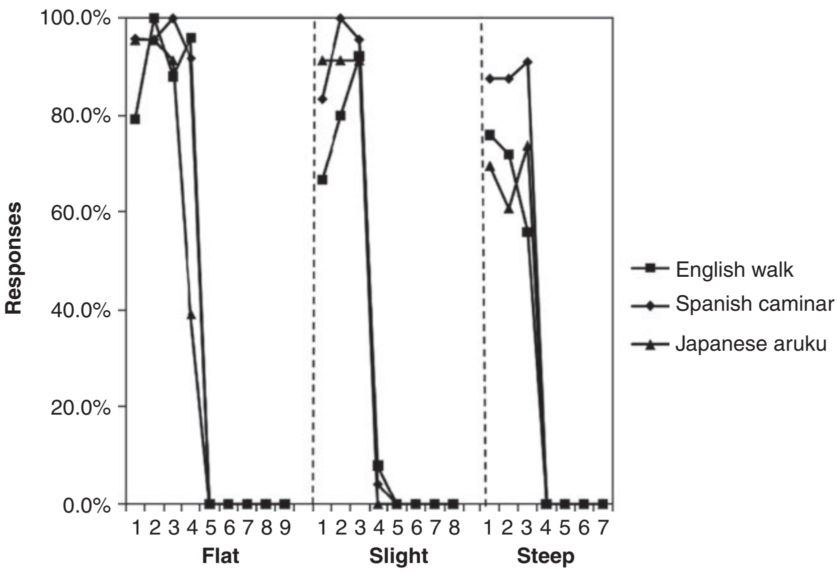


FIGURE 2.2. Distribution of responses to treadmill clips for terms encompassing biomechanical walking.

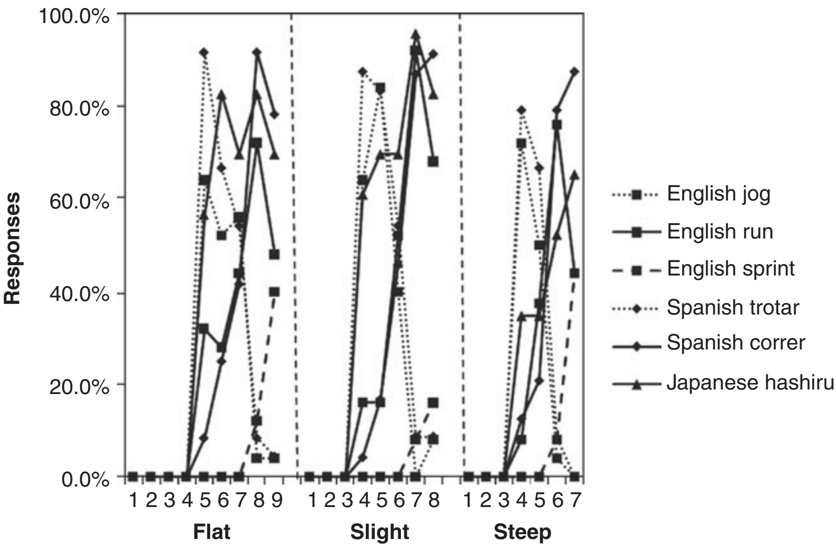


FIGURE 2.3. Distribution of responses to treadmill clips for terms encompassing biomechanical running.

the contrasting, all-or-none, trade-off in labels when crossing the biomechanical gait boundary does reflect a perception of the discontinuous nature of the stimulus space.

In sum, in this case where the world presents strong structure, and for a portion of the domain that is presumably important across cultures and is experienced in similar ways, the three languages made the same distinctions with manner verbs despite their differing linguistic histories and verb typologies. The verbs *walk* and *run* (with varied spellings) appear in English as early as the 1300s and were used in discourse contexts similar to their modern uses (Oxford English Dictionary, 1989). *Caminar* has its origins in pre-Roman Celtic vocabulary, and *correr* came into Spanish from Latin (Real Academia Española Staff, 2001). *Aruku* and *hashiru* are represented in Japanese by characters that are Chinese in origin, but the words themselves may have predated the Chinese influence, which began in the sixth century AD. With such varied origins and long trajectories over which the meanings could evolve, it is particularly remarkable that the current uses of the words so closely match in observing the distinction between biomechanical gaits. It appears that structure in the world, when observed in a domain that is common and presumably important, constrains the mapping between words and the world.

Flexibility

The preceding data demonstrate how the mapping between experience and words to label it may be constrained by salient structure in a domain or portion of it. At the same time, we found some differences between the languages even in this domain, and, as we have discussed earlier, overall, there is considerable diversity in patterns of naming of various domains across languages. Some substantial portion of this diversity is not related to current cultural differences in any obvious way. Why does diversity arise even with the constraints that structure in the world may provide? A key observation, we believe, is that the kind of strong structure in the world that the walk–run biomechanical distinction presents is often not present. In many domains or parts of those domains, the distribution of properties across entities in the domain is much less tightly clustered. For instance, for artifacts such as the common household containers we have studied, new objects can be created with all sorts of combinations of values on dimensions ranging from size to shape to type of opening to use. Even in cases such as spatial relations, where the location of one object with respect to another is limited by the laws of physics, there may be no major discontinuities across the possible relations that would cause all languages to group the same ranges of relations together by name. In such cases, there is greater room for other factors to influence how the domain or part of a domain is lexically divided.

One factor that can create diversity across languages is the variable salience of entities to members of a culture due to the particular needs or interests of the culture. As we have noted, such variation may drive different languages to develop vocabulary in a domain to different extents, as seems to happen for color or wine terminology. Also, because people in different cultures may experience somewhat different entities as the manifestation of a domain (or similar entities but with different frequencies), the meanings associated with words are likely to be influenced by what is present or common to them. What is prototypical of a lexical category in one culture may be peripheral in another (e.g., Schwanenflugel & Rey, 1986).

Other factors may result in diversity that is not tightly linked to current cultural differences. As we touched on before, cultures evolve over time, and so the lexical distinctions in a language at a given time may be, in part, a product of past cultural needs, goals, interests, or experiences rather than current ones. The word meanings of a language themselves are also in a constant state of evolution (e.g., Hock, 1996; Traugott & Dasher, 2005), for reasons partially distinct from cultural goals, interests, or needs. Contact between different languages can introduce new words into a domain in a language, causing previously existing ones to expand, contract, or otherwise modify their meaning and patterns of application in ways that might differ from another language. For instance, English distinguishes between the live animal, *pig*, labeled by a word of Germanic origin, and the food, *pork*, the latter term having entered English from French (similar to the situation for *cow* vs. *beef*; Hock, 1996), whereas for chicken and fish, English makes no such distinction. Dutch, however, has only a single word for a pig and its flesh, following the pattern for chicken and fish (E. Ameel, personal communication).[3] Likewise, meanings may shift as words come to have new pragmatic functions in a language. For instance, a word such as *lady*, originally used only for women of the highest social standing, may begin to be applied to women of lesser standing out of politeness and thus eventually lose or even reverse its original status implication (Keller, 1994). A word such as *woman*, which may have previously contrasted with *lady*, might then expand to encompass those previously known as *lady*, or perhaps even become the signifier of higher status. And the meanings at any one time are a function of not only the forces that shape meanings but also of the particular input (the previously existing words and structures available in the language and their associated meanings) on which those forces operate, making the development of these meanings a dynamic process subject to multiple grammatical and pragmatic constraints (Keller, 1994; Traugott & Dasher, 2005). As a result, even two languages with similar characteristics spoken by members of similar cultures at some moment in history might have different patterns of lexicalizing a domain if the evolutionary paths of their vocabulary for the domain differed (see Roberson & Hanley, this volume, for a related argument for color terms).

Other linguistic factors from outside the realm of semantics per se may also shape how a domain is segmented by name. These are differences in the syntax and morphology of languages. For instance, languages that have gender marking systems are forced to make a lexicalized distinction between male and female cousins (e.g., *primo* vs. *prima* in Spanish; *cousin* vs. *cousine* in French), whereas languages without gender marking need not make that distinction. In our container data, we noted that Spanish speakers made more lexical distinctions among containers than English or Chinese speakers did. Spanish morphology makes it easy to form single-word names for containers by adding the *-ero/-era* and *-or* suffixes to root words (e.g., *talquera* for an object holding talc; *roceador* for an object for spraying), and this feature may contribute to the substantially larger number of lexical distinctions that exist in Spanish. In a related vein, diversity among languages in how manner and path of motion are expressed in words may stem at least partly from other differences in the morphosyntactic devices they make available for encoding the semantic ingredients common to all representations of motion events (Levin, Beavers, & Tham, 2004; see Wolff et al., this volume, for another compelling example of this sort of influence). Thus multiple interacting forces working over the course of a language's history are likely to shape the lexical resources available to speakers of a language at any point in time, independent of the particular interactions with the world speakers have or culturally shaped ways they may learn of thinking about a domain.

Locomotion on a Walkway The domain of locomotion allows us to test several ideas about where cross-linguistic diversity comes from. In a second study on naming of locomotion, we examined naming of a wider range of gaits produced under more natural conditions. One important idea to test, in the long run, is that where structure in the stimulus input is less clear, there is more opportunity for diversity among languages in naming patterns. Because the literature on biomechanical qualities of human gait focuses primarily on walking and running, however, we do not have as useful an objective indicator of where the most major structural discontinuities in a larger range of motions lie. Jumping and hopping have been referred to as separate gaits (Alexander, 2002). It is less clear whether there are multiple correlated properties that separate running from skipping, for instance, or whether the distinctions lie in fewer dimensions. We are currently collecting similarity judgments on the larger set of gait exemplars used in this study to establish what people see as the major physical similarities and dissimilarities among the stimuli. We can use those judgments to make predictions about where weaker structure is likely to lead to greater cross-linguistic diversity. For now, one prediction we can evaluate is that diversity should be greater within a gait than between gaits. The data from Study 1 that we discussed are consistent with this prediction. We can use the greater range of variations in the current stimulus set to further test this idea.

A second idea that can be tested in this study is that correspondence should be greater among the languages for the more central portion of the domain. Different types of locomotion are of substantially differing degrees of centrality to human experience. Regardless of culture or location, most of the time when a person observes or engages in human locomotion, the event will be of walking, and sometimes it will be of running. Much less often it will be of hopping, skipping, jumping, etc. Thus for all cultures, the need to make reference should be greatest for the more central parts of the domain (walking and running), and less so for the more peripheral parts of the domain (hopping, skipping, jumping, etc.).[4] If centrality to a culture's gait options (and attendant degree of need or likelihood of wanting to express an experience in language) affects what distinctions languages encode in their lexicon, we should expect a high degree of uniformity in drawing the distinction between walking and running across languages. What distinctions are lexically encoded in unique verbs for the more peripheral gaits should be more susceptible to variability brought about by the various other forces that shape lexicalization patterns over a language's history.

This domain also allows us to test further whether the differences between languages in how commonly manner is expressed in the verb can influence how a domain is lexicalized. Languages that less often express manner because the verb often encodes path instead may develop fewer verbs to encode manner distinctions overall (Slobin, 2004), despite the fact that they do make the important walk–run distinction. We therefore examined the data to determine if Spanish and Japanese speakers would produce fewer manner verbs than English speakers in naming these more varied gaits.

The gaits filmed came from two sources. One was a list of all the manner of motion verbs in English (provided to us by Dan Slobin). We had a student act out each of the verbs that named ways of moving forward, backward, or sideways bipedally or on one foot (eliminating many on his list such as *barge*, *bolt*, *bound*, *bump*, and *burst*, which seem to capture elements of movement such as speed, suddenness, or gracefulness but not gait per se). We selected for the final set of clips those that seemed visually distinct from one another. (For instance, the filmed versions of *trudge* and *plod* differed little, if at all, so we kept only one.) We also had informants from Japan and Argentina videotape culturally relevant movements not covered by the English terms. These included several distinctive military march-type movements, two traditional Japanese styles of walking, and two modern Argentinean styles of walking. The student who served as actor for filming the rest of the gaits viewed them and reproduced the actions on the walkway along with the rest of the gaits. The final stimuli consisted of 36 clips illustrating variants of gaits such as walking, running, marching, and jumping. Figure 2.4 illustrates several of the motions filmed.

As before, the clips were embedded in a web page that allowed participants to watch each one and then type in what they thought they would call the motion. Participants were native, largely monolingual speakers of the three languages (recruited in the United States, Japan, and Argentina) who had not participated in the treadmill study. As before, we tabulated the frequency of the verbs produced to each clip by speakers of each language and then focused on the use of verbs that were the dominant (i.e., most frequent) response for at least one clip.

Table 2.1 presents the most frequent response to each clip for each of the three languages, along with the proportion of responses it accounted for. (If no term accounted for at least one-third of responses to a clip, the response is listed as "mixed.") As before, we report main verbs without modifiers. In two cases, the Japanese dominant response was a verb formed from a noun plus light verb (*ashibumi-suru* and *sukippu-suru*); these are conventional verbs in Japanese. To make similarities and differences among the languages more apparent, the clips are grouped according to their dominant English name; we can then see to what extent the Spanish and Japanese distributions of names match the English and each other. As in Study 1, terms in all three languages segregated a variety of

FIGURE 2.4. Sample frames from clips in the walkway study.

pendulum-based limb motions from bounce-and-recoil motions. Thus this basic gait distinction is lexically observed across the more stylistically varied versions of the two gaits used in this study. This result supports the conclusion from Study 1 that a lexicalized gait distinction is shared across languages and is based on a shared perception of structure that exists in the world.

In contrast to the treadmill experiment, however, there is one stimulus (labeled "trot" in Table 2.1) in this portion of the domain in which the responses are variable rather than all-or-none (and, for Japanese speakers, the term associated with other walking clips dominated whereas for English and Spanish, a term associated with running clips dominated). Inspection of the motion involved reveals why this response pattern occurs. The movement is essentially pendulum based with one foot in contact with the ground at all times (as in other movements called by walking terms), but it has more knee bending at one point in each pace and therefore a bouncier quality than other motions called by walking terms (see Fig. 2.5). Responses reflect the mixed features of the stimulus.

We had predicted that diversity should be greater within a defined gait than between gaits because the structural differences creating variations of a gait will be much less sharp and perceptually less salient. For instance, languages should differ more in marking variations within the biomechanical category of walking than they do in marking the distinction between walking and running. The data support this suggestion. Japanese never applied their main walking term to walking backward, but Americans and Argentineans did with a high level of consistency. Japanese speakers never applied their main walking term to walking in place, although both Americans and Argentineans sometimes did. On

TABLE 2.1. Dominant Responses in English, Spanish, and Japanese and the Percentage of Participants Who Produced Each Response for 36 Examples of Locomotion on a Walkway[a]

	English		Spanish		Japanese	
Clip Name	Dominant Verb	Dominance (%)	Dominant Verb	Dominance (%)	Dominant Verb	Dominance (%)
Shuffle	Walking	0.50	Caminando	0.86	Aruku	0.92
Stroll	Walking	0.70	Caminando	0.91	Aruku	0.88
Trudge	Walking	0.50	Caminando	0.91	Aruku	0.80
Noh	Walking	0.40	Caminando	0.68	Aruku	0.68
Slink	Walking	0.80	Caminando	0.91	Aruku	0.88
Stride	Walking	0.43	Caminando	0.90	Aruku	0.80
Ghetto walk	Walking	0.43	Caminando	0.80	Aruku	0.64
Lumber	Walking	0.93	Caminando	1.00	Aruku	0.88
Strut	Walking	0.66	Caminando	0.86	Aruku	0.92
Heels Argentinean	Walking	0.80	Caminando	0.77	Aruku	0.72
Walk backward	Walking	0.93	Caminando	0.95	Aruku	0.48
Clomp	Walking	0.60	Caminando	0.95	Aruku	0.84
Pigeon toed	Walking	0.77	Caminando	0.76	Aruku	0.68
Heels	Walking	0.80	Caminando	0.67	Aruku	0.64
Walk in place	Walking	0.43	*Mixed*	0.32	Ashibumi-suru	0.60
Trot	Jogging	0.33	Trotando	0.50	Aruku	0.44
Jog	Jogging	0.70	Trotando	0.86	Hashiru	0.32
Run fast	Running	0.67	Corriendo	0.95	Hashiru	0.56
Run in place	Running	0.53	Trotando	0.68	*Mixed*	0.12
Goose step	Marching	0.43	Marchando	0.86	Aruku	0.32
March Japanese	Marching	0.67	Marchando	0.67	Aruku	0.32
March American	Marching	0.80	Marchando	0.76	Aruku	0.44
March Argentinean	Marching	0.48	Marchando	0.67	Aruku	0.36
March in place	Marching	0.83	Marchando	0.82	*Mixed*	0.08
Jump	Jumping	0.52	Saltando	0.73	*Mixed*	0.16
Jump in place	Jumping	0.67	Saltando	0.95	*Mixed*	0.28
Hop	Hopping	0.70	Saltando	0.68	*Mixed*	0.28
Hop in place	Hopping	0.73	Saltando	1.00	*Mixed*	0.28
Skip	Skipping	0.93	*Mixed*	0.23	Sukippu-suru	0.84
Tiptoe	Tiptoeing	0.37	Caminando	0.95	Aruku	0.68
Stomp	Stomping	0.45	Caminando	0.57	Aruku	0.68
Gallop	Galloping	0.45	*Mixed*	0.27	*Mixed*	0.12
Leap	Leaping	0.40	Saltando	0.36	*Mixed*	0.16
Sneak	Creeping	0.40	Caminando	0.77	Aruku	0.48
Power walk	Power walking	0.40	Caminando	0.86	Aruku	0.68
Step sideways	*Mixed*	0.07	*Mixed*	0.25	*Mixed*	0.28

[a]The dominant response is listed as "mixed" if fewer than one-third of participants produced the same name. "Clip name" is an experimenter-given description of the clip.

the other hand, Americans rarely used their main walking term for three other clips for which Spanish and Japanese speakers readily did. Americans used the names *tiptoe*, *creep*, and *stomp*, respectively, for these more than any other name, but with a low degree of consensus. It appears that Spanish and especially Japanese treat forward movement as a more central part of the meaning of their walking term than English does, whereas English may tolerate more variation in path of movement but less in details of the style. Also, all three languages used their main walking term for some actions that varied from a strictly pendulum motion, but there was variation in which ones they extended the term to. Japanese speakers used their main walking term

FIGURE 2.5. The "trot" stimulus.

for two forms of marching and for stomping, and Spanish speakers did just for stomping—actions involving more pronounced knee bends—but English speakers did not. Again, the Spanish and Japanese terms in particular seem to allow variation from the prototypical movement in their use. Spanish and Japanese speakers were substantially less willing than the English speakers to apply their main running term to running in place.

We had predicted that the languages would agree more in the composition of the lexical categories covering the more central, commonly experienced, portion of the domain (walking and running) than the more peripheral, lower frequency portion (other gaits). We evaluated this prediction by first separating the clips into those that showed conformity to the biomechanical definitions of walking or running and those that showed other types of gaits. This resulted in 15 clips being classified as belonging to the central part of the domain (13 walking clips and two running clips). The remaining 21 clips were considered to be the peripheral part of the domain and included actions such as those described in English as *marching*, *hopping*, *skipping*, and *jumping*, and actions performed in place or moving sideways or backward.

We then examined the naming pattern in each language by determining, for each pair of clips, whether they received the same or a different dominant name in that language (that is, were placed in the same lexical category). So, for instance, if Clips 1 and 2 were called *walk* by Americans and Clip 3 was called *run*, then Clips 1 and 2 were counted as having been placed in the same lexical category, Clips 1 and 3 were counted as having been placed in different ones, and Clips 2 and 3 were also counted as being in different ones. We could then compare the languages to see to what extent they showed similar patterns of placing clips into the same or different lexical categories. We coded pairs with a shared category as "1" and those with different categories as "0" and correlated the resulting arrays between each pair of languages, looking separately at the clips in the central portion and those in the peripheral portion. For clips in the central portion, the average correlation between languages was 0.83; for clips in the peripheral portion, it was only 0.31. Each pair of languages individually showed greater agreement in the central than the peripheral portion. This analysis thus supports the contention that languages will tend to diverge more in naming for stimuli that receive less attention and where lexically encoding certain observable distinctions may be less important.

We also predicted that English, as a manner verb language that frequently encodes manner of motion in the main verb, would show greater lexical differentiation of the gaits than would Spanish (a path verb language) and Japanese (a path-and-ground language), but that in particular the difference would be greater in the more peripheral part of the domain. We evaluated this possibility by counting the number of verbs that were dominant for at least one stimulus, for each language. For English speakers, four different verbs emerged as dominant for at least one clip apiece in the central portion (*tiptoe*, *walk*, *jog*, and *run*) and an additional nine did in the peripheral portion (*hop*, *skip*, *jump*, *march*, *gallop*, *creep*, *leap*, *stomp*, *and power-walk*). For Spanish speakers, three were dominant in the central portion (*caminar*, *trotar*, *correr*) and only two additional verbs emerged as dominant in the peripheral portion (*marchar* and *saltar*). Clearly, Spanish speakers made many fewer discriminations via unique verbs than English speakers did, and the difference appears primarily for the less common gaits.

[As noted earlier, this observation does not mean that Spanish speakers are incapable of expressing finer distinctions among the gaits. For instance, although having only a single verb, *saltar*, that was applied to gaits that Americans called *hop*, *skip*, *jump*, and *leap*, Spanish speakers often used additional descriptors such as *en un pie* (= on one foot) or *con los dos pies juntos* (= with both feet together) to describe the gaits more fully.]

Japanese presented a particularly interesting case. Japanese speakers had only two dominant verbs in the central portion (*aruku* and *hashiru*) of the domain, and two others occurred as dominant within the clips, *ashibumi-suru* and *sukippu-suru*.[5] The contrast with English in the extent of diversity within the dominant verbs for these clips is thus even more striking than for Spanish. However, it is not that Japanese speakers failed to discriminate among the actions in their verbs. Rather, Japanese affords speakers a variety of different ways of encoding motion in verbs. For instance, almost 25% of all responses were various noun + light verb (*suru*) combinations, but these were formed from a variety of nouns including loan words as well as traditional Japanese words, producing low consensus in the main verb of responses. Japanese, then, is even sparser than Spanish in the existence of high-consensus verbs for drawing distinctions in the peripheral part of the gait domain, but this sparseness is compensated for, and in fact possibly due to, the availability of a variety of ways for conveying manner of motion in a verb phrase.

The data from this study are consistent with suggestions we put forward about where flexibility enters into the mapping between knowledge of the world and knowledge of words. Where structure in the world is less clear, there are more opportunities for languages to diverge in their patterns of lexically grouping stimuli. Where structure matters less because less attention is paid to some portion of the domain, there is likely to be more diversity. And independently existing characteristics of languages may influence the development of vocabulary in a domain. In the case of these data, it seems that not only may the often-discussed manner-path difference among languages influence the development of manner verbs used with consistency by speakers, but so may the availability of other options for expressing manner distinctions. In fact, a larger point here may be that there is a trade-off between the degree to which languages have acquired large stocks of unique vocabulary words for specific motions versus have developed ways of expressing subtleties of meaning through morphological and syntactic complexity. English has an extremely large number of vocabulary words overall as a result of the many different languages that have contributed to it (e.g., Crystal, 2003), and so it encodes many distinctions in single-word lexical items. On the other hand, Spanish and Japanese are languages with greater morphological and syntactic complexity (Talmy, 1983, 1985) and so may tend to encode more distinctions via multiword phrases. If salient or commonly mentioned phenomena are most likely to be encoded simply by virtue of their frequent use (Bybee, 2003), these differences among the languages may become most apparent in those domains or parts of a domain in which structural distinctions are less striking or less frequently talked about.

IMPLICATIONS AND CONCLUSIONS

In the preceding section, we discussed evidence indicating that there can be a dissociation between how people think about objects in a domain and how they label them, in at least some cases. In other words, the mapping between words and experience of the world is not always very tight. We proposed some mechanisms by which the mapping might be partially constrained but yet at the same time flexible, and we presented data consistent with this proposal. Given this evidence, and our discussion of the tight and loose mapping possibilities that preceded it, what conclusions are suggested about the mapping between knowledge of the world and knowledge of words?

First, the way that people think about some domain or portion of it—in terms of what they see as the important distinctions within it, and what they might indicate belong together or share the greatest similarity in a laboratory task—may at times correspond very well to the way that the words of their language group things. In such cases, we can say that there is indeed a tight mapping between words and concepts. The presence of such a correspondence can be anticipated by the presence of significant discontinuities or structural distinctions in the stimulus array. Where there is such strong structure, languages will tend to correspond to one another in the way they lexically divide the domain. Correspondences that come about for this reason do not indicate that language has shaped thought. Rather, they indicate that the world has shaped thought, which in turn has shaped language.

Second, the way that people think about some domain may at other times not correspond well to the way that the words of their language group things. In such cases, we can say that there is a loose mapping between words and concepts. The presence of such looseness can be anticipated by the presence of weak or little structure in a stimulus array, which leaves the evolution of vocabulary for the domain susceptible to a range of other influences. The same set of influences will have the potential to shape the vocabularies of every language, but the impact of each factor and the outcomes that result will be highly variable from language to language because of the multiple interactions among factors and the way that the state of a language and culture at one moment feeds into outcomes at the next moment in the language's evolution. The evidence that we presented for the influence of these multiple factors in the gait data did not directly provide evidence that the resulting patterns of naming are dissociated from how speakers of the languages think about the domain. However, other studies we discussed demonstrated clearly that such dissociations can occur. Furthermore, such dissociations make sense from the perspective of the arguments we raised earlier about why language may not always be the determinant of how people think about some things, including the myriad ways that people learn about some parts of the world aside from language, and the fact that attention to certain distinctions or lack thereof must be shaped in part by the utility of making such discriminations for functioning in the world.

In light of these two points, we cannot conclude that the mapping between words and concepts is best characterized either as tight across the board or as loose. Our proposal of a constrained but flexible mapping allows for the existence of both possibilities while providing some suggestions about when each might come about. We now consider some broader implications of our perspective.

The Importance of Working out Causal Paths

We began this chapter by considering the widely held view that words map onto coherent packets of nonlinguistic knowledge constituting concepts, and we raised the puzzle of how such a close mapping could come about in light of recent research demonstrating pervasive cross-linguistic differences in how people talk about the world. If linguistic diversity is paralleled by conceptual diversity, where do the parallels come from? If concepts are the causal agent, we run up against the problem that the differing cultural needs, conditions, and experiences of current speakers do not seem to explain many of the types of differences that occur across languages, and conversely, shared linguistic patterns are exhibited by speakers of the same language who live under widely varying cultural conditions. Perhaps language is the causal agent instead. But, we noted, under this hypothesis it is necessary to explain what causes the languages to be different in the first place. Suggesting that the answer is different concepts held by speakers of the languages leads to circularity. A better alternative is to consider that arbitrary linguistic differences arise through mechanisms independent of conceptual differences. Once it is acknowledged that these differences are arbitrary with respect to how people go about interacting with the world, though, it is apparent that these differences might not be useful for guiding the way people think about the world, and that other sources of information for how to think about the world may dominate, in at least some domains.

The different scenarios we have sketched out may not be the only ones available to help explain what the relation is between words and concepts. However, we think it is an important exercise to try to be explicit about what the story is for how linguistic diversity comes about. It is also important to be explicit about how that story relates to what is known about cultural differences and the types of conceptual differences they may lead to, as well how it relates to the particular instances of linguistic diversity that have been observed. A clear account of how linguistic diversity arises and what aspects of word meanings vary across languages will provide clues to the causal relationship between language and thought. If there are viable alternatives to the possibilities we have raised, making them explicit should yield further insights about the relations.

Diachronic as Well as Synchronic Processes Are Relevant

One point that our suggested account of the relations makes salient is that the origins of linguistic diversity observed at any point in time may be due to events long past such as language contact and the shaping of word meanings through the entry of new lexical items into the vocabulary, as well as historic cultural practices or conditions not directly reflected in current lifestyles. Psychologists are interested in the mental representations and processes of individuals and how they are shaped over the course of an individual's development, and so they have naturally tended to take a synchronic perspective on the language–thought relationship. That is, they have assumed that the word–concept mappings held by individuals arise through processes occurring at a given moment or within the individual's lifespan, without regard to larger historical context. From this perspective, it is natural to assume either direct causation (in one direction or the other) between the word knowledge and conceptual knowledge held by individuals, or else an independence between them. Each of these possibilities leaves some portion of the empirical evidence about patterns of word use, concepts, or the relation between them in individuals impossible to explain. Taking a longer-term view of how differences among languages arise permits breaking out of the paradoxes that arise under a completely synchronic approach. Adding the diachronic perspective helps explain why current patterns of word use can differ from those an individual speaker would impose on a domain if she or he were simply given an array of exemplars of the domain and the task of grouping them to assign to lexical items (see Slobin, 2001, for a related argument concerning grammaticalization), and why, in turn, these patterns of word use may not always be effective in shaping the perception of the domain.

Language as a Window into the Mind

Chomsky (1968) considered language to be a window into the mind and linguistics therefore a branch of cognitive psychology. Pinker (2007) echoes this sentiment in referring to language as a window into human nature, and cognitive linguists make a similar assumption (e.g., Lakoff, 1987). Chomsky's original comments concerned how syntax and the possibility of a universal grammar could provide insights into the architecture of the mind, revealing universal capacities such as the all-important ability to handle recursion. Pinker (2007) and cognitive linguists, however, go beyond syntax to argue that metaphors and even the meanings of individual words and paradigms of meanings in a language's repertoire reveal something fundamental about cultures and individuals. These views are not based on the idea that language shapes thought but rather assume the alternative version of causation: The way people think is reflected in the words of the language they speak. The words are an external manifestation of the internal workings of the mind. This view assumes the synchronicity of the causal link between concepts and words, whereas we have argued that the link may be more distant and mediated by nonconceptual influences. If our view is right, word meanings may still be useful in understanding the nature of the human mind, but the understanding of what it is that they reveal may need to be more nuanced.

Implications for Language Learning

The perspective we have advocated also helps to make sense of an apparent paradox in language learning that is posed by our data on common nouns and on verbs of locomotion. Gentner (1982) has noted that nouns are represented in children's early vocabulary to a much greater extent than verbs are, and this tendency holds across languages. She (1982, 2003, 2005) argues that concrete nouns are easy to learn because they refer to entities easily segregated from their background. She points out that, on the other hand, to understand a verb, a child has to determine which elements of a scene are encoded into the verb, and this varies from language to language—it might be the manner, path, figure, or ground of the action (Talmy, 1983, 1985), or some combination. In this way, the meanings of verbs are more linguistically embedded than the meanings of nouns, and their meanings are more language specific. Consistent with this conclusion, Gleitman (1990) argues that verb meanings may be inferred more from the argument structures in which they occur within discourse than from direct experience.

But the data we have presented make two points that at first glance might seem contradictory to this line of reasoning about why nouns should be easier to learn than verbs. First, we found that the sets of objects referred to by common nouns vary substantially across languages (and hence the meaning associated with the nouns presumably does too), and those sets are not well predicted by perception of similarity among the objects. Thus the meanings of concrete, common nouns are not necessarily so readily derived by observation of the world. Indeed, Andersen (1975) and Ameel, Malt, and Storms (2008) demonstrate that children learning their native language do not fully converge on adult usage for common nouns until age 12 years or beyond. Second, we found that the usage patterns of several verbs of locomotion (and hence presumably the meaning associated with the words) were quite strongly shared across three disparate languages, and this shared nature is well predicted by structure presented to the observer by the world. Thus our data suggest that (some) concrete noun meanings can be relatively difficult to derive from observation and (some) verb meanings may be relatively easy.

We suggest that our data are fully compatible with Gentner's proposal if a distinction is made between early and later aspects of word learning. Gentner has focused on how readily individual referents of a word can be segregated from the backgrounds against which they are embedded. Here, the learning challenges may be as she describes: The notion that nouns label whole objects may be simple to grasp (Markman, 1990), and identifying the individual object encoded by a noun in a given context may be perceptually and conceptually a simple task (Hollich, Golinkoff, & Hirsh-Pasek, 2007). Thus understanding what kind of mapping exists between nouns and referents, and acquiring prototypes providing an initial fast mapping of meaning (Carey, 1978) to concrete nouns, may be relatively easy to achieve. Meanwhile, figuring out whether the input language has a bias to encode manner, path, or some other element of meaning in its verbs may take more work, and segregating that component from other cooccurring aspects of an observed action in progress may be perceptually and/or conceptually much trickier. However, grasping what the basic nature of the mapping for a grammatical category is and identifying some initial referents of words to provide a rudimentary sense of meaning are only part of the job. Children must ultimately infer the broader conditions of applicability of the word that allow them to use the word for referents besides the one(s) initially observed (see also Ameel et al., 2008; Imai, Haryu, Okada, Li, & Shigematsu, 2006; Saji, Saalbach, Imai, Zhang, Shu, & Okada, 2008). As we have seen, the forces that interact over a language's evolution to determine the extension of a word can conspire to create simpler or more complex patterns of use, and the degree of complexity need not correspond neatly to the noun–verb distinction. (See Golinkoff, Chung, Hirsh-Pasek, Liu, Bertenthal, Brand, Maguire, and Hannon, 2002, for evidence that 3-year-old children readily perceive and label gait differences represented only by point-light displays; see also Ma, Golinkoff, Hirsh-Pasek, McDonough, and Tardiff, 2009.) Fully understanding how the child's knowledge of the world comes to be translated to knowledge of words may require considering two aspects of the task: the relation of the concepts being acquired by the child to the words they are learning to use and the nature of the full target word use. The first considerations may be most revealing for understanding earliest word use and the second for understanding the process of achieving fully mature use of the words.

Nonlinguistic Knowledge: Concepts or Something Else?

To this point, we have mostly used standard terminology and asked about the relation between language and concepts or conceptual representations. Indeed, we began this chapter by raising the standard assumption that words map onto coherent packets of nonlinguistic knowledge constituting concepts. But the perspective we have argued for suggests that the traditional way of talking about the relation of language to nonlinguistic knowledge may benefit from some adjustment. Words may package together certain elements of nonlinguistic knowledge for communicative purposes (so, hearing or using the word *bottle* evokes one subset of knowledge and hearing or using *jar* evokes a partially different subset), but if it is granted that the words do not necessarily dictate the shape or content of nonlinguistic representation, it may be useful to set aside the idea that such representations come in packets worthy of the name *concepts*. For domains in which the stimulus space is not highly structured, the nonlinguistic knowledge may not have any inherent boundaries or coherent packets that resemble what is traditionally thought of as concepts. [It may be word use itself that creates the impression that it does (Sloman & Malt, 2003; Malt & Sloman, 2007).] It may be more useful to think about the nonlinguistic knowledge in terms of smaller components or features (which may constitute "primitives" such as manner and path of motion; e.g., Parish-Morris, Pruden, Ma, Hirsh-Pasek, & Golinkoff, this volume) and to consider separately to what extent there are correlations among the features and to what extent languages package these features together. With this more fine-grained notion of representation less tightly tied to word use, it is easy to imagine, for instance, how different tasks may tap different elements of the representations, producing experimental results that vary in the extent to which linguistic performance and nonlinguistic performance mirror each other (Gennari et al., 2002; Saalbach & Imai, 2007). Greater progress in understanding the mapping between knowledge of the world and knowledge of words may come from looking beyond how words relate to "concepts" per se.

ACKNOWLEDGMENTS

We thank Asifa Majid for helpful comments on an earlier draft, Kristine Schuster for developing the web experiments on locomotion, Stephanie Sterrett for serving as actor in the walkway clips, Caitilin Gallagher for serving as actor in the treadmill clips, and Hiroyuki Okada and Naoki Tsuda for preparing the Japanese versions of the web experiments. Stephanie Sterrett, Caitilin Gallagher, Lisa Guest, Celina Hayes, Allison Baker, Erin Howard, Ruohong Wei, Sandra Cvejanov, Leonor Tecker, and Minako Ushino provided valuable assistance in collecting and analyzing data. This work was supported in part by a Ministry of Education grants-in-aid for scientific research awarded to Mutsumi Imai (#15300088 and #18300089) and by research grants from Keio University (Keio Gijuku Academic Development Fund and Keio Gijuku Mori Memorial Research Fund) to Imai.

Notes

1. Some studies have found that linguistic factors influence similarity judgments (e.g., Gennari et al., 2002; Saalbach & Imai, 2008). We do not suggest that such influences can never be found (see also the discussion in the text), but rather that they are not an inevitable consequence of a causal relation between language and thought. The appearance of such effects may depend on particular task demands.
2. Note that our goal here is not to examine the manner-path distinction per se, nor its pattern of use in language production, as some past studies have done. Instead, we take advantage of this well-documented typological difference to ask: *When* languages make manner distinctions, what is the nature of their lexical categories for manner, and does their typological status as a manner or path verb language have an implication for the answer to this question?
3. An example of how borrowings get adopted with different results in different languages is as follows: French *boeuf* came into English as *beef*, distinguished from *cow*, which presumably was then elaborated into *beefsteak* for a specific cut of beef and then borrowed back into French as *bifteck*, also for a specific cut of beef. *Bifteck* or *beefsteak* was then borrowed into Dutch to create *biefstuk*, but Dutch already had a specific term for the meat of cows, *rundvlees* (as well as a less commonly used term, *koeievlees*, which compositionally means "cow meat," E. Ameel, personal communication) and perhaps, as a result, *biefstuk* in Dutch refers to the flesh of any animal, not just that of cows.
4. If some forms of locomotion are typical, it could be argued that they should be unmarked and therefore less likely to receive lexical status. However, any trend in this direction is likely outweighed by the absolute likelihood of the different motions. For instance, an English speaker is more likely to say *I walked to the library* than *I skipped* simply because it would be very rare for the form of locomotion to be skipping. English word frequency counts indicate that *walk* and *run* have a much higher frequency in text than *hop*, *skip*, and *jump* (Baayan, Piepenbrock, & Gulikers, 1995).
5. Japanese does have several other single-word locomotion verbs that did not emerge as dominant for our particular clips, such as *tobu*, which is used for leaping actions.

References

Alexander, R. M. (2002). *Principles of animal locomotion*. Princeton: Princeton University Press.

Ameel, E., Malt, B. C., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language*, *58*, 262–285.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, *53*, 60–80.

Andersen, E. S. (1975). Cups and glasses: Learning that boundaries are vague. *Journal of Child Language*, *2*, 79–103.

Au, T. K. (1983). Chinese and English counterfactuals: The Sapir-Whorf hypothesis revisited. *Cognition*, *15*, 155–187.

Baayan, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (release 2) [CD-ROM]. Philadelphia, PA: Linguistics Data Consortium, University of Pennsylvania [Distributor].

Barr, D. J. (2005). The emergence of conventions in language communities. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 331–344). Cambridge, MA: MIT Press.

Bennett, M. B. (1992). Empirical studies of walking and running. In R. McN. Alexander (Ed.), *Comparative and environmental physiology, vol. 11: Mechanics of animal locomotion*. Berlin: Springer-Verlag.

Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.

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

Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In B. D. Joseph & J. Janda (Eds.), *The handbook of historical linguistics* (pp. 602–623). Oxford: Blackwell.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Carey, S. (2001). Whorf versus continuity theorists: Bringing data to bear on the debate. In M. Bowerman and S. C. Levison (Eds.), *Language acquisition and conceptual development* (pp. 185–214). Cambridge, UK: Cambridge University Press.

Carroll, J. B., & Casagrande, J. B. (1958). The function of language classifications in behavior. In E. E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (pp. 18–31). New York: Henry Holt.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *42*, 83–121.

Chomsky, N. (1968). *Language and mind*. New York: Harcourt, Brace, Jovanovich, Inc.

Clark, E. V. (1983). Meanings and concepts. In P. H. Mussen (Series Ed.) & J. H. Flavell & E. M. Markman (Vol. Eds.), *Handbook of child psychology*, Vol. 3: *Cognitive development* (pp. 787–840). New York: John Wiley.

Clark, E. V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences*, *8*, 472–478.

Crystal, D. (1987). *Cambridge encyclopedia of language*. Cambridge, UK: Cambridge University Press.

Crystal, D. (2003). *Cambridge encyclopedia of the English language*. Cambridge, UK: Cambridge University Press.

Devine, J. (1985). The versatility of human locomotion. *American Anthropologist*, *87*, 550–570.

Diedrich, F. J., & Warren, W. H. (1995). Why change gaits? Dynamics of the walk-run transition. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 183–202.

Dougherty, J. W. D. (1978). Salience and relativity in classification. *American Ethnologist*, *15*, 66–80.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

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

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development, Vol. 2: Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–236). Cambridge, MA: MIT Press.

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison, (Eds.), *Building object categories in developmental time* (pp. 245–275). Hillsdale, NJ: Lawrence Erlbaum.

Gentner, D., & Goldin-Meadow, S. (Eds.) (2003). *Language in mind: Advances in the study of language and thought* (pp. 195–236). Cambridge, MA: MIT Press.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition: A Journal of Developmental Linguistics*, *1*, 3–55.

Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., Maguire, M. J., & Hannon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology*, *38*, 604–614.

Greenberg, J. (1966). *Language universals*. The Hague: Mouton.

Gumperz, J. J., & Levinson, S. C. (Eds.). (1996). *Rethinking linguistic relativity*. Cambridge, UK: Cambridge University Press.

Gumperz, J. J., & Levinson, S. C. (1996). Introduction to Part 1. In J. J. Gumperz and S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 21–36). Cambridge, UK: Cambridge University Press.

Hock, H. H. (1996). *Language history, language change, and language relationship: An introduction to historical and comparative linguistics*. Berlin: Walter de Gruyter & Co.

Hollich, G., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology*, *43*, 1051–1061.

Hunn, E. (1977). *Tzeltal folk zoology: The classification of discontinuities in nature*. New York: Academic Press.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theory of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, *9*, 45–75.

Imai, M., Haryu, E., Okada, H., Li, L., & Shigematsu, J. (2006). Revisting the noun-verb debate: A cross-linguistic comparison of novel noun and verb learning in English-, Japanese- and Chinese-speaking children. In K. Hirsh-Pasek and R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 450–476). New York: Oxford University Press.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*, 54–65.

Imai, M., & Mazuka, R. (2003). Re-evaluating linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In D. Gentner & S. Goldin-Meadow (Eds.) *Language in mind: Advances in the study of language and thought* (pp. 429–464). Cambridge, MA: MIT Press.

Kay, P., Berlin, B., Maffi, L., & Merrifield, W. (1997). Color naming across languages. In C. L. Hardin & L. Maffi (Eds.), *Color categories in language and thought*. Cambridge, UK: Cambridge University Press.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, *54*, 610–646.

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Science*, *100*, 9085–9089.

Keller, R. (1994). *On language change: The invisible hand in language*. London: Routledge.

Kronenfeld, D. B., Armstrong, J. D., & Wilmoth, S. (1985). Exploring the internal structure of linguistic categories: An extensionist semantic view. In J. W. D. Dougherty (Ed.), *Directions in cognitive anthropology* (pp. 91–113). Urbana: University of Illinois Press.

Laible, D., & Song, J. (2006). Constructing emotional and relational understanding: The role of affect and mother-child discourse. *Merrill Palmer Quarterly*, *52*, 44–69.

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

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

Lehrer, A. (1983). *Wine and conversation*. Bloomington: Indiana University Press.

Levin, B., Beavers, J., & Tham, S. W. (2004). *A morphosyntactic basis for variation in the encoding of motion events.* Paper presented at the Conference on Diversity and Universals in Language: The Consequences of Variation, Stanford University, Stanford, CA.

Levinson, S. (1997). From outer to inner space: Linguistic categories and non-linguistic thinking. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 13–45). Cambridge, UK: Cambridge University Press.

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language*, *36*, 405–423.

Majid, A., Bowerman, M., van Staden, M., & Boster, J. S. (2007). The semantic categories of cutting and breaking events: A cross-linguistic perspective. *Cognitive Linguistics*, *18*, 133–152.

Majid, A., Enfield, N. J., & van Staden, M. (2006). Parts of the body: Cross-linguistic categorisation. Special issue of *Language Sciences*, 28(2–3).

Malt, B. C., Gennari, S., Imai, M., Ameel, E., Tsuda, N., & Majid, A. (2008). Talking about walking: Biomechanics and the language of locomotion. *Psychological Science*, *19*, 232–240.

Malt, B. C., & Sloman, S. A. (2007). Artifact categorization: The good, the bad, and the ugly. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 85–123). New York: Oxford University Press.

Malt, B. C., Sloman, S. A., & Gennari, S. (2003a). Universality and language-specificity in object naming. *Journal of Memory and Language*, *49*, 20–42.

Malt, B. C., Sloman, S. A., & Gennari, S. (2003b). Speaking vs. thinking about objects and actions. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 81–111). Cambridge, MA: MIT Press.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*, 230–262.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*, 57–77.

Muehleisen, V., & Imai, M. (1997). Transitivity and incorporation of ground information in Japanese path verbs. In M. Verspoor, Ke. D. Lee, & E. Sweetser (Eds.), *Lexical and syntactical constructions and the construction of meaning*. Amsterdam: John Benjamins.

Munnich, E., Landau, B., & Dosher, B.A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition*, *81*, 171–208.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. *Psychological Review*, *81*, 267–285.

Oxford English Dictionary, 2nd Edition (1989). Oxford: Oxford University Press.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155–184.

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

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition, 98*, B75–87.

Pavlenko, A., & Jarvis, S. (2001). Conceptual transfer: New perspectives on the study of crosslinguistic influence. In E. Németh (Ed.), *Cognition in language use: Selected papers from the 7th International Pragmatics Conference, Volume 1* (pp. 288–301). Antwerp: International Pragmatics Association.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking.

Pullum, G. K. (1991). The great Eskimo vocabulary hoax. In G. K. Pullum (Ed.), *The great Eskimo vocabulary hoax and other irreverent essays on the study of language* (pp. 159–171). Chicago: Chicago University Press.

Real Academia Española Staff. (2001). *Diccionario de la Lengua Española*. Spain: Espasa Calpe, S.A.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, *104*, 1436–1441.

Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, *88*, 313–338.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum.

Saalbach, H., & Imai, M. (2007). The scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, *136*, 485–501.

Saji, N., Saalbach, H., Imai, M., Zhang, Y., Shu, H., & Okada, H. (2008). Fast-mapping and reorganization: Development of verb meanings as a system. In *Proceedings of the 30th Annual Cognitive Science Society Meeting* (pp. 35–40). Mahwah, NJ: Lawrence Erlbaum.

Schwanenflugel, P. J., & Rey, M. (1986). Relationship between category typicality and concept familiarity: Evidence from Spanish- and English-speaking monolinguals. *Memory & Cognition, 14*, 150–163.

Slobin, D. (1996a). Two ways of travel: Verbs of motion in English and Spanish. In M. Shibatani & S. Thompson (Eds.), *Grammatical constructions: Their form and meaning*. Oxford: Clarendon Press.

Slobin, D. (1996b). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, UK: Cambridge University Press.

Slobin, D. (2001). Form-function relations: How do children find out what they are? In M. Bowerman & S. C. Levison (Eds.), *Language acquisition and conceptual development* (pp. 406–449). Cambridge, UK: Cambridge University Press.

Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Vol. 2: Typological and contextual perspectives* (pp. 219–257). Hillsdale, NJ: Lawrence Erlbaum.

Sloman, S. A., & Malt, B. C. (2003). Artifacts are not ascribed essences, nor are they treated as belonging to kinds. *Language and Cognitive Processes [Special Issue: Conceptual Representation], 18*, 563–582.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, *38*, 179–211.

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredelo (Eds.), *Spatial orientation: Theory, research, and applications* (pp. 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, Vol. 3: Grammatical categories and lexicon* (pp. 56–149). Cambridge, UK: Cambridge University Press.

Traugott, E. C., & Dasher, R. B. (2005). *Regularity in semantic change*. Cambridge, UK: Cambridge University Press.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General, 134*, 501–520.

Waxman, S. R., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*, 139–156.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. J. B. Carroll (Ed.). Cambridge, MA: MIT Press.

Wierzbicka, A. (1992). *Semantics, culture, and cognition*. Oxford: Oxford University Press.

Wolff, P., Medin, D.L., & Pankratz, C. (1999). Evolution and devolution of folkbiological knowledge. *Cognition, 73*, 177–204.

3 WORDS FOR PARTS OF THE BODY

Asifa Majid

The human body like any other object is perceived through our senses, but it is unique in that it alone is internally as well as externally apprehended. It serves as a special focus of attention. It is ritually dressed and adorned in a variety of ways in different cultures—from tattoos, piercings, and scarification to plastic surgery and make-up. It is widely regarded as the source for many of our concepts, particularly in relation to space, time, and emotions. Yet comparatively little attention is paid to the basic vocabulary for the body.

The body appears to be a discrete and independent semantic domain in memory. Neuropsychological studies show that in both production and comprehension, lexical-semantic knowledge of body parts can be relatively preserved (Coslett, Saffran, & Schwoebel, 2002; Shelton, Fouch, & Caramazza, 1998) or impaired (Dennis, 1976; Suzuki, Yamadori, & Fuji, 1997) in comparison to other semantic domains. Imaging studies also provide converging evidence that there are distinct cortical areas responsible for processing semantic knowledge of body parts, regardless of input modality (written, spoken) or language (for English-French bilinguals; see Le Clec'H et al., 2000). Shelton, Fouch, & Caramazza (1998; Caramazza & Shelton, 1998) have speculated that as a result of evolutionary pressures, body parts might be a specialized module neurally and functionally. According to their account, semantic representations are organized into domains because there have been specific adaptations to quickly classify and respond to objects relevant for survival value. The body would be one of these domains because body parts play a key role in interacting with the environment—"hands for grasping, legs for movement, eyes for seeing, mouths for ingesting, etc." (Shelton et al., 1998, p. 348). In addition to the evolutionary arguments, developmental evidence suggests that the body is special. Infants less than an hour old imitate facial movements (Meltzoff & Moore, 1983), and within a few weeks they can even imitate simple manual gestures (Meltzoff & Moore, 1977), suggesting an innate ability to perceive and interpret body parts.

The ontogenetic data and evolutionary arguments, along with evidence from perception that will be reviewed, suggest a fundamental categorization of the body into parts—i.e., into head, hands, arms, feet, legs, etc. The question addressed in this chapter is whether there are "basic" body parts that are recognized across cultures in how people talk about the body. Or to put it another way, how do people from different communities come to conceptualize the body in the languages they speak? Are the terms for parts of the body across the world's languages the same or different? And are there constraints on naming; if so, what are these constraints?

The particular semantic component on which I focus is the reference of body part terms, that is, what is the exact extension of body part terms across languages; how similar or different are the categorization systems?

This is but one component of this semantic domain. According to Kemmerer and Tranel (2008), additional components include information about the spatial organization of body parts, their characteristic functions, and their typical cultural associations. Although we may expect differences between languages in characteristic functions or cultural associations of body parts, many have predicted universals in body part categorization. I argue in this chapter that there is considerably more variation in the naming of body parts than is acknowledged, but that this variation is constrained.

BODY PARTS IN PERCEPTION

A dominant view held by many psychologists, linguists, and anthropologists is that body part categories are "given" by visual perceptual discontinuities, and that words are merely labels for these predetermined parts (e.g., Andersen, 1978; Biederman, 1987; Brown, 1976; Lakoff, 1987). The mapping is taken to be straightforward and obvious. There is one salient partitioning of the body into parts and all a speaker has to do is to identify which of these parts is associated with which particular label in their own language.

In many current theories of object recognition, objects are represented by parts (Biederman, 1987; Hoffman & Richards, 1984; Marr, 1982). Theories differ in detail, such as whether objects are segmented according to general purpose geometric constraints (Hoffman & Richards, 1984; Singh, Seyranian, & Hoffman, 1999; Xu & Singh, 2002) or into volumetric parts corresponding to shape primitives (e.g., Biederman, 1987; Marr & Nishihara, 1978). Nonetheless, there is consensus on core points: namely, that objects are segmented at discontinuities, and that there is a nested hierarchy of partitions, with parts lower in the hierarchy being smaller than parts higher in the hierarchy (Hoffman & Richards, 1984; Marr, 1982; Marr & Nishihara, 1978; Palmer, 1977).

For the human body these theories come to the same segmentation. According to Marr (1982), for example, visual processing begins with an image that has an array of intensity values, and after a number of processing stages this results in a three-dimensional model that is hierarchical in structure. The model for the human body consists of a number of generalized cylinders or cones. At the highest level of granularity the whole body can be represented as a single cylinder. At the next level the head, trunk, arms, and legs would each be represented by a separate cylinder. Then the arms and legs would further be subdivided into smaller cylinders, corresponding to upper-arm, lower-arm, upper-leg, and lower-leg, etc. (see Figure 3.1). In Biederman's (1987) formulation, these parts would be referred to as geons, and the head would be represented as a sphere, rather than a cylinder.[1]

Discussions about "natural" segments of the body have privileged vision over the other senses; however consideration of additional senses is highly pertinent to the issue of a perceptual partitioning of the body. As discussed in the introduction, the body is uniquely apprehended, being an object of internal perception through proprioception and somesthetic inputs, as well as an object of external perception through vision. Current psychological research takes it for granted that body parts named in language reflect the "true" and unique partitioning of the body (Schwoebel & Coslett, 2005; Sirigu, Grafman, Bressler, & Sunderland, 1991). But this may be too simplistic. There is now an emerging literature on how body parts are represented and organized

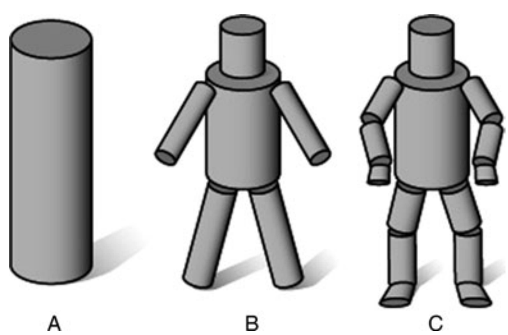


FIGURE 3.1. In visual perception the human body is represented by a three-dimensional hierarchical model (A–C) consisting of a series of cylinders.

in different perceptual modalities, as well as how these sensorial representations are pooled together to create an integrated and holistic representation of the body and its parts (de Vignemont, Majid, Jolla, & Haggard, 2009; de Vignemont, Tsakiris, & Haggard, 2005). Although there are some common principles in how partonomies are derived in different representational systems, there may not be one unique partitioning of the body into parts, but multiple partonomies subserving different functions.

The skin, for example, is a continuous undifferentiated sheet, but the primary somatosensory cortex has receptive fields that generally follow neuroanatomical divisions, such as finger and arm. Nonetheless, these categorical divisions are the result of our cumulative experiences and are highly malleable. They can become blurred under certain types of tactile experience. Braille readers, for example, who move multiple adjacent fingers simultaneously to read dot patterns raised on a surface show "smearing" of the neural representation of their digits. When asked to identify *which* finger is being touched in a psychophysical task, they make many more confusion errors between fingers, consistent with the topographical disarrangement of the cortex (Sterr et al., 1998a,b). These results are not due to a loss of sensitivity as demonstrated by the fact that the same people have lower tactile thresholds (i.e., higher sensitivity) than controls when asked to indicate merely *when* they have been touched (Sterr et al., 1998a,b). Similar results of cortical reorganization and digit mislocalization have been found for normal participants who have had finger and thumb simultaneously stimulated for an hour a day over a 4-week period (Braun, Schweizer, Elbert, Birbaumer, & Taub, 2000).

Action provides another basis for body part segmentation. Intentional action imposes a functional, as opposed to a spatial, organization to body parts. If I raise my arm, my hand follows. But if you tap me on my arm, this does not tell me anything about the sensation in my hand. So the tactile partonomy and motor partonomy appear to give rise to different representations of body parts. For example, tactile perception exhibits a categorical boundary effect. When two tactile stimuli are applied to a single body part (hand or arm), those points are perceived to be closer together than when the two points are presented across body parts (one point on the hand and the other on the arm). However, if participants are made to move their hands (by flexing and extending the wrist joint), the category boundary effect is attenuated. The distance between the hand and arm is perceived to be closer than when the parts were static (de Vignemont et al., 2009). Overall, then, action appears to unify discrete parts.

Different partonomies, thus, exist for different representational systems—visual, somatosensory, motor—each of which is more or less differentiated and more or less malleable. Nevertheless, in all the systems, joints appear to be landmarks for segmentation of the body. As Bermudez (1998, p. 156) argues: "Individual body parts are paradigmatically defined in terms of hinges. The forearm, for example, is the volume between the elbow and the wrist . . . Using hinges provides a nonarbitrary way of segmenting the body that accords pretty closely with how we classify body parts in everyday thought and speech." Or as Bloom (2000, p. 109) says: "objects are parsed into natural parts through a sensitivity to discontinuities in surface contour . . . A finger, for instance, is an excellent part because unpleasant as it is to think about—it is seen as having a potential separateness from the rest of the body, in that it can be cleanly severed." Vision provides a more nuanced, hierarchical structure, but in general the senses appear to be aligned in their default segmentational strategies.

So what can we predict about how languages label these perceptual parts? Figure 3.1 provides a good basis to lay out the possibility space for how languages name the body and its parts. Beginning with level A, the classic review articles on nomenclature for the body by Brown (1976) and Andersen (1978) propose that the geon on this level, i.e., the BODY, will be universally labeled.

The next level—level B—is often taken as the "basic level" as illustrated in this quote from Tversky (1989, pp. 993–994): "an extraterrestrial being, with a cognitive system similar to ours, may, in trying to comprehend Homo Sapiens, decompose a human first into head, trunk, arms, and legs." The notion of a basic level for parts (as opposed to kinds) is problematic, but has a certain utility here in order to test hypotheses about naming patterns cross-linguistically. For example, Andersen (1978) has claimed that all languages will label HEAD, TRUNK, ARM (and HAND), and LEG (and FOOT) (see also Brown 1976 and Wierzbicka, 2007). The content in parentheses is informative here, since it already suggests a variation in naming across languages—that "There is not always exact correspondence across languages for the reference of a given term" (Andersen, 1978, p. 351). At the next level, level C, we can ask whether languages name the UPPER-ARM, LOWER-ARM, HAND, UPPER-LEG, LOWER-LEG, and FOOT with distinct terms. According to the most simplistic hypothesis every visually segmented body part would be named in language. This could be read from claims such as made by Hoffman and Richards (1984): "It is probably no accident that the parts defined by minima are often easily assigned verbal labels" (p. 82).

In the next section, I will illustrate how languages differ in which parts they single out for naming. Nevertheless, naming of body parts is not unconstrained by the segmentations provided by perception, a point to which we will come back in the final section.

VARIATION OF BODY PART CATEGORIES ACROSS LANGUAGES

Before examining how well words across the world's languages map onto perceptually derived parts of the body, some thought needs to be given as to what sorts of linguistic units we wish to consider. At the heart of the problem is the fact that the same referential entity can be expressed in language in different ways. Within a single language choice of expression can convey a different perspective, for example, the family dog could felicitously be referred to as "*Rover, the family pet, our dog, Tim's pup, the rubbish-bin, the destroyer of shoes,* or even *the vacuum-cleaner*" (Clark, 1997, p. 2); however not all of these expressions are on par; rather some seem to be more fundamental or "basic."

For example, in Tarascan (spoken in Mexico) one common way to talk about body parts is through a set of suffixes. These provide information regarding the spatial relationship between objects and can be used to describe the location of an experience. For example, the verb root *p'ame* 'feel pain, ache' can be combined with different body part suffixes to express where the pain is felt, *p'ame-**cha**-ni* 'to feel pain in the throat', *p'ame-**k'u**-rha-ni* 'to feel pain in the hand(s)', *p'ame-**a**-rha-ni* 'to feel pain in the stomach', *p'ame-**ndi**-*ni 'to feel pain in the ear(s)', *p'ame-**t'a**-rha-ni* 'to feel pain in the leg(s)', *p'ame-**ŋarhi**-ni* 'to feel pain in the face or eyes', etc. (Friedrich, 1971; Mendoza, 2007). Tarascan is not unique in having such a set of body part morphemes. Body part verb suffixes appear in Totonac (Levy, 1999), and there are verb prefixes in North American languages for doing things 'by hand', 'by foot', etc.

One interesting thing about these body part suffixes is that they can have unusual referents. For example, the nose and forehead are conflated under a single suffix,*-ṇu* in Tarascan, which leads Andersen (1978) to claim that in this language the term for nose includes the forehead. But Tarascan also has a distinct noun that refers to the nose and another for forehead (Lathrop, 2007). In fact, there appears to be a dual semantic system for body part categorization in Tarascan and the other languages above: one that is expressed in nouns and another more schematic one that is expressed in grammaticalized morphemes. For a complete perspective of how languages categorize body parts, it would be important to consider these non-nouns also. For now, though, we ask the simpler question of how nouns in different languages categorize body parts, with the goal of examining whether they partition the body in the same way across languages. Literally descriptive expressions, such as the *right foot* or *the back of the knee,* in which the meaning of the whole is a direct combination of the meaning of the parts, will be excluded from consideration. Rather, the focus will be on expressions that are conventionalized. This would include complex expressions that are decomposable, but in which the whole meaning is not descriptive. For example, *underarm* in English is decomposable into *under* plus *arm*, but the meaning of this phrase is not predictable, as can be seen by the fact that it means ARMPIT in English but the cognate *onderarm* in Dutch means FOREARM.

Having established what words we will be comparing, the question then, is whether all languages have a term to refer to the parts at every level of the part hierarchy in Figure 3.1? Recall that Andersen (1978), Brown (1976), and others propose that (1) all languages will label level A, the BODY, (2) all languages should label HEAD, TRUNK, ARM, and LEG in level B because these are "basic" parts, and (3) UPPER-ARM, LOWER-ARM, HAND, UPPER-LEG, LOWER-LEG, and FOOT should receive distinct labels since they are parts defined by minima.

To be able to determine whether languages have terms for the body parts listed above, detailed descriptions of body part naming systems from geographically, genealogically, and typologically distinct languages are required. By ensuring a broad and diverse language sample, we can be more confident that any generalizations discovered hold beyond the particular languages studied. Detailed descriptions are required because consulting a dictionary or word list from a grammar will often not provide enough information to establish the exact meaning of a term. For these reasons, a concerted effort was made to provide detailed descriptions of body part terminology at the Language and Cognition group, Max Planck Institute for Psycholinguistics (Majid, Enfield, & van Staden, 2006). A team of field researchers used a standardized battery of linguistic tests to collect information about body part terminologies in a broad sample of languages (see Table 3.1). The languages studied were mostly from small-scale "traditional" societies (with the exception of Punjabi and American Sign Language). Researchers were experts on the languages: they were either speakers of the languages described themselves or had conducted long-term linguistic and ethnographic work on the language communities. There are two main advantages of this method of data collection: (1) because the same battery of tests is used in different languages, maximal comparability is ensured, and (2) because language experts conduct the study in different settings, language-specific nuances are more likely to be detected and can therefore be respected in the comparative endeavor.

In the following sections, I will draw primarily on these languages and consider each of the above proposed claims of body part naming in turn.

Proposal 1: All Languages Will Label the Body

This is not a universal. There are languages that do not have a term for BODY, the leftmost cylinder in Figure 3.1, on the highest level of the perceptual hierarchy. That is, there is not a

TABLE 3.1. Contributions to Special Issue of *Language Sciences* Parts of the Body: Cross-Linguistic Categorization[a]

Language	Country Spoken	Researcher
Jahai	Malaysia	Niclas Burenhult
Lao	Laos	N. J. Enfield
Kuuk Thaayorre	Australia	Alice R. Gaby
Yélî Dnye	Papua New Guinea	Stephen C. Levinson
Punjabi	Pakistan, India	Asifa Majid
Tiriyó	Brazil/Surinam	Sergio Meira
American Sign Language	United States	Jennie E. Pyers
Lavukaleve	Solomon Islands	Angela Terrill
Tidore	Indonesia	Miriam van Staden
Savosavo	Solomon Islands	Claudia Wegener

[a]Majid et al. (2006).

term that refers to the uniquely physical appearance—as opposed to the social dimension—of a person. This is the case in Tidore, a Papuan language spoken on the island of Tidore in the North Moluccas. In Tidore, the term *mansia*, which could be used to refer to level A, has a wider scope, meaning 'person' or 'human being' (van Staden, 2006). This appears to be a common pattern, and has been reported for Tiriyó, spoken in the northern Amazonia (Meira, 2006) and Kuuk Thaayorre, spoken on the west coast of Cape York, Australia (Gaby, 2006; see also Wilkins, 1996; Evans & Wilkins, 2001).

These reported counterexamples have been challenged by Wierzbicka (2007; cf. Goddard, 2001), who claim that in these languages, the term for 'person' is polysemous, with one sense referring to the physical body and another sense to the person. If we accept this argument, then the universal of naming of the body could be upheld and Proposal (1) would remain intact. Since this is a crucial point, let us consider the argument more closely.

Both Wierzbicka and Goddard suggest that if the same word has distinct interpretations in different syntactic constructions then that word must have distinct senses stored in the mental lexicon. But distinct interpretations can be generated on the fly (i.e., pragmatically generated), rather than stored as separate lexicalized entries. For the languages cited, there could be a general meaning corresponding roughly to 'person', with 'body' being understood within a specific situation. Ordinarily, however, a general interpretation would be sufficient for communication—for example, Evans and Wilkins (2001) describe how Arrente-English bilingual speakers translated a notice on diabetes, which in English read "all these things are bad for the body," to "thing this all bad *tyerrtye* (= 'person/body')." There is no need to specify further—a general interpretation is sufficient—since what is good/bad for the body is good/bad for the person. They conclude that "the distinction between the 'body' sense and the 'person' sense of Arrente *tyerrtye* becomes blurred (and is inconsequential for adequate comprehension)" (Evans & Wilkins, 2001, p. 502). Or to take an example from a different domain: When I use the word *bird* any member of the category could be meant (it is a general term), but a specific member might be inferred in the right context (e.g., *I saw a bird stick its head in the sand* = ostrich). This does not mean that the meaning of bird is polysemous. In the same way, we could say that *tyerrtye* is not polysemous.

Wierzbicka and Goddard argue that a polysemous interpretation is necessary unless a unitary definition—which can account for the range of the word's usages—can be provided. Critically, they require that the unitary definition should be a paraphrase in natural language.[2] This is the cornerstone of the Natural Semantic Metalanguage (NSM) approach, according to which all word meanings can be defined by a set of simpler words—"primes"—that are innate, universal, and themselves not definable (Wierzbicka 1972, 1996; Goddard, Chapter 4, this volume). This insistence on a single definition in natural language comes only from practitioners of NSM; it is not a generally accepted requirement. Meaning as reduction to simpler components is not widely accepted in the cognitive and linguistic sciences today; rather many take an "embodiment" or "simulation" viewpoint instead (e.g., Barsalou, 1999; see Kemmerer, Chapter 14, this volume). The fact is that in ordinary usage of language it may not be necessary to have a distinctly lexicalized expression for BODY, as the above example from Arrente demonstrates.

In Tidore there is no indigenous word for BODY but speakers can use the Indonesian loan word *badan* 'body' to specify the purely physical component. Whereas many would take the borrowing as evidence for an existing lexical gap in the language, Goddard (2001, p.15) argues that "some languages have borrowed terms for semantic primes, presumably replacing the earlier indigenous words." So, the argument goes, Tidore did have a term for body but just replaced it with a new word from Indonesian. It is hard to see why a single body part term would be borrowed (if there was no gap), particularly since other body part terms are indigenous (cf. Hale, 1994).

The crux of Wierzbicka and Goddard's objections to abandoning Proposal 1 is that within NSM no distinction is drawn between semantic and conceptual representations. The BODY is viewed as a conceptual universal and to deny that a language has a word meaning 'body' is to deny that the language community has the concept BODY. But this conclusion follows only if we conflate linguistic meaning with nonlinguistic representations; we can deny that there is a specific word with the semantics 'body' without denying that a person could entertain that concept. Speakers of languages such as Tidore, Tiriyó, and Kuuk Thaayorre do not have words for the body, but that does not entail that they do not have the concept BODY.

Proposal 2: All Languages Will Label the "Basic" Parts HEAD, TRUNK, ARM, and LEG

After the BODY, the next level of the hierarchy in Figure 3.1 has the major subdivisions of the body. This level may conceivably be thought of as the "basic" level, or the level with the most salient parts (Tversky, 1989). Do languages label each of these generalized cones? On one version of this hypothesis, there would be separate words for each of the cones present. No theorist predicts this, of course, since symmetrical parts are not expected to be lexicalized distinctly. Nevertheless, distinct words for the HEAD, TRUNK, ARMS, and LEGS may be expected to exist in every language of the world, but do not.

Jahai, an Aslian language of Malaysia (Burenhult, 2006), does not have a term for the head. The closest contender is the term *kuy*, which in every day discourse refers to the top part of the head, not the whole generalized cone. We may wonder if this is just an idiosyncratic lexical gap in this language, but the lack of a term for head seems to be consistently absent in many other Aslian languages, including Semelai, Mah Meri, and Ceq Wong (N. Kruspe, personal communication).

Jahai *kuy* (and its cognates in the languages above) is the closest equivalent to head because if someone is beheaded this would be the term that would be used to refer to the disembodied head. But this usage is most probably a case of metonymic extension since all other evidence confirms the word has a much narrower sense than HEAD. When speakers of Jahai, Semelai, and the other Aslian languages are asked to color in the head on a line drawing of a body they color only that part of the head that is covered with hair. Corpus evidence also supports a narrower sense for the head term. For example, when a Mah Meri speaker says *baci? koy*[3] 'look.through head' it means to look through someone's hair for lice or dandruff and *təc koy* 'cut head' means to cut someone's hair, **not** behead someone.[4]

Moving to the next "basic" part—the TRUNK—it appears that this body part is not highly salient for naming purposes. Many languages, such as Jahai, Tiriyó, Tidore, Punjabi, and Savosavo (a Papuan language spoken on the Solomon Islands), lack a distinct term for the trunk. Commonly—across languages—the same term is used for torso as for the whole body, as is the case in Yélî Dnye, for example. It is much rarer to find a distinct term for the torso alone, although it does happen (e.g., Kuuk Thaayorre *rerngk*). A closer look at the exact meaning of the trunk terms reveals further fine-grained differences between languages. For example, when speakers of Yélî Dnye are asked to construct a partonomy of the body, under the 'trunk' sense of *pââ* they include chest, belly, and buttocks (Levinson, 2006). English speakers, on the other hand, do not consider the buttocks to be part of the torso but part of the legs instead.

Terms for LEGS and ARMS likewise show considerable variation across languages. Lavukaleve (another Papuan language of the Solomon Islands) has one general term for ARMS and LEGS, thus categorizing together spatially discontinuous parts (Terrill, 2006). This term is somewhat reminiscent of English *limb*, but unlike English there is no other specific word to refer to the arm or leg, respectively.[5] Lavukaleve contradicts previous claims that the arm and leg are always given distinct terms (Andersen, 1978; Brown, 1976).

At the other extreme, Jahai has a very fine-grained categorization of the limbs. It also lacks terms for ARMS and LEGS at the "basic" level of the hierarchy. There are no superordinate terms, as found in Lavukaleve, but also no distinct terms for ARMS and LEGS separately. Jahai has a much more fine-grain categorization system for the limbs, as will be discussed in the next section.

For languages that do have terms for ARM and LEG, one issue that remains to be determined is whether these words refer to the whole geon in level B of Figure 3.1, or whether they have a more restricted range. The issue is whether HAND and FOOT are included in the referential range of ARM and LEG: Does *arm* in English end at the wrist and *leg* at the ankle, or do they extend to include the extremities? We'll come back to this issue in a later section.

To summarize, it appears that there is not a universal "basic level" for body parts that includes HEAD, TRUNK, ARMS, and LEGS. The closest equivalents to these terms can have differing extensional ranges, and the first level of categorization of the body can be more general, for example, collapsing the distinction between ARMS and LEGS or more specific, as in the Jahai system. The next section examines whether all parts defined by minima are named across languages.

Proposal 3: Parts Defined by Minima, e.g., UPPER-ARM, LOWER-ARM, HAND, UPPER-LEG, LOWER-LEG, and FOOT, Will Be Named in Languages

As just mentioned, there are languages, such as Jahai, that name the limbs at a fine level of granularity with separate terms for UPPER-ARM, LOWER-ARM, HAND, UPPER-LEG, LOWER-LEG, and FOOT. Recall that Jahai lacks a term for HEAD and TRUNK too. The language system favors naming at a finer level across the board for body parts (Burenhult, 2006). Granularity of naming is not always consistent within a language, however. In Hopi, for example, there are discrete terms for UPPER-LEG (*qá:si*), LOWER-LEG (*hóky̆a*), and FOOT (*kŭkŭ*), as in Jahai, but HAND-ARM receives a single label (*má:ʔa*), with no further elaboration (Swanson & Witkowski, 1977). Similarly, in Yélî Dnye there are more distinctions made for the lower body than the upper body.[6] But different again from Jahai and Hopi, while the UPPER-LEG is singled out for naming, the LOWER LEG and FOOT distinction is collapsed under a single term.

Perhaps the most salient discontinuities to be recognized at this level of granularity are those distinguishing HAND and FOOT. The discontinuities are as salient for their functional significance as for their perceptual distinctness: hands for manipulating objects; feet for walking. Two-thirds of the world's languages have a distinct word for HAND. But the remaining one-third does not make this distinction, collapsing HAND and ARM or HAND and LOWER ARM (Brown, 2005; Witkowski & Brown, 1985). Where HAND and ARM are not distinguished, the proper analysis may be that the whole upper limb geon in level B is named, but the HAND geon in level C is not distinctly labeled. Or to put it another way, the HAND geon in level C is not singled out as a distinct part in linguistic categorization. This is the case in Savosavo (Wegener, 2006). The word *kakau* is general over hand-arm. If someone says that they broke the *kakau* or that their *kakau* is dirty it could refer to any part of the geon in level B (see also Liston, 1972, on Serbo-Croatian).

As with body/person, Wierzbicka (2007) has argued that words such as *kakau* are polysemous between 'arm' and 'hand'; there are two distinct senses. Her argument is that since hand is a crucial concept required to explicate many other word meanings, such as *slap*, *stroke*, and *tear*, it must be universally lexicalized. In fact, to perform any of these actions requires not just the hand but the whole hand-arm.[7] Although English speakers associate such verbs with the hand (Maouene, Hidaka, & Smith, 2008), in motor terms the hand-arm is a single coordinated entity that achieves the action. Logically, then, there is no reason why the unified hand-arm could not be used to explicate "manual" actions. Or, where required, more specific parts could be recruited: *slap* could be done with the palm and *tear* and *stroke* with the fingers, etc. To be able to distinguish these accounts requires careful investigation and experimentation.

There are further subtle differences within languages that do distinguish HAND and ARM in precisely which perceptual geon is selected. Punjabi and Dutch speakers both have distinct terms for hand and arm (and foot and leg), but exhibit divergent intuitions about the relationship between these parts. Punjabi speakers are clear that the hand is not part of the arm and the foot is not included in the leg, but Dutch speakers are not so sure about this. Some Dutch speakers include the hand within the scope of the arm term whereas others exclude it (and likewise with the leg–foot). A similar situation exists in English. According to some tests hand appears to be included in the meaning of arm—one can say that *the hand is part of arm* or *an arm has a hand*. But according to other tests hand is not integral to the meaning of arm, as illustrated by the following example: A: *Did you find the arm?* B: *Yes, but the hand was missing*. B: *Yes, but the forearm was missing* (Cruse, 1986).[8]

PRINCIPLES OF BODY PART CATEGORIZATION ACROSS LANGUAGES

Although the same visual partonomy (i.e., the full set of "geons") in Figure 3.1 is perceptually available to speakers of different languages, not all of them—or even some subset—is universally singled out for the purposes of linguistic categorization. Lavukaleve and Jahai, for example, represent two very different solutions for how to refer to the arms and legs; in the former language only a general word exists to refer to the limbs, whereas in Jahai there are no "basic" terms. Instead UPPER-ARM, LOWER-ARM, HAND, UPPER-LEG, LOWER-LEG, and FOOT are all distinctly labeled. English has distinct terms for ARM and LEG versus HAND and FOOT, whereas Savosavo has only general words for ARM–HAND and LEG–FOOT. Speakers of Yélî Dnye have a distinct word for UPPER-LEG and another word for LOWER-LEG-FOOT. And so on. Naming at each one of the levels of the hierarchy seems to be an independent choice, as can be seen from Figure 3.2. Whereas Lavukaleve names at the "superordinate" level and Jahai at the "subordinate" level, speakers of Punjabi have conventionalized expressions at all levels.

Universals in body part naming are not to be found in the precise parts that are labeled then. But perhaps discontinuities provide constraints in how body part words can refer to the world. Rather than geons waiting to be labeled by body part words, perhaps words select possible extensions and the word's potential reference is delimited by the boundaries in the perceptual field. This would mean that body part terms may vary in the precise extension they have, but they should never defy the discontinuities provided by perception. This proposal is not an obvious fact. van Staden (2006), for example, argues that in Tidore the 'leg' term in Tidore, *yohu*, begins at the foot but ends three-quarters of the way up the thigh, where there is no perceptual discontinuity in the body image. Despite the lack of a natural boundary, there may be some other perceptually salient feature at play here. Perhaps Tidore speakers wear shorts or skirts that end at the mid-thigh and the clothing thus creates an artificial perceptual boundary. Alas this is not the case. Tidore speakers wear sarongs that come down at least to the knee. van Staden suggests that the boundary of *yohu* is not a physical one but a social or moral one—the boundary of *yohu* marks the beginning of the taboo genital area. So it is an empirical question as to what extent perceptual boundaries constrain the reference of body part words across languages.

Given that joints play such a fundamental role in our visual, tactile, and motor representations of the body, the possibility that they may not in our semantic representation of the body is striking. To investigate this possibility further and find the precise mapping of body part words onto the body, my colleagues and I used a very simple method to establish the extension of body part terms in different languages. We asked a number of speakers from a wide array of languages to color in a selection of body part terms from their own language (van Staden & Majid, 2006). Figure 3.3 shows the outcome of the coloring-in task when speakers of Dutch, Japanese, and Indonesian were asked to color in the arm.

It is important to know that the three languages differ in their naming patterns for the

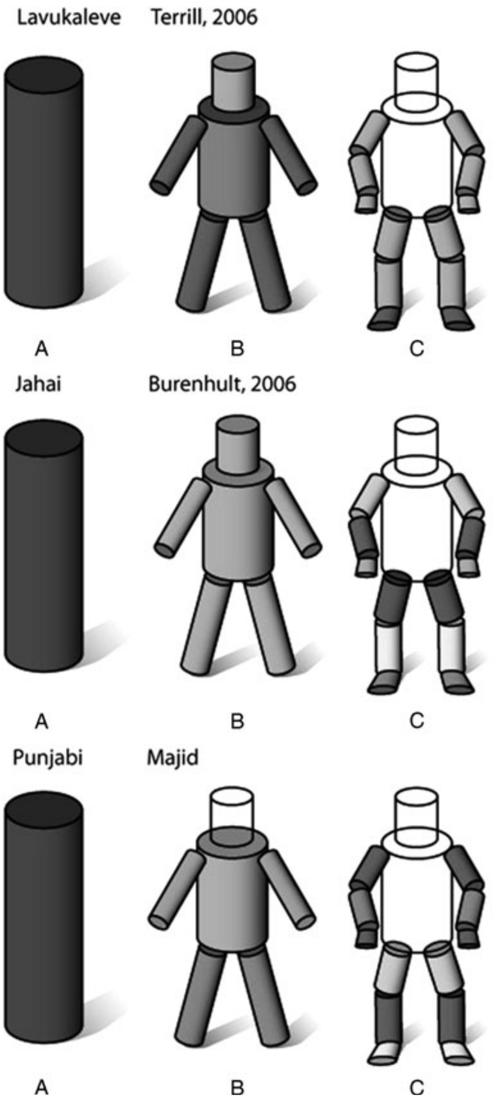


FIGURE 3.2. Grayscale rendition of Color Plate 1 illustrating three different languages and how they name parts of the body (A–C). See Color Plate 1 for interpretation. A gray geon means that there is no conventionalized means for talking about that body part. So, for example, in Jahai there is no word for HEAD, TRUNK, ARM, or LEG. Within a language, geons with the same color are referred to with the same word. Thus Lavukaleve speakers use *tau* to refer to ARM and LEG. Note that while Lavukalave names body parts at level B, Jahai names at level C, and Punjabi names at all levels suggesting that naming of geons at each level of the hierarchy is an independent choice.

limbs. Dutch, like English, has distinct words for HAND, ARM, FOOT, and LEG. Japanese distinguishes HAND from ARM but has a single term covering FOOT–LEG. Indonesian is less differentiating again and has a single term for HAND–ARM and another for FOOT–LEG. For all of these body parts, speakers colored in parts largely respecting the perceptual discontinuities provided by the joints. Most Indonesian speakers colored in from the fingertips to the shoulder joint; some colored in only the hand to the wrist, suggesting that hand is the primary meaning of *tangan* for them. Japanese speakers showed more variability in their coloring in.

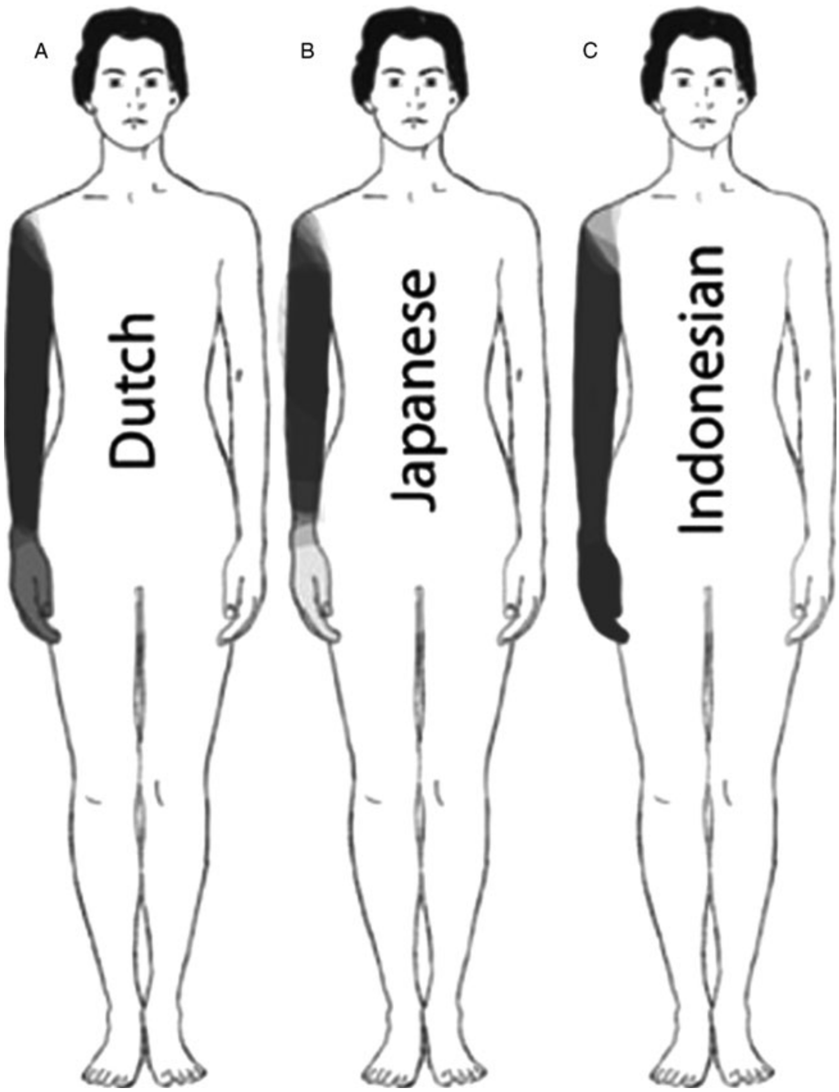


FIGURE 3.3. Grayscale rendition of Color Plate 2. See Color Plate 2 for interpretation. Eight Dutch, Japanese, and Indonesian speakers were asked to color in parts of the body. Their responses were then layered into a single image so that points of consensus could be viewed. The darker the image, the more speakers colored in that part of the body; the lighter the image, the fewer who included that part. These are the results when Dutch speakers were asked to color in the *arm*, Japanese speakers the *ude*, and Indonesian speakers the *tangan*.

Remember, Japanese distinguished HAND from ARM with distinct terms. In their coloring in, only one speaker included the HAND in the extension of *ude*, the term for ARM. Most colored from wrist to shoulder, one or two colored from elbow to shoulder, and one or two others showed no clear adherence to the joints as landmarks. Dutch speakers showed a split pattern in their coloring; half colored from fingertips to the shoulder and half from the wrist to the shoulder. So even though Dutch has a distinct term for HAND, many speakers think that the HAND is included in the reference of *arm*. The notable result from this is that despite variations in how many lexical distinctions are made in the language, speakers still respect the boundaries provided by the joints.

The composite pictures in Figure 3.3 graphically illustrate how participants respect boundaries at joints. These figures were created by superimposing all the pictures colored-in to the translation-equivalent of arm by speakers of the three languages. Remember that Dutch and Japanese have distinct terms for arm and hand, whereas Indonesian has a single hand–arm term. Regardless of this difference, speakers from all three languages colored up to the shoulder joint and down to either the wrist or the fingertips.

These results support the claim that joints constitute delimitation points for the extension of body part terms. Regularities in body part naming come not from which geons are selected for reference. Instead granularity and depth of naming of body parts differ across languages, with perception helping to provide constraints on the precise reference of the terms.

CONCLUSIONS

Faced with the variable mapping of language onto the body, it could be concluded that body part categories expressed in language are not formed on a perceptual basis. But this does not do justice to the mappings we observe. The variability that we see in the mappings of words to world is constrained by perception. Even though the precise segments selected by different languages vary—limb versus upper arm, lower arm, etc—the terms pick out constrained spaces. Visual discontinuities (and other perceptual cues) can help in categorizing body parts. At the same time, the cross-linguistic variability we see in the meaning of body part terms suggests that different parts of the body are open to interpretation, making the system of meaning associated with body part terms partially arbitrary. Individuals have to learn the linguistic conventions of their community to determine which of these discontinuities are relevant for the language they are learning. The externalized words, or signs, of a language help provide a way of coordinating individual representations so that of the myriad different solutions they could adopt, speakers within a community can come to a common solution for referring to the body (Belpaeme & Steels, 2005). Thus members of each language community must learn a system that in part is grounded in perception and in part is a function of local interpretation.

ACKNOWLEDGMENTS

Thanks to Barbara Malt and Phillip Wolff for critical comments and guidance. Special thanks to Niclas Burenhult for his input throughout and to colleagues and collaborators at the Language and Cognition group, Max Planck Institute for Psycholinguistics.

Notes

1. This conforms rather nicely to Wierzbicka's (2007) analysis of head, according to which ROUND is an important component of the meaning of *head*.
2. The criterion is stronger than this; the unitary meaning should be explicated in "an independently justified set of semantic primes" (Wierzbicka, 2007, p. 30).
3. This is cognate to the Jahai *kuy*.
4. Thanks to Nicole Kruspe for kindly providing these examples.
5. If necessary, phrasal expressions can be used to specify whether the arm or the leg is intended—*tau furi me* (literally 'lower/west limb') and *tau vego me* (literally 'upper/east limb'). (East-up and west-down are commonly conflated in the languages of the Solomon Islands.) But in common discourse, the general term *tau* is used without further specification.
6. These languages form counterexamples to the proposed universal by Andersen (1978) that languages make more distinctions for the upper body than the lower body.
7. See de Vignemont et al. (2009) for evidence that action unifies body parts.
8. Another example from Cruse (1986) serves to further illustrate this point. *There were burns on his fingers* entails *There were burns on his hand*. However, *there were burns on his hands* does not entail *There were burns on his arms*.

References

Andersen, E. S. (1978). Lexical universals of body-part terminology. In J. H. Greenberg (Ed.), *Universals of human language* (pp. 335–368). Stanford: Stanford University Press.

Bermudez, J. L. (1998). The paradox of self-consciousness. Cambridge, MA: MIT Press.

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

Bloom, P. (2000). How children learn the meaning of words. Cambridge, MA: MIT Press.

Braun, C., Schweizer, R., Elbert, T., Birbaumer, N., & Taub, E. (2000). Differential activation in somatosensory cortex for different discrimination tasks. *Journal of Neuroscience, 20*, 446–450.

Brown, C. H. (1976). General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature. *American Ethnologist, 3*, 400–424.

Brown, C. H. (2005). Hand and arm. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), The world atlas of language structures (pp. 522–525). Oxford: Oxford University Press.

Burenhult, N. (2006). Body part terms in Jahai. *Language Sciences*, *28*, 162–180.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1–34.

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64, 1–37.

Coslett, H. B., Saffran, E. M., & Schwoebel, J. (2002). Knowledge of the human body: A distinct semantic domain. *Neurology, 59*, 357–363.

Cruse, D. A. (1986). Lexical semantics. Cambridge: Cambridge University Press.

de Vignemont, F., Tsakiris, M., & Haggard, P. (2005). Body mereology. In G. Knoblich, I. M. Thorton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 147–170). New York: Oxford University Press.

de Vignemont, F., Majid, A., Jolla, C., & Haggard, P. (2009). Segmenting the body into parts: Evidence from biases in tactile perception. *Quarterly Journal of Experimental Psychology*, 62, 500–512.

Dennis, M. (1976). Dissociated naming and location of body parts after left anterior temporal lobe resection: An experimental case study. *Brain and Language*, *3*, 147–163.

Evans, N., & Wilkins, D. P. (2001). The complete person: Networking the physical and the social. In J. Simpson, D. Nash, M. Laughren, P. Austin & B. Alpher (Eds.), *Forty years on: Ken Hale and Australian languages* (Vol. 512, pp. 493–521). Canberra: Pacific Linguistics.

Friedrich, P. (1971). The Tarascan suffixes of locative space: Meaning and morphotactics. The Hague, Netherlands: Mouton & Co.

Gaby, A. R. (2006). The Thaayorre 'true man': Lexicon of the human body in an Australian language. *Language Sciences*, *28*, 201–220.

Goddard, C. (2001). Lexico-semantic universals: A critical overview. *Linguistic Typology*, *5*, 1–65.

Hale, K. (1994). Preliminary observations on lexical and semantic primitives in the Misumalpan languages of Nicaragua. In C. Goddard, & A. Wierzbicka (Eds.), Semantic and lexical universals (pp. 263–283). Amsterdam: John Benjamins Publishing Company.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.

Kemmerer, D., & Tranel, D. (2008). Searching for the elusive neural substrates of body part terms: A neuropsychological study. *Cognitive Neuropsychology*, *25*, 601–629.

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

Lathrop, M. (2007). Vocabulario del idioma Purépecha. Tlalpan, D.F.: Instituto Lingüístico de Verano, A. C.

Le Clec'H, G., Dehaene, S., Cohen, L., Mehler, J., Dupoux, E., Poline, J. B., et al. (2000). Distinct cortical areas for names of numbers and body parts independent of language and input modality. *NeuroImage*, *12*, 381–391.

Levinson, S. C. (2006). Parts of the body in Yélî Dnye, the Papuan language of Rossel Island. *Language Sciences*, *28*, 221–240.

Levy, P. (1999). From 'part' to 'shape': Incorporation in Totonac and the issue of classification by verbs. *International Journal of American Linguistics*, *65*, 127–175.

Liston, J. L. (1972). Semantic structure of body-part terms in Serbo-Croatian.1. Part-whole hierarchy. *Anthropological Linguistics*, *14*, 323–338.

Majid, A., Enfield, N. J., & van Staden, M. (2006). Parts of the body: Cross-linguistic categorization. *Language Sciences*, *28(2–3)*, 137–147.

Maouene, J., Hidaka, S., & Smith, L. B. (2008). Body parts and early-learned verbs. *Cognitive Science*, *32*, 1200–1216.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. New York: W. H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *200*, 269–294.

Meira, S. (2006). Tiriyo body part terms. *Language Sciences*, *28*, 262–279.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.

Meltzoff, A. N., & Moore, M. K. (1983). Infants imitate adult facial gestures. *Child Development*, *54*, 702–709.

Mendoza, M. (2007). Derivational resources in P'urhepecha: Morphological complexity and verb formation. *Acta Linguistica Hungarica*, *54*, 157–172.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9, 441–474.

Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience*, *17*, 543–553.

Shelton, J. R., Fouch, E., & Caramazza, A. (1998). The selective sparing of body part knowledge: A case study. *Neurocase*, *4*, 339–351.

Singh, M., Seyranian, G., & Hoffman, D. D. (1999). Parsing silhouettes: The short-cut rule. *Perception and Psychophysics*, *9*, 636–660.

Sirigu, A., Grafman, J., Bressler, K., & Sunderland, T. (1991). Multiple representations contribute to body knowledge processing. *Brain*, *114*, 629–642.

Steels, L., & Belpaeme, T. (2005). The semiotic dynamics of color. *Behavioral and Brain Sciences*, *28*, 515–529.

Sterr, A., Müller, M. M., Elbert, T., Rockstroh, B., Pantev, C., & Taub, E. (1998a). Changed perceptions in Braille readers. *Nature*, *391*, 134–135.

Sterr, A., Müller, M. M., Elbert, T., Rockstroh, B., Pantev, C., & Taub, E. (1998b). Perceptual correlates of changes in cortical representation of fingers in blind multifinger Braille readers. *Journal of Neuroscience*, *18*, 4417–4423.

Suzuki, K., Yamadori, A., & Fuji, T. (1997). Category-specific comprehension deficit restricted to body parts. *Neurocase*, *3*, 193–200.

Swanson, R. A., & Witkowski, S. (1977). Hopi ethnoanatomy: A comparative treatment. *Proceedings of the American Philosophical Society*, *121*, 320–337.

Terrill, A. (2006). Body part terms in Lavukaleve, a Papuan language of the Solomon Islands. *Language Sciences*, *28*, 304–322.

Tversky, B. (1989). Parts, partonomies and taxonomies. *Developmental Psychology*, *25*, 983–995.

van Staden, M. (2006). The body and its parts in Tidore, a Papuan language of Eastern Indonesia. *Language Sciences*, *28*, 323–343.

van Staden, M., & Majid, A. (2006). Body coloring task. *Language Sciences*, *28*, 158–161.

Wegener, C. (2006). Savosavo body part terminology. *Language Sciences*, *28*, 344–359.

Wierzbicka, A. (1972). Semantic primitives. Frankfurt/Main: Athenäum.

Wierzbicka, A. (1996). Semantics: Primes and universals. Oxford: Oxford University Press.

Wierzbicka, A. (2007). Bodies and their parts: An NSM approach to semantic typology. *Language Sciences*, *29*, 14–65.

Wilkins, D. P. (1996). Natural tendencies of semantic change and the search for cognates. In M. Durie & M. Ross (Eds.), The comparative method reviewed: Regularity and irregularity in language change (pp. 264–304). Oxford: Oxford University Press.

Witkowski, S. R., & Brown, C. H. (1985). Climate, clothing, and body-part nomenclature. *Ethnology*, *24*, 197–214.

Xu, Y., & Singh, M. (2002). Early computation of part structure: Evidence from visual search. *Perception and Psychophysics*, *64*, 1039–1054.

4 UNIVERSALS AND VARIATION IN THE LEXICON OF MENTAL STATE CONCEPTS

Cliff Goddard

In a global perspective, the language of mental state concepts displays a great deal of variation—much greater than imagined by most cognitive scientists. Almost all the words in the English lexicon of emotion and cognition are language and culture specific, i.e., they do not have exact meaning equivalents in many, perhaps most, other languages. This includes words for emotional and attitudinal states, such as *sad*, *angry*, *surprised*, *anxiety*, and *grief*, words for epistemic states and cognitive processes, such as *believe*, *doubt*, and *remember*, and words for ethnopsychological constructs, such as *mind*, *heart*, *psyche*, and *memory* (Russell, 1991; Wierzbicka, 1999; Harkins & Wierzbicka, 2001; Palmer, Goddard, & Lee, 2003; Shweder, 2004; Amberber, 2007; Schalley & Khlentzos, 2007).

Most cognitive scientists underestimate not only the scale of semantic variation across languages, but also the theoretical and methodological challenges it poses. In theorizing and discussing emotional states, they tend to take English for granted, effectively absolutizing the English lexicon of emotion and cognition (for example, assuming that words such as *sadness*, *anger*, and *surprise* represent natural psychological categories), while denying the same privilege to the lexical categories of other languages. At the methodological level, many researchers seem to regard the "problem of translation" as a mere nuisance that can easily be overcome by tagging indigenous concepts with English glosses. In the process, many cross-cultural studies are seriously flawed by inaccurate translations and concomitant "terminological ethnocentrism." I will return to these theoretical and methodological issues at some length toward the end of the chapter. In the meantime, however, there is a more pressing question. Presumably, to describe mental states at all, we must make some use of words, and these words must be drawn from some particular language. If adopting the lexicon of any language inevitably brings with it a culture-conceptual bias, does this not make some degree of terminological ethnocentrism unavoidable?

Fortunately it does not. Empirical evidence from cross-linguistic semantics indicates that despite the tremendous variation, a very small number of simple meanings connected with cognitive processes are shared across languages, namely (to list them using English terms) THINK, FEEL, WANT, and KNOW. These are among the 63 universal semantic primes identified over the past 20 years of cross-linguistic research by Anna Wierzbicka and colleagues in the Natural Semantic Metalanguage (NSM) framework (Wierzbicka, 1996; Goddard & Wierzbicka, 1994, 2002; Goddard, 2008a). The NSM system of meaning analysis, based on these shared simple meanings, is a powerful analytical tool for understanding the mental state concepts of different languages.

In the first two sections of this chapter, I provide an overview of NSM research and findings, with a particular focus on mental state concepts. In the next two sections I show how NSM techniques make it possible to reveal complex and culture-specific meanings in detail and in terms that are readily transposable across languages. Examples will include emotion terms, epistemic verbs, and ethnopsychological constructs in English, Chinese, Russian, and Korean. The next section discusses the relationship between linguistic meanings (word meanings) and cognition and elucidates the theoretical and methodological implications for cognitive science. The chapter concludes with the suggestion that people's subjective emotional experience can be shaped or colored to some extent by the lexical categories of their language.

CROSS-LINGUISTIC SEMANTICS

The fundamental methodological problem in lexical semantic analysis is the problem of metalanguage. To state the meaning of any word, it is necessary to use other words or equivalent symbols. How then can we avoid falling into circularity, or, just as bad, into an infinite regress? These problems were well known to seventeenth-century philosophers, including Descartes, Leibniz, Locke, and Pascal. Arnauld expressed a common view among them, when he wrote:

> I say it would impossible to define every word. For in order to define a word it is necessary to use other words designating the idea we want to connect to the word being defined. And if we again wished to define the words used to explain that word, we would need still others, and so on to infinity. Consequently, we necessarily have to stop at primitive terms which are undefined. [Arnauld & Nicole, 1996 (1662), p. 64]

Though they recognized the need for "primitive terms," the seventeenth-century philosophers did not make much progress toward identifying an optimal set. That task fell off the agenda until modern times, when it was revived by scholars working in the new discipline of linguistics. Structuralists and early generativists alike (e.g., Hjelmslev, 1961; Katz & Fodor, 1963) recognized that semantic analysis necessarily presupposes some inventory of semantic primitives (atomic concepts, elementary concepts, or the like). Neither the structuralist nor the generativist program made much headway, however, principally because they lost sight of the basic point that any viable set of semantic primitives must be constructed (or rather, discovered) from the meanings of ordinary nontechnical words. Technical words, logical symbols, etc. are not suitable for the purpose, because they necessarily require further definition or explanation. Of course, simply recognizing that viable semantic primitives must be meanings of ordinary nontechnical words does not get us very far, for there are thousands of such words and they cannot all be primitive. To arrive at a viable set of hypotheses about semantic primitives requires analytical work—wide-ranging experimentation with different sets of hypothetical primitives, in an effort to determine the minimal set of terms that would be adequate to explain the meanings of all the other terms. An additional important desideratum for any system of meaning analysis is universality, i.e., applicability not only to English, but to all the languages of the world. Ideally, semantic primitives should be composed of simple word meanings that are found not only in English, but in all other languages as well. Clearly, to identify such meanings—even to ascertain whether any such meanings exist—requires serious empirical research across many languages. This brings us to the program of semantic research now known as the Natural Semantic Metalanguage (NSM) approach.

Semantic Primes

Since it was inaugurated by Anna Wierzbicka (1972) in her book *Semantic Primitives*, linguists working in the NSM approach have been developing, testing, and refining a system of meaning analysis based on simple word meanings in natural language. Though it is still controversial in linguistics, the system can lay claim to being the best developed and productive on the contemporary scene. The current NSM model recognizes 63 indefinable semantic primes, which on current evidence appear in all languages. They are meanings such as SOMEONE, SOMETHING, DO, HAPPEN, WANT, KNOW, SAY, GOOD, BAD, BIG, SMALL, BECAUSE, and IF, and about 50 others. There is a large body of descriptive-analytical work in the framework, not only about English, but about Russian, Polish, French, Spanish, Malay, Japanese, Chinese, Korean, Ewe, East Cree, and many other languages (Wierzbicka, 1996; Goddard & Wierzbicka, 1994, 2002; Peeters, 2006; Goddard, 2008a; among other works).[1]

The NSM metalanguage is essentially a small standardized subset of natural language: a minivocabulary of simple word meanings, together with a minigrammar consisting of their associated grammatical properties. In effect, it represents the lexical and grammatical intersection of all languages. NSM researchers claim—and have sought to demonstrate in scores of published studies—that all the thousands of complex word meanings in the world's languages can ultimately be paraphrased into configurations of semantic primes. The NSM metalanguage is described in great detail in various publications, especially in Goddard and Wierzbicka (2002). Introductory overviews are available in standard sources such as the *Encyclopedia of Languages and Linguistics* (Brown, 2006) and *The Oxford Handbook of Linguistic Analysis* (Heine & Narrog, 2009). Goddard (1998) is an introductory textbook. The full NSM lexicon of universal semantic primes is set out in summary form in Table 4.1.

The inventory of semantic primes has been arrived at after a long period of experimentation with reductive paraphrase, which it is obviously impossible to recapitulate here. An example may be helpful, however. Consider the word *say*, in sentences such as *Mary said something to me*. How could we paraphrase the meaning of *say* in this context, using simpler words? An expression such as *verbally express* would not do, because terms such as *verbally* and *express* are more complex and difficult to understand than *say* is in the first place. The only plausible line of explication would be something like "Mary did something, because she wanted me to know something"; but this fails because there are many actions a person could undertake because of wanting someone to know something, aside from saying. Because of its resistance to reductive paraphrase, SAY is a good candidate for the status of semantic prime. Furthermore, SAY is clearly required for the

TABLE 4.1. Semantic Primes, Grouped into Related Categories[a]

I, YOU, SOMEONE, SOMETHING/THING, PEOPLE, BODY	Substantives
KIND, PART	Relational substantives
THIS, THE SAME, OTHER/ELSE	Determiners
ONE, TWO, SOME, ALL, MUCH/MANY	Quantifiers
GOOD, BAD	Evaluators
BIG, SMALL	Descriptors
KNOW, THINK, WANT, FEEL, SEE, HEAR	Mental predicates
SAY, WORDS, TRUE	Speech
DO, HAPPEN, MOVE, TOUCH	Actions, events, movement, contact
BE (SOMEWHERE), THERE IS, HAVE, BE (SOMEONE/SOMETHING)	Location, existence, possession, specification
LIVE, DIE	Life and death
WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT	Time
WHERE/PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE	Space
NOT, MAYBE, CAN, BECAUSE, IF	Logical concepts
VERY, MORE	Intensifier, augmentor
LIKE/WAY	Similarity

[a]Primes exist as the meanings of lexical units (not at the level of lexemes). Exponents of primes may be words, bound morphemes, or phrasemes. They can be formally complex, and they can have combinatorial variants (allolexes). Each prime has well-specified syntactic (combinatorial) properties.

explication of many other lexical items involving speaking and communication, especially speech–act verbs, as well as many discourse particles, and, on available evidence, exponents of SAY can be found in all languages. (On the other hand, if we take a word such as *ask*, as in *Mary asked me something*, it seems readily paraphrasable in simpler terms, including SAY, WANT, and KNOW: "Mary said something to me because she wanted to know something; she wanted me to say something because of this." Because it is decomposable into a combination of simpler terms, *ask* would be a nonstarter as a semantic prime.)

As mentioned, there is a substantial body of cross-linguistic research on how semantic primes manifest themselves in languages from different locations, different language families, and different linguistic types. In particular, "whole metalanguage" studies have been carried out for English, Polish, Russian, Spanish, French, Amharic, Lao, Malay, Mandarin Chinese, Mbula, Korean, and East Cree (Goddard & Wierzbicka, 2002; Peeters, 2006; Goddard, 2008a). These studies have been conducted by linguistic specialists in the language concerned (in many cases, native-speaker linguists). The evidence indicates that semantic primes are lexical universals in the sense of having an exact translation in every human language (the term "lexical" as used here includes short phrases and bound morphemes, as well as words proper).

The robustness of the NSM primes is dramatized when they are compared with various impressionistically basic—but non-primitive—items of English vocabulary such as "go," "eat," "hot," "tree," and "bird." In most cases, any possibility that such items enjoy universal status in the world's languages collapses on even a small sample of languages (Goddard, 2001b). A few brief examples may be useful. Many languages, including European languages, lack an exact equivalent of "go" because they distinguish between motion via a vehicle (horse, etc.) versus motion on foot. Languages such as Kalam, Warlpiri, and Shanghainese do not have equivalents of "eat" because they use a single verb to cover "ingesting," without distinguishing between eating and drinking. French and Spanish lack exact equivalents of "hot" because their nearest words (*chaud*, *caliente*) apply to what in English would be described as "warm," as well as "hot." Many Australian languages do not distinguish between what are termed in English "trees" and "bushes." In sharp contrast, however, semantic primes appear to have exact equivalents in all languages. This means that they provide a kind of stable lingua franca for cross-linguistic semantics, in the following sense: If we can explicate the content of any language-specific word into semantic primes, we can be sure that the explication can be freely transposed across languages without distortion of meaning.

Semantic Primes for Mental States

For the purposes of this chapter, I will concentrate on four semantic primes that can be seen as preeminently "cognitive" in character—namely, THINK, FEEL, WANT, and KNOW. Table 4.2[2] shows how these primes are realized in six languages from five language families: Indo-European (Spanish, Polish), Austronesian (Malay), Niger-Congo (Ewe), Tai-Kadai (Lao), and Sinitic (Mandarin Chinese). A word that realizes a particular prime in a given language is known as an "exponent" of that prime.

TABLE 4.2. Exponents of Cognitive Primes in Six Languages[a]

	Spanish	Polish	Malay	Ewe	Lao	Chinese
THINK	*pensar*	*myśleć*	*fikir*	*súsú*	*khùt*	*xiǎng*
FEEL	*sentir*	*czuć*	*rasa*	*se le làme*	*huu4.sùk2*	*gǎnjué*
WANT	*querer*	*cheieć*	*mahu*	*dí*	*jaak5*	*yào*
KNOW	*saber*	*wiedzić*	*tahu*	*nyá*	*huu4*	*zhìdao*

[a]See Note 2 for sources of the data in this table.

Many (perhaps most) words in any language do not have one single meaning, but rather a set of interrelated meanings (a situation known as lexical polysemy), so it is not uncommon for exponents of the same prime to have additional meanings (polysemic extensions), which can differ from language to language.[3] I will now give thumbnail sketches of THINK, FEEL, WANT, and KNOW, specifying for each one the range of grammatical contexts or syntactic frames in which it can occur. It is important to know the syntactic frames for each prime because they govern how the primes can be combined into phrases and sentences. Evidence reported in the collective volumes mentioned earlier indicates that all the grammatical frames described below for the individual primes are universally available.

THINK. Exponents of semantic prime THINK commonly have polysemic extensions such as "worry," "long for," "count," and "intend."[4] The minimal frame for THINK has a "topic," as in (1). A substantive complement can be added, as in (2). In a third and very productive frame, THINK can take a "quasiquotational" complement introduced by "like this," as in (3). Finally, a *that*-complement is possible if the verb is accompanied by a temporal expression, such as "at this time" or "now," tying it to a particular time.

(1) someone thinks about something/someone
(2) someone thinks something (good/bad) about something/someone
(3) someone thinks like this: " – – "
(4) at this time (or: now), someone thinks that [——]s

[English permits the use of *think* in a general "opinion" sense, but this is a peculiarity of English, not shared by many other languages of the world (Goddard, 2003; Goddard & Karlsson, 2008).]

FEEL. Exponents of semantic prime FEEL commonly have polysemic extensions to meanings such as "taste" and "smell," and to "hold an opinion." Sometimes the exponents of FEEL and HEAR are identical in form, or else the exponent of FEEL is formally related to "hear" in some other way. Three basic syntactic frames are posited for FEEL, as shown in (5)–(7).[5]

(5) someone feels like this
(6) someone feels something good/bad
(7) someone feels something good/bad toward someone else

FEEL is the semantic foundation for both emotion words (*sad, angry, excited,* etc.) and sensation words (*hungry, thirsty, hot, cold,* etc.). The meanings of emotion words also involve cognitive verbs such as THINK and WANT, whereas the meanings of sensation words involve these and, in addition, the prime BODY.

WANT. Exponents of semantic prime WANT often have polysemic extensions to meanings such as "like" and "love," or to "seek." The simplest frame has a substantive complement, as in (8). A distinctive property of WANT is the difference between the complement types shown in (9), on the one hand, and in (10a) and (10b), on the other. In the former the complement verb has no subject, whereas in the latter it does; compare "I want to do something" with "I want *you* to do something." Many languages employ different grammar in these two types of clausal complements. In Spanish, for example, the complement of *querer* WANT in a sentence such as (9) takes an infinitive verb, but in a sentence such as (10a) it takes a finite subjunctive verb and is introduced by *que* "that"; compare: *Quiero ir* [I-want to-go] and *Quiero que tu vayas* [I-want that you go (subjunctive)]. In some languages, different lexical exponents (allolexes) of WANT are used in these contexts (Harkins, 1995).

(8) someone wants something
(9) someone wants to do (know, say, . . .) something
(10a) someone wants someone else to do (know, say, . . .) something
(10b) someone wants something to happen

KNOW. Some languages have two exponents of KNOW, one for noun-phrase complements and one for sentential complements, as with German *kennen* and *wissen*. Leaving aside the issue of "knowing a person" (Wierzbicka, 2002), one syntactic frame for KNOW has a substantive complement with an optional "topic" (realized in English by a prepositional phrase introduced with *about*), as in (11). KNOW can also take a sentential complement (a so-called *that*-complement), as in (12). Finally, there is a set of frames focusing on the category of knowledge, as illustrated in (13a)–(13c).

(11) someone knows something (about someone/something)
(12) someone knows that [——]$_S$
(13a) someone knows *who* did it
(13b) someone knows *what* happened
(13c) someone knows *where* it happened

This completes the sketch of the "cognitive" meanings in the NSM metalanguage. For a fuller account, see Goddard (2007). The complete metalanguage of course includes another 50 or so elements, all of which have their own grammatical properties.

A common reaction of those encountering the NSM system for the first time is surprise at the small size of the posited metalanguage and scepticism that such a restricted set of terms would be sufficient to explicate the numerous mental state concepts in the world's many languages, and to capture the subtle differences between them. In the remainder of this chapter, I want to answer some of these doubts, while at the same time showing that many English terms that are regarded as unproblematical in mainstream cognitive science are in fact semantically complex and deeply culture specific.

CROSS-LINGUISTIC VARIATION IN EMOTION TERMS: ENGLISH AND CHINESE

Background

Following Paul Ekman and colleagues (cf. Ekman, 1992, 2004; Ekman & Davidson, 1994; among other works), many psychologists believe in the existence of a small set of biologically inbuilt human emotions, which are usually designated with English words such as *happiness, sadness, fear, anger,* and *surprise*. Weighing against these claims, evidence from linguistic anthropology, psychological anthropology, and cultural psychology attests to the cultural variability of emotional experience (Briggs, 1970; Harré, 1986; Kitayama & Markus, 1994; Lutz, 1988; Rosaldo, 1980; Schwartz, White, & Lutz, 1992; Shore, 1996; Shweder, 1991, 2004; Shweder & Haidt, 2000; White & Kirkpatrick, 1985).

As far as word meanings (lexicalized concepts) are concerned, research in cross-linguistic semantics shows that emotion meanings seldom have exact equivalents across languages, and often diverge very considerably (Wierzbicka, 1992, 1999; Harkins & Wierzbicka, 2001; Enfield & Wierzbicka, 2002; Goddard, 1996; Hasada, 2000). I will explore the situation with *sadness* and its (non)equivalents in Chinese, but first let me mention several other examples. German *Angst* is substantially different from English *fear* (Wierzbicka, 1999, Ch. 3). Malay has no equivalent for English *surprised*, because the closest words it has always imply something bad, whereas English *surprise* can be bad, good, or neutral (Goddard, 1997). Many languages do not have any word exactly equivalent to English *happy*; its closest counterpart in Russian, for example, implies a more intense emotion (Wierzbicka, 2004).

I will now discuss several studies about "sadness" and related words, in an effort to illustrate how the NSM methodology of reductive paraphrase makes it possible to pin down the precise similarities and differences in the lexical semantics of emotion words. It will not be possible to provide full justifications for the content of every explication, but this is not necessary for the main purpose of the exercise. In general terms, the primary evidence for linguistic semantics comes from native speaker interpretations of the use of linguistic expressions in context (including their entailments and implications), from naturalistic observation of language in use, from direct consultations with native speakers, and from the distribution of linguistic expressions, i.e., patterns of usage, cooccurrence, and collocation.

English *Sad* and *Unhappy*

It will be helpful to introduce the basic approach by way of English examples. What is the difference between *sad* and *unhappy*? Conventional dictionaries cannot shed much light on this question, because their definitions are flawed by circularity, reliance on more obscure terms than the ones being defined, and other lexicographic problems. For example, the *Australian Oxford Dictionary* defines *sad* as "unhappy, feeling sorrow or regret," *unhappy* as "not happy, miserable," and *miserable* as "wretchedly unhappy or uncomfortable." On this account, the words *sad* and *unhappy* would be pretty much the same in meaning, but this does not tally with the intuitions of native speakers or with the evidence of usage. There are contexts in which we could use one word but not the other, as shown in (14), or in which the choice of one word or the other would lead to different implications, as suggested by the examples in (15). These examples, and the explications that follow, are based on Wierzbicka, (1999, pp. 60–63).

(14) *I miss you a lot at work.... I feel so sad (*unhappy) about what's happening to you.* [said to a colleague in the hospital who is dying of cancer]
(15a) *I was feeling unhappy at work.* [suggests dissatisfaction]
(15b) *I was feeling sad at work.* [suggests depression, sorrow, etc.]

The basic NSM technique for explicating emotion meanings depends on linking a feeling (usually a good feeling or a bad feeling) with a characteristic or prototypical cognitive scenario involving thoughts and wants (Wierzbicka, 1972). The scenario serves as a kind of reference situation by which the nature of the associated feeling can be identified. For example, *joy* is a very good feeling linked prototypically with the thought "something very good is happening now"; *remorse* is a bad feeling linked with the thought "I did something bad." The content of the prototypical scenarios can be—and often is—quite detailed, although it can be spelled out entirely in semantic primes. The prototypical scenario helps us to explicate the meaning of any particular emotion word.

Consider the explication in [A] below for English *sad*. It says that when we say that someone feels *sad*, we are saying that this person feels something bad and we are indicating the quality of the feeling by appealing to how people can feel when they are thinking in a certain way, namely, as set out in section (c) of the explication (the prototypical cognitive scenario). In the case of *sad*, the prototypical scenario involves awareness that "something bad happened" (not necessarily to oneself), which the prototypical experiencer did not want to happen but is prepared to accept, in the sense of recognizing that he or she cannot do anything about it (an attitude akin to "resignation"). This explication is compatible with the wide range of use of *sad*; for example, that I may feel *sad* when I hear that my friend's dog died, or when I think about some unpleasant bickering in my workplace.[6]

[A] *Someone X felt sad*

a. someone X felt something bad
b. people can feel something like this when
 they think like this:
c. "I know that something bad happened
 I don't want things like this to happen
 I can't think like this: I will do something
  because of it now
 I know that I can't do anything"
d. this someone felt something like this

This format of explication enables subtle differences in meaning to be modeled, as we can see when we turn to the question of how to capture the meaning of *unhappy*. Consider some of the ways in which feeling *unhappy* differs from feeling *sad*. Perhaps the most significant difference is that *unhappy* has a more personal character: We can be *sad* because of bad things that have happened to other people, but if we feel *unhappy*, it is because of bad things that have happened to us personally. Relatedly, feeling *unhappy* focuses on some thwarted desires. The attitude is not exactly active, but it is not passive either. It suggests something like dissatisfaction, rather than resignation. These properties are modeled in the different, but related, prototypical cognitive scenario for *unhappy* given in explication [B]. Notice that the final component—component (d)—is a little more elaborate than the corresponding component for *sad*. It implies not only that the experiencer feels as we would imagine people would feel when having the kind of thoughts spelled out in the prototypical cognitive scenario, but that the experiencer actually had some thoughts like these. To put it in more technical language, feeling *unhappy* implies a more cognitively active state than feeling *sad*. (Although we can say *I feel sad, I don't know why*, it would be a little odd to say *I feel unhappy, I don't know why*.)

[B] *Someone X felt unhappy*

a. someone X felt something bad
b. people can feel something like this when they think like this:
c. "some bad things happened to me
   I wanted things like this not to happen to me
   I can't not think about it"
d. this someone felt something like this, because this someone thought like this

Explications [A] and [B] are phrased entirely in semantic primes. Needless to say, they hinge critically on the primes FEEL, THINK, WANT, and KNOW (as well as on others, such as SOMEONE, PEOPLE, BAD, HAPPEN, DO, LIKE, VERY, CAN, and NOT).

Now let us apply a similar approach to explicating emotion concepts of another language.

No Basic-Level Equivalent to *Sadness* in Chinese

Ye (2001) studied the meaning of three Chinese words that seem to partially overlap with the English concept of *sadness*, while differing from it significantly in other respects (cf. Russell & Yik, 1996). I will summarize her key claims with respect to two of these words: *bei* "tragic fatalistic sadness" and *chou* "confused sadness/worry/melancholy" (tones not indicated; Ye's glosses). There is a third salient word, *ai* "ethical and altruistic grief, mourning," which I will not consider here.

Ye's semantic analysis of *bei* is shown in [C] (slightly modified). Comparing it with the explication for *sad* in [A], it can be noted that both are bad (or in the case of *bei*, very bad) feelings triggered by the recognition that something bad has happened; but after that the prototypical cognitive scenarios diverge markedly. *Bei* has a distinctly fatalistic and pessimistic tone. The experiencer of *bei* confronts the apparent certainty that "after this, good things will not happen anymore" and that this cannot be prevented. The experiencer's helplessness is not seen as narrowly personal, because it appears set in a wider perspective ("no one can do anything when things like this happen") which suggests people's powerlessness before the laws of nature. As might be expected, the feeling of *bei* is apt to be evoked by the contemplation of the inevitability of aging and death.

[C] Someone X felt *bei* ("tragic fatalistic sadness")

a. someone X felt something very bad
b. people can feel something like this when they think like this:
c. "something bad happened now
   I know that after this good things will not happen anymore
   I don't want things like this to happen
   because of this, I want to do something if I can
   I know that I can't do anything
   because I know that no one can do anything when things like this happen"
d. this someone felt something like this, because this someone thought like this

In her commentary, Ye (2001) characterizes the feeling tone of *bei* as "poignant bitterness, pain and somberness," and she does not hesitate to link it with traditional Chinese views of life and the world. The emotion of *bei*, she points out, has long been linked in Chinese literature and poetry with the season of *qiu* "autumn," which suggests a cyclic decline in nature. "*Beiqiu* (lit. '*bei* autumn'), which is a lexicalized item, has been one of the most important and enduring themes in Chinese literature...Bleak and desolate autumn scenes suggest unavoidable ageing and death" (Ye, 2001, pp. 370–371).

Clearly, Chinese *bei* is far from identical to English *sad*, but is there no better candidate? Ye (2001) also presents an analysis of *chou*—a much more personal and everyday word than *bei*, and one that in some contexts seems to overlap with uses of English *sad*. Her account makes it quite clear, however, that *chou* is at least as close to English *worry* as to *sad*. Dictionaries more frequently gloss it as "worry, be anxious" than as "sadness, sorrow." According to Ye, a person experiences *chou* when confronting a personal predicament that is forced on him or her by the circumstances: "It is a struggle between wanting to find a solution to change a situation and not knowing what to do, leaving the experiencer caught in a dilemma—wanting to overcome a difficult situation, yet not finding a solution" (2001, p. 379).

From Ye's explication in [D], we can see a certain affiliation with *unhappy*, namely, that *chou* represents a bad feeling triggered by the thought that "something very bad is happening to me" (compare with *unhappy*: "some bad things happened to me"). In its future-orientation, however, and the lack of anything resembling resignation, *chou* is very different from both *sad* and *unhappy*. Rather, the experiencer is in a state similar to that suggested by English words such as *worried* and *preoccupied*, insofar as he or she wants to do something but cannot think what to do, and, moreover, cannot stop thinking about it.

[D] Someone X felt *chou* ("confused sadness/worry/melancholy")

a. someone X felt something bad
b. people can feel something like this when they think like this for some time:
c. "something very bad is happening to me
   before this, I did not think that this would happen to me
   I don't want things like this to happen to me
   because of this, I want to do something if I can
   I don't know what I can do
   I can't not think about this all the time"
d. this someone felt something like this, because this someone thought like this

Though not as deep or philosophical as *bei*, the word *chou* is associated with a rich and vivid phraseology in the Chinese language. Both words, but particularly *bei*, are culturally very salient. Many texts in the traditional Chinese philosophical-literary canon about emotion include *bei* among the "basic emotions" (Russell & Yik, 1996, pp. 82–83). Ye (2001, 2006) stresses that what is considered basic to English speakers is quite different from what seems basic to Chinese people.

It seems clear that Chinese *bei* and *chou* have no close equivalents in English, and conversely, that *sad* and *unhappy* have no close equivalents in Chinese (at least, not at the level of basic vocabulary items).[7] Another language that has been studied in close semantic detail in relation to "sadness" is Russian. Although it has two words (*grust'* and *pečal'*) usually translated into English as *sad*, Wierzbicka (1998) demonstrates that neither is an exact equivalent of the English word. Similar demonstrations could be adduced for many other languages and for many other areas of the emotion lexicon.

Emotion Terms Are Culture Specific

Research in cross-linguistic semantics shows definitively that emotion terms are semantically complex, and that the meanings of emotion terms in the world's languages are culture specific. They represent local interpretations, local construals, of how people can feel in response to particular cognitive and social scenarios.[8] As Wierzbicka (1999, p. 15) puts it: "Every language...has lexically encoded some scenarios involving both thoughts and feelings and serving as a reference point for the identification of what speakers of this language see as distinct kinds of feelings." The differences between these different culture-specific ways of thinking about emotions are not, however, incommensurable—far from it. They can be explicated quite precisely in terms of shared conceptual universals (semantic primes), including FEEL, THINK, WANT, and KNOW.

At this point, a common reaction is: "All this is about emotion words, but words are just words, not emotions as such." Though it is of course true that emotion words are not emotions, I will argue later in the chapter that this response misses the point that we need to understand emotion words much better if we are to have any chance of understanding emotions "as such." In particular, we need to avoid taking English-specific lexical categories—with their English-specific semantics—as default models for human beings in general, and we need to understand better how language-specific lexical categories may enter into people's thinking about their subjective experience.

I now want to briefly canvass another lexical domain connected with mental states and experiences, and to show that in this domain too, there are substantial semantic differences between languages.

CROSS-LINGUISTIC VARIATION IN EPISTEMIC VERBS: ENGLISH AND RUSSIAN

The term "epistemic verb" refers to verbs such as English *believe*, *doubt*, *assume*, and *suppose* (they are also known as propositional attitude verbs). The most important of these verbs in English is *believe*, with its related noun *belief*. In philosophy in particular, but also in cognitive psychology, the importance of the notion of *beliefs* can hardly be overestimated, given the pervasive discourses about "belief states" and the distinction between *knowledge* and *belief*. What usually goes unrecognized is that *believe* is highly English specific in its meaning. I will demonstrate this by comparing English with Russian, based on work by Wierzbicka (2006, Ch. 7) and Gladkova (2007), respectively. For a comparable cross-linguistic study of English and Swedish, see Goddard and Karlsson (2008).

English *believe* is a polysemous word that occurs in several grammatical frames. I will confine myself to the grammatical frame in which it takes a *that*-complement, i.e., *believe that*. . ., as in the examples in (15) below. Roughly speaking, *to believe that* conveys a considered conviction or commitment. It also conveys a certain gravitas, as shown by the fact that it can collocate with adverbs such as *strongly* (e.g., *I strongly believe that* . . .).[9]

(15a) *I believe that they shouldn't have the vote.*
(15b) *I believe that ultimately we're going to have to get ourselves into a position where we* . . .
(15c) *I believe that that is the key factor to being successful in this particular HEO role.*

Wierzbicka (2006, pp. 216–218) explicates *believe that* . . . as in [E]. The first component expresses something like a considered thought, using a *that*-complement of the semantic prime THINK. Then follows acknowledgment or awareness of the possible existence of another point of view ("I know that someone else can think not like this"). Subsequent components express the speaker's apparent confidence that he or she can provide some kind of justification for thinking this way (for example, evidence that the proposition is true, that it has been vouched for by a reliable source), and finally, the speaker's apparent confidence that he or she can explain or defend the value of thinking like this.[10]

[E] *I believe that – –:*

a. when I think about it, I think that – –
b. I know that someone else can think not like this
c. I can say why I think like this
d. I can say why it is good if someone thinks like this

In her semantic study of propositional attitude verbs in Russian, Gladkova (2007) shows that the Russian word *sčitat'*, frequently glossed as "believe," does not correspond exactly to English *believe*. Actually, as shown below, there are several Russian verbs that could be used to render the English meaning, but *sčitat'* is the most common:

Ja sčitaju, čto . . .	"I firmly believe/consider that . . ."
Ja verju, čto . . .	"I believe/trust that . . ."
Ja ubežden, čto . . .	"I am convinced/sure that . . ."

Gladkova (2007) quotes the eminent Russian semanticist Jurij Apresjan on the kind of mental processes involved in *sčitat'* (Apresjan, 2000, p. 149):

> serious conditions are required to develop an opinion (that which we *sčitaem* [consider]). An opinion is usually the result of a fairly long and thorough process of consideration of all observable facts (note the original idea of *sčet* [counting] which is present in *sčitat'*), weighing up other possible interpretations of them and selecting [by an act of will] the interpretation which best accords with the accumulated personal experience of the subject and which he is prepared to uphold as correct.

From an English point of view, *sčitat'* can seem decidedly egocentric. It presents an opinion about which there is no longer any room for doubt, and does not allow for or envisage other credible options and opinions. Gladkova (2007, p. 75) proposes the explication in [F]. Components (e) to (g) portray a strong determination to stick to the view expressed. (Notice also that *I believe that* is more "open" in the sense that it has the component "I can say why I think like this," as if the person is expecting to be asked to justify his or her belief.[11])

[F] *Ja sčitaju, čto – –:*

a. when I think about it, I think that – –
b. I have thought about it for some time
c. I have thought about things like this before
d. I want to think like this
e. I know why I want to think like this
f. I don't want to think about it in any other way
g. it is good to think about it like this

One striking phraseological difference between *sčitat'* and *believe* is that the former cannot be qualified (strengthened) with intensifiers. That is, although it is perfectly normal to say in English "I strongly believe that . . ." and "I firmly believe that . . .," *sčitat'* cannot be used with any of the words that can intensify opinions in Russian—*gluboko* "deeply," *sil'no* "strongly," or *tverdo* "firmly." Sentences such as **Ja gluboko sčitaju, čto* . . . and **Ja sil'no sčitaju, čto* . . . are ungrammatical. This is because the verb *sčitat'*, with its component—"I don't want to think about it in any other way"—expresses a level of "belief" that cannot be extended further. Gladkova (2007) links the directness and "absoluteness" of *sčitat'* to the value Russians place on forcefully speaking their mind and on *govorit' pravdu* "telling the truth" for its own sake (cf. Wierzbicka, 2002).

This brief exploration shows again that the apparently basic words of your home language need not have equivalents in other languages. *Believe* has no precise equivalent in Russian, and *sčitat'* has no precise equivalent in English.

CROSS-LINGUISTIC VARIATION IN ETHNOPSYCHOLOGICAL CONSTRUCTS: ENGLISH AND KOREAN

Most languages have some kind of nominal expressions designating nonphysical parts of a person, akin to English *mind, heart, soul,* and *spirit*. There is a significant NSM literature on the cross-linguistic semantics of such ethnopsychological constructs. Perhaps the single most discussed example is Russian *duša* (roughly) "soul" (cf. Wierzbicka, 1992, 2005; cf. Pesmen, 2000), but other work has examined English *heart*, French *âme* and German *Seele* (Wierzbicka, 1992, pp. 55–59), Malay *hati* (Goddard, 2001a, 2008b), Japanese *kokoro* (Hasada, 2000, pp. 115–116), and Korean *maum* and *kasum* (Yoon 2006). It is clear from this literature that *mind* is an English-specific construct, without precise equivalents even in European languages such as French, German, and Russian. To date, however, this fact has had a negligible impact on academic discourse, and *mind* continues to be a key word in philosophy, psychology, and cognitive science. The following section presents NSM semantic analyses for English *mind* and *heart*, followed in the subsequent section by an analysis of the Korean concept of *maum*. Again, my purpose is two-fold: to demonstrate the complexity and culture specificity of such concepts and to demonstrate that the NSM reductive paraphrase technique can resolve the complexity into simple cross-translatable terms.

English *Mind* and *Heart*

From a cross-linguistic point of view, English *mind* is unusual in its "rationalistic" character, i.e., its focus on thinking and knowing, to the exclusion of feeling. Feelings, especially feelings toward other people, are allocated to another ethnopsychological location, namely, the *heart*. Compare the implication of the expressions *a good mind* and *a good heart*. The explication in [G] is closely based on Wierzbicka (1992, 2005). According to this explication, the word *mind* suggests a dualistic way of thinking about a person, in which the *mind* is an invisible counterpart of a person's body, on account of which people can think and can know things, and in which something happens when a person thinks.

[G] *someone's mind*

a. one of two parts of this someone (one part is the body, this is the other part)
b. people cannot see this part
c. because people have this part, people can think about things
d. because people have this part, people can know things
e. when someone thinks about something, something happens in this part

As for English *heart*, it is a little more complex but can be depicted as in explication [H] below (Goddard, 2008b).[12] The initial components characterize it simply as an invisible part of a person, rather than as "one of two parts" because the word *heart* does not invoke the same dualistic opposition to *body*, as does *mind*. Subsequent components present the *heart* as the locus of interpersonal feelings, and, in particular, the source of positive feelings toward other people (cf. the acceptability of sentences such as *His heart was full of joy/bitterness/sadness* vs. the oddness of **His heart was full of surprise/interest*). These positive feelings can give rise to the desire to "do good things for other people," hence the link with *kindness*, *compassion*, *generosity*, etc., and their converses, *meanness*, *stinginess*, etc. The final components articulate the link with the physical body-part *heart* ("heart$_1$"). They state that people can think of the *heart* as a part of a person's body, and that the location of this part can be thought of as close to where the physical heart is (consistent with this, people sometimes touch their chests when speaking about feelings of the heart).[13]

[H] *someone's heart*

a. one part of this someone
b. people cannot see this part
c. because people have this part, when they think about other people, they can feel many things, these things can be good things, these things can be bad things
d. because people have this part, they can feel good things toward other people
e. because people have this part, they can want other people not to feel bad things
f. because people have this part, they can want to do good things for other people
g. people can think about this part like this:
 it is like a part of someone's body
 this part is near the place where this someone's heart$_1$ [M] is

Korean *Maum*

The culture specificity of these English concepts is cast into striking relief by comparison with *maum*, the major ethnopsychological construct of Korean (Yoon, 2006). *Maum* is a common basic word in everyday Korean. It is usually rendered into English as either "mind" or "heart," but it differs semantically from both of these. Dictionaries indicate that the meaning is a broad one, but are of little help in determining it with any precision. One bilingual (Korean-English) dictionary gives:

maum: 1. an entity for all kinds of mental activities including thinking, feeling, etc.; 2. a personality; 3. a mental activity for judging right and wrong; 4. an inner thought that is not expressed; 5. some feeling that can be changed according to the situation; 6. a thoughtful attitude towards other people; 7. an interest in something

Like *mind* (in this respect), *maum* is often counterposed with *mom* "body," both in everyday colloquial use and in the scholarly and general literature. There are numerous book titles such as *Thunthunhan mom kenkanghan maum* "The healthy body and the healthy *maum*," *Maumkwa momuy pyeng* "Diseases of body and *maum*," *Momkwa maumuy kwankye* "The relationship between the *maum* and the body." The symmetry of *mom* (body) and *maum* suggests that Koreans think of the *maum* as one of the two parts of a human being, the invisible psychological counterpart to the physical *mom* "body." Nevertheless, paradoxical as it may seem to speakers of English, the *maum* is conceptualized as located somewhere inside the chest area (like English *heart*, in this respect).

The importance of the *maum* concept is reflected in the existence of a large number of collocations and set phrases. The following are some metaphorical expressions that highlight the role of the *maum* in intention and attention.

maumi kata [lit. *maum* goes] 'tend to be attracted', *maumi nata* [lit. *maum* comes up] 'want to do something', *maumul ssuta* [lit. use *maum* for someone/something] 'pay attention to someone/something', *maumul mekta* [lit. eat *maum*] 'decide to do something', *maumul colita* [lit. troubles one's *maum*] 'be anxious about or be concerned about'.

It would not be right, however, to regard the *maum* as a figurative "organ of thinking," in the way that English *mind* can be regarded, because in Korean ethnopsychology this function is actually ascribed to the *meli* [lit. head]. It is perfectly possible to think without using the *maum*, but such thinking is morally suspect. To think using the *maum* implies something like altruism or virtue. The importance of the *maum* to a moral life is highlighted by the proverb *Maum-ul calkacimyen cwuketo olhun kwisini toynta,* "One can be a good ghost after death if one has lived with a good *maum*." A large number of attributive combinations with *maum* imply evaluation of a person's morality or character. For example, the expression "with a warm *maum*" means something like "warm-heartedly" and "with a beautiful *maum*" means something like "kind-heartedly" or "tender-heartedly." A person's *maum* can be good at times and bad at different times. When your *maum* is good, you tend to do good things for other people; when it is bad, you would behave oppositely (thus the expressions "a good *maum*" and "a bad *maum*" can mean something like "with good intentions" and "with bad intentions," respectively). Koreans believe that they can act with *maum* when carrying out various activities, including occupational duties, craft, or even daily routines.

On the basis of this and a range of other evidence, Yoon (2006) proposes the following explication, presented here in a somewhat modified form.[14]

[I] someone's *maum* [Korean]

a. one of two parts of this someone (one part is the body, this is the other part)
b. people cannot see this part
c. because people have this part, when they think about something they can feel many things
d. because people have this part, they can do many good things
e. when this part of someone is good, this someone wants to do good things
f. when this part of someone is bad, this someone wants to do bad things
g. when someone thinks about someone else, it is good if this someone thinks with this part
h. when someone does something, it is good if this someone thinks with this part like this: "I want to do this thing well"

It should be apparent that the ethnopsychological constructs of different languages, when appropriately analyzed, are revelatory of widely differing folk models of "psychological anatomy."

IMPLICATIONS: WHAT DO WORDS MATTER?

So far I hope to have shown three main things: (1) almost all the words in the English lexicon of mental state concepts (including emotion terms, epistemic verbs, and ethnopsychological constructs) are semantically complex and culture specific; (2) in contrast, the semantic primes THINK, FEEL, WANT, and KNOW are true semantic universals and thus provide a firm foothold in efforts to explore and map out cross-linguistic semantic variations in mental state concepts; (3) the NSM technique of reductive paraphrase into semantic primes provides the necessary analytical tool for explicating mental state concepts across languages with subtlety and precision.

For many years NSM researchers have been urging cognitive scientists to take seriously the cross-linguistic variability of mental state concepts, especially in relation to emotions, but with limited success. As mentioned earlier, one common response is along the following lines: "Words are not emotions, only representations of emotions. Our interest is in the emotions themselves, not mere words." Another frequently heard comment is that even though word meanings may be regarded as concepts in a limited sense (linguistic concepts), we can hardly assume that they are real psychological concepts in any larger (i.e., nonlinguistic) sense. In this final section, I want to argue that such responses seriously underestimate the implications of language issues for cognitive science. Over half a century ago Edward Sapir warned: "The philosopher needs to understand language if only to protect himself against his own language habits" (1949, p. 165). Echoing Sapir's concern, Shweder and Haidt (2000, p. 410) argue that when thinking about human emotions in general there is a constant danger of "assimilating them in misleading ways to an *a priori* set of lexical items available in the language of the researcher." The issues are both metatheoretical and epistemological, on the one hand, and methodological and practical, on the other. I will discuss them under four headings.

The Need to Avoid Ethnocentrism in Theory Construction

Wierzbicka has long argued that theory construction in mainstream (Anglophone) psychology has been implicitly ethnocentric, because it has tended to assume that certain English-specific folk categories, such as *emotion*, *memory*, and *mind*, are objective categories of psychological reality. Much of modern Western psychology, she argues, has emerged out of the folk psychology embedded in the English language and continues to carry that conceptual baggage, even after it has been articulated in technical terms and augmented with techniques of scientific observation and hypothesis testing. A key target of Wierzbicka's critique has been the "basic emotions" theory. Ekman (2004) and colleagues continue to maintain that there are objective biopsychological correlates for English basic emotion words, such as *happy*, *sad*, *angry*, *afraid*, and *surprised*, while denying the same privilege to the basic emotion words of other languages. This amounts to saying that the English lexicon of emotion just happens to have "got it right" (i.e., that it cuts nature at its joints), while the emotion lexicons of other languages are for some reason in a state of noncorrespondence with the emotional reality of their speakers. This means attributing a kind of epistemological superiority to English. Despite its great significance to theory building in cognitive science I will not labor this issue here, but refer the reader to sources such as the inaugural issue of the journal *Emotion Review*, where the debate is played out in Wierzbicka (2009) and commentaries thereon.

Language as a Source of Data on Subjective Experience

Ultimately all data on other people's mental states as such, i.e., subjective experience, rely in a crucial way on self-reports, and self-reports require a language. Even facial expressions, behavioral observations, and other nonverbal data cannot be correlated with particular mental states without some kind of subjective attributions, and such attributions must be made using language. Likewise, without some connection to subjective phenomenology, objective data about physiological states (skin conductance, hormonal states, neurological states, etc.) remain just that—data about physiological states. I would not dispute that the human "emotional potential" rests in part upon a bodily, physiological base and that to fully understand human emotions we have to explore and understand this physiological base. But this does not mean that emotions can be reduced to physiology or that the subjective can be reduced to the objective (cf. Searle, 2004; Barrett et al., 2007). Emotional experience can be accessed only from the inside, and it can be reported and communicated only via language. Barrett et al. (2007, pp. 374–377) spell out the logic as follows:

> Describing how emotion experiences are caused does not substitute for a description of what is felt, and in fact, an adequate description of what people feel is required so that scientists know what to explain in the first place . . . To know what emotion feels like, it is necessary to ask people what they experience . . . to examine people's verbal behaviors regarding their own mental state, in the form of self-reports . . . [self-reports] are useful—and indeed essential—for revealing the ontological structure of consciousness . . .

When interpreting the self-reports of people whose native language is not English, if we simply convert their words into their assumed English counterparts, we are in effect "recoding" those reports, and in the process altering them. Such recoding is a routine practice in psychological studies of emotion across cultures. For example, Scherer and collaborators administered a questionnaire in eight European countries with the aim of assessing the frequency and quality of emotional experience (Scherer, Wallbott, & Summerfield, 1986). Information was sought on four supposedly universal categories, each characterized by a pair of English words: *joy/happiness*, *sadness/grief*, *fear/fright*, *anger/rage*. The technique of translation and back-translation was used "to guarantee equivalence across languages" (Aebischer & Wallbott, 1986, p. 32), but as shown by the semantic studies summarized in this chapter, and by numerous others, ordinary translation and back-translation cannot guarantee full equivalence of meaning. In the German version of the questionnaire, for example, one of the words used for the category *fear/fright*, with the sanction of the back-translation procedure, was *Angst* (Scherer, 1986, p. 177), but the meaning of *Angst* is quite different from that of either *fear* or *fright* (Wierzbicka, 1999, pp. 123–165).

In my view, this flawed translation methodology should be abandoned altogether. We can—indeed, must—continue to gather self-reports about mental states in terms of indigenous categories, but what we cannot do any longer is to assume that the content of these categories can be matched in a simple fashion with the categories of the English language. The semantic content of language-specific categories must be analyzed and explicated into configurations of semantic primes, which can then be transposed without distortion across languages.[15]

Word Meanings Are "Real Concepts," Hence the Importance of Ethnopsychology

When people speak and communicate in any language, they are necessarily processing and manipulating, in real time, the language-specific semantic content of the words, grammatical constructions, intonation patterns, and so on, of their language [cf. Slobin's (1996, 2000) concept of "thinking for speaking"]. In my view, therefore, it can hardly be denied that word meanings (linguistic concepts) fully deserve to be called "concepts." The only qualification I would add is that linguistic concepts are not exclusively personal and individualistic, but also social and intersubjectively available. When speakers formulate or report their thoughts or feelings, they normally express themselves in terms of these intersubjectively available concepts. To put it bluntly, the existence of a word (and a corresponding word meaning) is evidence for the existence of a psychologically real concept. Most such concepts are of course culture specific, but this does not make them any less real.

Given that word meanings are psychologically real concepts that people use to categorize, organize, and communicate about their mental lives, they ought to be a valued object of study in their own right. As Fehr and Russell (1984, p. 483) put it: "Part of the psychologist's job . . . is to understand emotion concepts as people use them in everyday life." Although I would not go so far as to embrace the slogan that "all psychology is ethnopsychology," surely no valid scientific model of human mental life can be based solely or mainly on people from a single language and culture. It follows that the areas of inquiry covered by cross-linguistic semantics, ethnopsychology, cultural psychology, and psychological anthropology are indispensable to the cognitive science enterprise.

Words and Cultural Concepts Contribute to Emotional Experience

But still, we might ask, "Even granted that there are very great differences in how people think about their subjective experience, and in how they talk about their subjective experience, how can we know whether or not it actually *feels* different?" Two eminent psychologists who believe that emotional and cognitive experience does differ across languages and cultures are Jerome Bruner and Richard Shweder. Bruner (1990) stresses how the "experience-near" concepts of indigenous folk psychology enter into people's life narratives and self-understandings. Shweder (2003, p. 28) sees everyday words and expressions as contributing to the "implicit meanings (the goals, values, and pictures of the world) that give shape to psychological processes." In view of the epistemological priority of self-report, we must attach particular importance to the experience of "deep bilinguals," i.e., people who have lived their lives through two or more languages. Here a substantial body of life writing—both in fiction and memoir—testifies that the caliber and quality of emotional and cognitive experience indeed vary with the linguistic and cultural environment in which it is lived out (cf. Besemeres, 2002; Besemeres & Wierzbicka, 2007; Pavlenko, 2005, 2006). Even so, it must be true to say that there is a long way to go before such reports and such experiences can be brought fully within the ambit of scientific understanding.

CONCLUDING REMARKS

The lexicon of mental state concepts varies greatly across the world's languages, and though posing serious problems in some respects, this situation also offers tremendous research possibilities. If we can unpack, analyze, and understand the mental state lexicons of different languages and cultures, we gain access to diverse ways in which mental and emotional experience can be construed and interpreted. The NSM technique of cross-linguistic semantic analysis provides the necessary tools, opening up new possibilities for the scientific understanding of subjectivity and psychological experience.

ACKNOWLEDGMENTS

I would like to thank Andrea Schalley, Anna Wierzbicka, Zhengdao Ye, and Kyung-Joo Yoon for comments that have helped improve this paper. An earlier version was presented at the conference "Words and the World: How Words Capture Human Experience," Lehigh University, Bethlehem, Pennsylvania, June 6–7, 2005 and I am grateful for comments received from the participants. I am also most grateful for the comments of the editors of this volume. This research was supported by the Australian Research Council.

Notes

1. A bibliography of NSM publications, along with a number of downloadable papers, is available at the NSM Homepage [www.une.edu.au/bcss/linguistics/nsm].
2. Sources for the data in Table 4.2 are as follows (from left to right): Travis (2002), Wierzbicka (2002), Goddard (2002), Ameka (1994), Enfield (2002), and Chappell (2002).

3. Since semantic primes are unitary meanings, their lexical exponents should be typically morphologically simple, and generally speaking, this expectation is borne out. On the other hand, there is no absolute requirement for exponents of primes to be morphologically simple, and occasionally a prime is encoded by a morphologically complex item, e.g., the Lao and Ewe exponents of FEEL, *huu4.sùk2* and *se le làme*, respectively. It is not difficult to show that these meanings are unitary and cannot be composed from the (apparent) meanings of the constituent items. For example, the Lao exponent of FEEL *huu4.sùk2* contains the form *huu4*, which as an independent word is the exponent of KNOW, but the putative "extra" morpheme *sùk2* does not occur elsewhere, and as Enfield (2002, p. 176) says: "the expression is not semantically analysable into 'know' plus something else." The situation is somewhat different in Ewe, because each element of *se le làme* is meaningful if taken separately: *se* means "hear" and *le làme* is a prepositional phrase meaning literally "in (the) body." Even so, the expression as a whole cannot be interpreted as "hear in the body" (nor is it confined to bodily feelings), but means simply FEEL (Ameka, 1994).
4. For data on the polysemic extensions reported in this and the following paragraphs on THINK, FEEL, and WANT see individual chapters in Goddard and Wierzbicka (1994, 2002), Peeters (2006), and Goddard (2008a). For THINK, relevant languages include Amharic, Ewe, Lao, Longgu, Mandarin, and Swedish; for FEEL, they include Acehnese, Amharic, Ewe, French, Malay, Mandarin, and Polish; for WANT, they include Acehnese, Amharic, Ewe, Spanish, Thai, and Ulwa.
5. The third of these frames (i.e., "feel something good/bad towards someone else") has not been as well tested cross-linguistically as the other two, and is therefore somewhat provisional in status.
6. Explications [A] and [B] are for the expressions *X felt sad* and *X felt unhappy*. Explications for the corresponding expressions *X was sad* and *X was unhappy* would have an additional final component "other people could know this," reflecting their more "objective" tone.
7. Chinese has a great number of compound expressions differentiating fine shades of emotional meaning. Without a comprehensive study, the possibility that one of these might correspond to English *sadness* cannot be completely ruled out.
8. Proponents of "basic emotions" sometimes link them to universal themes or situations. For example, referring to the work of his colleagues Jerry Boucher and Klaus Scherer, Ekman (2004, p. 23) writes: "in every culture loss of something important was the trigger for sadness, but exactly what that loss was reported to be varied from one culture to another." As argued by Wierzbicka (2009), however, this proposal is equally subject to the researcher's home language bias. The English word *loss* refers, in its basic meaning, to loss of one's possessions, which has become a pervasive metaphor for misfortune in general—for "losing" someone through death, and for all kinds of other misfortunes and disappointments. Wierzbicka (2009) notes the existence of English book titles such as *Grief and Loss*, *Loss and Trauma*, *Adaptation to Loss*, and so on, and comments: "There are no such titles in French, Spanish, or German. Similarly, while one can do a PhD in *loss*, one cannot do a PhD in *la perte*, *perdida* or *Verlust* . . . This is a highly culture-specific way of looking at human life."
9. There is also another distinct use of *believe* in the conversational formula *I believe*, which has a "lighter" meaning. This usage cannot combine with *strongly* or other similar adverbs, nor can it be matched with any talk of *beliefs*. For example, I could say, using the conversational formula *I believe*, something like: *Oh, I believe the Library has a copy*; but one would hardly expect me to add any adverbs such as *strongly* to a statement of this kind, nor could I refer to *my belief* that the Library has a copy. (Other grammatical frames for *believe*, not considered here, are *to believe someone* and *to believe in someone*.)
10. Wierzbicka (2006) has (d) as "I think that it is good if someone thinks like this."
11. Gladkova (2007) has a variant of the final component of this explication as component (d).
12. For a larger study of English *heart*, see Goddard (2008b). Note that although the explication presented here accounts for a large number of collocations, typical usages and derivatives of *heart*, it does not account for every single one. For example, the word *heartless* indicates an emotional or moral "failure" that is more specific than simply not being able to "feel good things towards other people"—it indicates something more like a lack of compassion or empathy. Similarly, an expression such as *He knew it in his heart* (for example, that some

course of action was wrong) indicates something specific about how the person knew it (not by thinking, but because of feeling something). There are also expressions that contrast *heart* and *head*, in which references to the *heart* imply a lack of careful or sensible thought. In short, many frequent uses of the word *heart* occur as part of fixed expressions that require separate, albeit related, explications.

13. The word "heart" is not of course a semantic prime; rather it is a complex word functioning as a "semantic molecule." Semantic molecules are complex words that function as intermediate semantic units in the explications of yet more complex meanings; cf. Goddard (1998, pp. 254–255) and Wierzbicka (2002, 2004).
14. The status in the NSM metalanguage of the locution "think with" is unclear. Research is needed to ascertain whether comparable expressions exist across a wide range of languages.
15. On a further methodological note, there can be great value in investigating mental states using scenarios and reporting protocols framed directly in terms of semantic primes. For example, rather than asking people how often and in what circumstances they *feel angry*, we can ask about how often and under what circumstances they think something like "someone did something bad to me, I don't want this" and about how often and in what circumstances they think something like "I want to do something bad to this person because of it." Likewise, rather than asking people how often and under what circumstances they *feel happy* or *feel satisfied*, we can ask about how often and under what circumstances they think something like "some good things happened to me, I don't want anything else now" or "some very good things happened to me, I can't want anything else now," and about how often and under what circumstances they "feel something good" or "feel something very good." Framing the reporting protocols in semantic primes makes them cross-translatable and, at the same time, more fine-grained.

References

Aebischer, V., & Wallbott, H. G. (1986). Measuring emotional experiences: Questionnaire design and procedure, and the nature of the sample. In K. Scherer, H. G. Wallbott, & A. B. Summerfield (Eds.), *Experiencing emotion. A cross-cultural study* (pp. 28–38). Cambridge, UK: Cambridge University Press.

Amberber, M. (Ed.). (2007). *The language of memory in a cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Ameka, F. (1994). Ewe. In C. Goddard & A. Wierzbicka (Eds.), *Semantic and lexical universals—Theory and empirical findings* (pp. 57–86). Amsterdam: John Benjamins.

Apresjan, J. D. (2000). *Systematic lexicography* (K. Windle, Trans.). Oxford: Oxford University Press.

Arnauld, A., & Nicole, P. (1996 [1662]). *Logic or the art of thinking* (J. V. Buroker, Trans.). Cambridge, UK: Cambridge University Press.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*, 373–403.

Besemeres, M. (2002). *Translating one's self: Language and selfhood in cross-cultural autobiography*. Oxford: Peter Lang.

Besemeres, M., & Wierzbicka, A. (Eds.). (2007). *Translating lives. Living with two languages and cultures*. St Lucia: University of Queensland Press.

Briggs, J. L. (1970). *Never in anger: Portrait of an Eskimo family*. Cambridge, MA: Harvard University Press.

Brown, K. (Ed.). (2006). *Encyclopedia of language and linguistics* (2nd ed.). Oxford: Elsevier.

Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Chappell, H. (2002). The universal syntax of semantic primes in Mandarin Chinese. In C. Goddard & A. Wierzbicka (Eds.), *Meaning and universal grammar—Theory and empirical findings* (Vol. I, pp. 243–322). Amsterdam/Philadelphia: John Benjamins.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6(3/4)*, 169–200.

Ekman, P. (2004). *Emotions revealed*. New York: Henry Holt.

Ekman, P., & Davidson, R. J. (Eds.). (1994). *The nature of emotion: Fundamental questions*. New York/Oxford: Oxford University Press.

Enfield, N. J. (2002). Combinatoric properties of natural semantic metalanguage expressions in Lao. In C. Goddard & A. Wierzbicka (Eds.), *Meaning and universal grammar—Theory and empirical findings* (Vol. II, pp. 145–256). Amsterdam/Philadelphia: John Benjamins.

Enfield, N. J., & Wierzbicka, A. (Eds.) (2002). *The body in the description of emotion*. Special issue of *Pragmatics & Cognition, 10(1/2)*.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113(3),* 464–486.

Gladkova, A. (2007). Universal and language-specific aspects of "propositional attitudes": Russian vs. English. In A. C. Schalley & D. Khlentzos (Eds.), *Mental states: Volume 2: Language and cognitive structure* (pp. 61–83). Amsterdam: John Benjamins.

Goddard, C. (1996). The "social emotions" of Malay (Bahasa Melayu). *Ethos, 24(3),* 426–464.

Goddard, C. (1997). Contrastive semantics and cultural psychology: 'Surprise' in Malay and English. *Culture & Psychology, 3(2),* 153–181.

Goddard, C. (1998). *Semantic analysis: A practical introduction.* Oxford: Oxford University Press.

Goddard, C. (2001a). *Hati*: A key word in the Malay vocabulary of emotion. In J. Harkins & A. Wierzbicka (Eds.), *Emotions in crosslinguistic perspective* (pp. 167–196). Berlin: Mouton de Gruyter.

Goddard, C. (2001b). Lexico-semantic universals: A critical overview. *Linguistic Typology, 5(1),* 1–66.

Goddard, C. (2002). Semantic primes and universal grammar in Malay (Bahasa Melayu). In C. Goddard & A. Wierzbicka (Eds.), *Meaning and universal grammar—Theory and empirical findings* (Vol. I, pp. 87–172). Amsterdam/Philadelphia: John Benjamins.

Goddard, C. (2003). 'Thinking' across languages and cultures: Six dimensions of variation. *Cognitive Linguistics, 14(2/3),* 109–140.

Goddard, C. (2007). A culture-neutral metalanguage for mental state concepts. In A. C. Schalley & D. Khlentzos (Eds.), *Mental states. Volume 2: Language and cognitive structure* (pp. 11–35). Amsterdam: John Benjamins.

Goddard, C. (Ed.). (2008a). *Cross-linguistic semantics.* Amsterdam: John Benjamins.

Goddard, C. (2008b). Contrastive semantics and cultural psychology: English *heart* vs. Malay *hati.* In F. Sharifian, R. Dirven, N. Yu, & S. Neimeier (Eds.), *Culture, body, and language: Conceptualizations of internal body organs across cultures and languages* (pp. 77–103). Berlin: Mouton de Gruyter.

Goddard, C., & Karlsson, S. (2008). Re-thinking THINK in contrastive perspective: Swedish vs. English. In C. Goddard (Ed.), *Cross-linguistic semantics* (pp. 225–240). Amsterdam: John Benjamins.

Goddard, C., & Wierzbicka, A. (Eds.). (1994). *Semantic and lexical universals—Theory and empirical findings.* Amsterdam: John Benjamins.

Goddard, C., & Wierzbicka, A. (Eds.). (2002). *Meaning and universal grammar—Theory and empirical findings, Vols. I and II.* Amsterdam/Philadelphia: John Benjamins.

Harkins, J. (1995). *Desire in language and thought: A study in cross-cultural semantics.* Ph.D. thesis. Canberra: The Australian National University.

Harkins, J., & Wierzbicka, A. (Eds.). (2001). *Emotions in crosslinguistic perspective.* Berlin: Mouton de Gruyter.

Harré, R. (Ed.). (1986). *The social construction of emotions.* Oxford: Basil Blackwell.

Hasada, R. (2000). *An exploratory study of expression of emotions in Japanese: Towards a semantic interpretation.* Ph.D. thesis. Canberra: The Australian National University.

Heine, B., & Narrog, H. (Eds.). (2009). *The Oxford handbook of linguistic analysis.* Oxford: Oxford University Press.

Hjelmslev, L. (1961). *Prolegomena to a theory of language,* (2nd ed., (F. J. Whitfield, Trans.). Madison/London: University of Wisconsin Press. [First published in Danish in 1943; first English translation published 1953, as *IJAL* Memoir 7.]

Katz, J. J., & Fodor, J. A. (1963). Structure of a semantic theory. *Language, 39,* 170–210.

Kitayama, S., & Markus, H. R. (Eds.). (1994). *Emotion and culture: Empirical studies of mutual influence.* Washington, DC: American Psychological Association.

Lutz, C. (1988). *Unnatural emotions: Everyday sentiments on a micronesian atoll and their challenge to western theory.* Chicago: Chicago University Press.

Palmer, G., Goddard, C., & Lee, P. (Eds.). (2003). *Talking about "thinking."* Special Issue of *Cognitive Linguistics, 14(2/3).*

Pavlenko, A. (2005). *Emotions and multilingualism.* Cambridge, UK: Cambridge University Press.

Pavlenko, A. (Ed.). (2006). *Bilingual minds: Emotional experiences, expressions, and representation.* Clevedon: Multilingual Matters.

Peeters, B. (Ed.). (2006). *Semantic primes and universal grammar: Empirical evidence from the romance languages.* Amsterdam/Philadelphia: John Benjamins.

Pesmen, D. (2000). *Russia and soul: An exploration.* Ithaca, NY: Cornell University Press, Rosaldo, M. Z. (1980). *Knowledge and passion: Ilongot notions of self and social life*. Cambridge, UK: Cambridge University Press.

Russell, J. A. (1991). Culture and the categorization of emotion. *Psychological Bulletin, 110*, 426–450.

Russell, J. A., & Yik, M. (1996). Emotion among the Chinese. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 166–188). Hong Kong: Oxford University Press.

Sapir, E. (1949). *Selected writings of Edward Sapir in language, culture and personality*. (Edited by D. Mandelbaum). Berkeley: University of California Press.

Schalley, A. C., & Khlentzos, D. (Eds.). (2007). *Mental states. Volume 2: Language and cognitive structure.* Amsterdam: John Benjamins.

Scherer, K. R. (1986). Studying emotion empirically: Issues and a paradigm for research. In K. R. Scherer, H. G. Wallbot, & A. B. Summerfield (Eds.), *Experiencing emotion: A cross-cultural study* (pp. 3–27). Cambridge, UK: Cambridge University Press.

Scherer, K. R., Wallbott, H. G., & Summerfield, A. B. (Eds.). (1986). *Experiencing emotion. A cross-cultural study.* Cambridge, UK: Cambridge University Press.

Schwartz, T., White, G. M., & Lutz, C. A. (Eds.). (1992). *New directions in psychological anthropology*. Cambridge, UK: Cambridge University Press.

Searle, J. R. (2004). *Mind: A brief introduction.* New York: Oxford University Press.

Shore, B. (1996). *Culture in mind: Cognition, culture and the problem of meaning*. New York: Oxford University Press.

Shweder, R. A. (1991). *Thinking through cultures: Expeditions in cultural psychology*. Cambridge, MA: Harvard University Press.

Shweder, R. A. (2003). *Why do men barbecue? Recipes for cultural psychology.* Cambridge, MA: Harvard University Press.

Shweder, R. A. (2004). Deconstructing the emotions for the sake of comparative research. In A. S. R. Manstead, N. Frijda, & A. Fischer (Eds.), *Feelings and emotions: The Amsterdam Symposium* (pp. 81–97). Cambridge, UK: Cambridge University Press.

Shweder, R. A., & Haidt, J. (2000). The cultural psychology of the emotions: Ancient and renewed. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed.) (pp. 397–414). New York: The Guilford Press.

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

Slobin, D. I. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107–138). Amsterdam: John Benjamins.

Travis, C. (2002). La Metalengua Semántica Natural: The Natural Semantic Metalanguage of Spanish. In C. Goddard & A. Wierzbicka (Eds.), *Meaning and universal grammar—Theory and empirical findings* (Vol. I, pp. 173–242). Amsterdam/Philadelphia: John Benjamins.

White, G. M., & Kirkpatrick, J. (Eds.). (1985). *Person, self and experience*. Berkeley: University of California Press.

Wierzbicka, A. (1972). *Semantic primitives*. Frankfurt: Athenaum.

Wierzbicka, A. (1992). *Semantics, culture and cognition.* Oxford: Oxford University Press.

Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford: Oxford University Press.

Wierzbicka, A. (1998). "Sadness" and "anger" in Russian: The non-universality of the so-called "basic human emotions." In A. Athanasiadou & E. Tabakowska (Eds.), *Speaking of emotions: Conceptualisation and expression* (pp. 3–28). Berlin: Mouton de Gruyter.

Wierzbicka, A. (1999). *Emotions across languages and cultures*. Cambridge, UK: Cambridge University Press.

Wierzbicka, A. (2002). Semantic primes and universal grammar in Polish. In C. Goddard & A. Wierzbicka (Eds.), *Meaning and universal grammar—Theory and empirical findings* (Vol. II, pp. 65–144). Amsterdam/Philadelphia: John Benjamins.

Wierzbicka, A. (2004). 'Happiness' in cross-linguistic and cross-cultural perspective. *Daedalus, Spring 2004*, 34–43.

Wierzbicka, A. (2005). Empirical universals of language as a basis for the study of other human universals and as a tool for exploring cross-cultural differences. *Ethos, 33(2)*, 256–291.

Wierzbicka, A. (2006). *English: Meaning and culture*. New York: Oxford University Press.

Wierzbicka, A. (2009). Language and metalanguage: Key issues in emotion research. *Emotion Review, 1(1)*, 3–14.

Ye, Z. (2001). An inquiry into "sadness" in Chinese. In J. Harkins & A. Wierzbicka (Eds.), *Emotions in crosslinguistic perspective* (pp. 359–404). Berlin: Mouton de Gruyter.

Ye, Z. (2006). Why are there two 'joy-like' 'basic' emotions in Chinese? Semantic theory and empirical findings. In P. Santangelo with D. Guida (Eds.), *Love, hatred, and other passions: Questions and themes on emotions in Chinese civilization* (pp. 59–77). Leiden: Brill.

Yoon, K-J. (2006). *Constructing a Korean natural semantic metalanguage*. Seoul: Hankook Publishing Company.

5 FORCE CREATION AND POSSIBLE CAUSERS ACROSS LANGUAGES

Phillip Wolff, Ga-hyun Jeon, Bianca Klettke, and Yu Li

People have strong expectations about the kinds of entities that can serve as causers. These expectations are a matter of psychology and philosophy but also, perhaps, a matter of linguistics. Consider, for example, the sentences in (1). In English, both (1a) and (1b) are acceptable descriptions of causal events, but (1c) is not. From the point of view of many current theories of causation, this is hard to explain. The problem is that many theories of causation define causers as a relation in which the presence of the causing entity increases the likelihood of an effect (e.g., Cheng & Novick, 1992; Woodward, 2007), and all of the causers specified in (1) meet this criterion.

(1) a. The boy broke the window.
  b. The knife cut the bread.
  c. #The fork lifted the potato.

The problem becomes more complex when we look across languages. As we will see, in many languages, only (1a) is acceptable and (1b) is as bad as (1c) (Alexiadou & Schäfer, 2006; Folli & Harley, 2007; Guilfoyle, 2000; Levin & Rappaport-Hovav, 2005; van Voorst, 1996). In other words, in some languages, simple objects such as "knives" and "forks" cannot be described as the causer of an event. If you ask a German speaker, for example, whether a knife can cut bread, they might point out that knives do not have "arms and hands that would enable them to do this," as one of our language consultants informed us.

In this chapter, we attempt to identify the semantic constraints on causers, first within English, and then across languages. We seek to show how causal expressions in English and other languages reveal an interaction between grammatical structure and word meaning and how this interaction may provide insights into the representation of causation in the mind. We view this kind of research as part of a broader attempt to understand the language–thought interface.

The chapter is structured as follows. Focusing on English, we identify the range of possible entities that can appear as the external argument of a causal sentence, which, in English, is coextensive with the subject of the sentence. We propose that the range of possible causers in English and in other languages can be understood in terms of a continuum of force creation, with languages differing on this continuum with respect to the kinds of causers that can appear as external arguments. We further show that where a language divides the continuum can be predicted by whether the language uses morphological case. Finally, we describe empirical evidence in support of our proposals in three lines of research. In Experiment 1, we demonstrate systematic differences across languages with respect to the range of acceptable external arguments, specifically, English and Chinese allow a greater range of causers than Korean. In Experiment 2, we demonstrate that constraints on causer external arguments in English, Mandarin Chinese, and Korean affect judgments of sentence acceptability in one-clause and two-clause descriptions of causal events. In Experiment 3, we show that such constraints also affect the selection of causal verbs by speakers of English, Russian, and German. We conclude with a discussion of the implications of these findings for the language–thought interface.

EXTERNAL ARGUMENTS AND CAUSERS IN ENGLISH

We are interested in the characteristics of entities that make good external arguments, or causers, in sentences that describe causal relationships. External arguments are typically associated with the semantic roles of "agent," "instrument," or "theme" (Radford, 1988). The concept of an external argument is defined in various approaches to grammars in configurational terms. For example, in the sentence *Alison petted the guinea pig* the second argument of the verb, *guinea pig*, is positioned inside the verb phrase, whereas the first argument, *Alison*, is positioned outside the verb phrase, and hence is referred to as an external argument (Radford, 1997). In generative grammar, the external arguments are effectively coextensive with the grammatical relation subject (Van Valin & LaPolla, 1997). In this chapter, however, we focus on external arguments rather than subjects because the behavioral and coding properties of "subjects" vary considerably across languages (Van Valin & LaPolla, 1997). Indeed, as argued by LaPolla (1993), Mandarin Chinese may completely lack the grammatical category of subject. We focus on the notion of external argument because it offers a more universal, language-neutral category of analysis, and we are especially interested in the range of possible causers across languages. Our claim will be that languages differ in the range of types of semantic roles that can appear in the external argument position. It is important to emphasize that the criteria that license external arguments in causal sentences may differ from the criteria that license external arguments in sentences describing noncausal relationships (Grimm, 2007). For example, whereas the sentence *The fork moved the potato* sounds odd, the sentence *The fork stabbed the potato* sounds fine; the first sentence describes a causal relationship and the latter does not.

English allows for a wide variety of entities in the external argument position of causal expressions. Although causer external arguments are often animate entities, animacy is not required (Fillmore, 1968), as shown in (2).

(2) a. Lightning killed the guard.
b. The wind opened the door.
c. The sun melted the ice.
d. The wave capsized the boat.

Though they lack intentionality, natural forces, like intentional agents, can initiate their own actions and generate their own energy (Cruse, 1973; Grimm, 2007; Alexiadou & Schäfer, 2006; Schlesinger, 1989). Consistent with this characterization, natural forces, unlike instruments (e.g., *keys, knives, drills*), sound odd in *with* phrases, for example, *The janitor opened the door with the wind*. *With* phrases imply human control, and natural forces, by definition, cannot be controlled by human agents (Nilsen, 1973).

Projectiles are another type of inanimate entity that can readily appear as the external argument of a causal sentence (Cruse, 1973; Grimm, 2007; Kearns, 2000), as exemplified in (3).

(3) a. The stone broke the window. (Kearns, 2000)
b. The bullet killed the president. (Schlesinger, 1989)
c. The cannon ball sank the ship.
d. The falling branch broke the car window.

Projectiles acquire their energy from an external rather than internal force, but that energy seems to be construed of as their own (Cruse, 1973; Grimm, 2007; Kearns, 2000; Alexiadou & Schaffer, 2006). It is interesting to note that when these kinds of entities appear as external arguments, the presence of an external, initiating agent is deemphasized (see Schlesinger, 1989); indeed, in the case of entities like branches (e.g., 3d), an external (sentient) agent may not be clearly obvious. But even when an external agent is present, projectiles sound somewhat odd expressed as instruments in a *with* phrase, as in the sentence *The assassin killed the president with a bullet*, which suggests that they are more similar to animate entities and natural forces than to instruments (see Alexiadou & Schaffer, 2006).

Other kinds of entities that can serve as a causer external argument are entities that are often referred to as "instruments," that is, entities that are used by a sentient entity to accomplish a task (Delancey, 1984; Schlesinger, 1989). In the lexical semantics literature, instruments have been divided into two types, *intermediary* and *facilitating*. Intermediary instruments are entities that can act in a manner that is at least partially independent of an agent's controlling influence (Alexiadou & Schaffer, 2006; Kamp & Rossdeutscher, 1994). As such, they can be viewed as acting as an "intermediary" in the causal chain (Levin & Rappaport, 1988; Marantz, 1984). As shown in (4), some intermediary instruments, although controlled by human agents, have their own power source, which makes them, as described by Grimm (2007), "semiautonomous."

(4) a. The forklift killed the construction worker.
b. The remote control opened the door.
c. The crane sank the boat.
d. The bomb shook the ground.

Other intermediary instruments do not generate their own energy, but can nevertheless appear in the external argument position, as shown in (5).

(5) a. The key opened the door.
b. The knife cut the bread. (Levin, 1994)
c. The axe split the log.
d. The diamond scratched the glass.
e. The wrench tightened three out of four bolts without incident.

Based on the discussion thus far, it might appear that the range of possible entities that can appear as an external argument is unconstrained. Importantly, however, there are certain entities that cannot appear as external arguments, despite their having an important role in the causal chain (as indicated by their acceptability in *with* phrases).

(6) a. #The snow shovel moved the snow. (adapted from Grimm, 2007)
b. #The fork lifted the potato.
c. #The spatula flipped the pancake.
d. #The chisel carved the statue. (Schlesinger, 1979)
e. #The razor shaved Max. (Reinhart, 2002)
f. #The broom cleaned the room.

The external arguments in (6) exemplify facilitating instruments (also referred to as "enabling" or "accessory" instruments). According to Kamp and Rossdeutscher (1994), a facilitating instrument is an entity that is not easily separated from the agent that handles it: there is no "causal complex" that includes the instrument that does not also include an external agent.

Given the high degree of similarity between the instruments in (5) and (6), an important question that arises is how to distinguish between intermediary and facilitating instruments. If an entity is able to generate its own energy or is able to transport externally acquired energy, it is likely that it can appear as an external argument in English. However, given that entities such as knives and keys can serve as external arguments, as shown in (5), the ability to generate or carry energy does not appear to be necessary. In the next section, we propose that the key difference between intermediary and facilitating instruments is the ability to create force. As discussed in the following, the notion of force creation offers not only distinctions that can account for the difference between different kinds of instruments but also a continuum that can account for the difference in the semantics of external arguments across languages.

FORCE CREATION

We propose that an entity can appear in the external argument position in English and other languages if it acts as a *force creator*. We further suggest that the relatively wide range of entities that can serve as causers in English follows from the relatively wide range of ways in which forces can be created. In this section, we describe three broad categories of force creation: energy conversion, physical contact, and force redirection.

First, a force can be created through energy conversion, that is, when energy is transformed from one form to another (Young & Freedman, 1999). For example, the forces involved in running, sneezing, and walking begin with a transformation of potential energy, in the form of chemical potential, to motion, kinetic energy. In internal combustion engines, energy conversion occurs when chemical potential energy in gasoline is transformed into kinetic energy. Many of the entities that make good causer external arguments (intentional agents, natural forces, power devices) generate their own kinetic force by converting energy. The creation of energy from potential energy seems to capture the notion of internal causation, which has been cited as playing an important role in distinguishing different classes of verbs (Levin & Rappaport-Hovav, 1995). According to Levin and Rappaport-Hovav (2005), internal causation arises from properties that are inherent to an entity. They note that although internal causation is prototypically associated with agentive entities (e.g., humans), it is also characteristic of nonagentive entities such as natural forces and complex machines, and that it is this property that allows an entity to be viewed as a causer, or "responsible" for the event.

Another way in which a force can be created is through physical contact. When an object approaches and hits another object, it imparts a force on the hit object. Importantly, the imparted force does not exist until the moment of the collision. We know that the force does not exist prior to the impact because the properties of a force depend on the object that is hit. For example, a bullet that hits and tears through a piece of tissue will impart less force on the contacted entity than a bullet that hits and is stopped by a lead block. In this way, force is a quantity that is created by the interaction of two entities.

Finally, a force can be created through force redirection, which occurs when the direction or magnitude of an applied force is changed. Mechanical devices that produce a change in direction or magnitude in a single motion are known as simple machines (Asimov, 1988; Cotterell & Kamminga, 1990). There are six classic simple machines: levers, pulleys, inclined planes, wedges, screws, and wheel and axles. Many everyday objects operate as simple machines (or combinations of these machines), including hammers, axes, bottle openers, shovels, crowbars, see-saws, wrenches, pliers, and scissors, among others. The notion of force modification offers an account of why instruments such as hammers and knives can appear as external arguments. It is important to keep in mind that a force is a vector, meaning it has only one direction and magnitude. As a consequence, a change in the direction of a force implies the creation of a new force.

To see how a simple machine creates new forces, consider the case of a knife. A knife is a wedge, that is, a triangular object that operates by converting a force applied in the direction of one edge into forces that are perpendicular to the applied force, as shown in Figure 5.1. Thus, when someone cuts a loaf of bread, the knife, in effect, creates two new forces perpendicular to the direction of the force from the agent. When

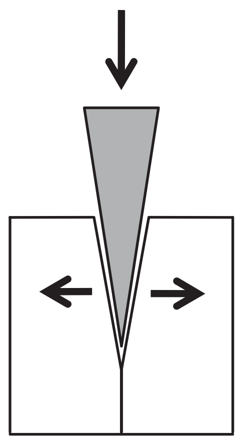


FIGURE 5.1. The downward force acting on the wedge is redirected to the sides of the material.

the operations of two or more simple machines are combined, the result is a complex or compound machine. One such machine is the shovel, which combines the operations of a wedge and a lever. The wedge part helps push the scoop of the shovel into the ground by pushing the soil open (like a knife). The lever part changes the direction of the applied force at the point at which the shovel pivots in the dirt.

According to our proposal, in English, entities can serve as causers when they create forces though energy conversion, physical contact, or force redirection. Thus, instruments such as knives, keys, and shovels, which can redirect forces, creating new forces, can appear as external arguments (see 5). Importantly, while instruments like forks, shovels, chisels, and sticks can be used as simple machines, they can also be used in manner in which they do not create new forces. On our account, when instruments are used in a manner that does not create a new force, they will not make acceptable external arguments. For example, *the shovel* can appear as an external argument when the verb implies that it is used to redirect forces (see 7a). In contrast, when it is clear that *the shovel* does not redirect forces, or the effect is not due to the force created by the shovel coming into contact with the effected entity, it should sound odd as an external argument (7b).

(7) a. . . . the shovel overturned clumps of soil. . . (Etzioni-Halevy, 2006)
b. #The snow shovel moved the snow. (Adapted from Grimm, 2007)

In general, we suggest that the reason why instruments sometimes do not make good external arguments is because they do not act as force creators. When instruments do not create forces, all that can be said of them is that they are "controlled" by the initiating agent and they fall into the category of "facilitating instruments." However, when the instrument is not only controlled, but also creates force, whether through physical contact or force redirection, the instrument falls into the category of "intermediary instruments."

CROSS-LINGUISTIC DIFFERENCES IN THE APPEARANCE OF NONAGENTIVE ENTITIES IN THE EXTERNAL ARGUMENT POSITION

It has long been observed that languages differ in the kinds of entities that can serve as external arguments (Comrie, 1989; Craig, 1977; DeLancey, 1984; Folli & Harley, 2007; Guilfoyle, 2000; Hawkins, 1985; Wolff & Ventura, 2009). In particular, it has been suggested that English may allow a much broader range of external arguments than Irish, Dutch, German, Russian, Jacaltec, Cora, Spanish, or Korean (Comrie, 1989; Craig, 1977; Fausey & Boroditsky, 2008; Guilfoyle, 2000; Hawkins, 1985; Soto, 2001; van Voorst, 1996; Wolff & Ventura, 2009), and probably also Japanese, French, Italian, and Hare (Achard, 2001; DeLancey, 1984; Folli & Harley, 2007).

In a discussion of the difference between English and German, Hawkins (1985) suggests how such differences might arise.[1] Hawkins notes that one prominent difference between German and English is the way in which grammatical relations are marked. In English, the grammatical function of a noun within a phrase or clause is indicated by relatively fixed word order and prepositions. German, in contrast, marks grammatical relations with morphological case, that is, morphological attachments or modifications to the noun. Common cases include nominative case, which indicates the subject of a finite verb; accusative case, which indicates the direct object of a verb; dative case, which indicates the indirect object of a verb; and instrumental case, which indicates the object used to perform an action. Like other languages with case systems, German has relatively free word order (Hawkins, 1985).

Hawkins (1985) hypothesized that word order rules interacting with certain pragmatic principles constrain the ranges of possible external arguments in German and English. For example, on the discourse level, it is preferable to position "given" information before "new" information (Lambrecht, 1994). This principle is easy to realize in German: Whether the given information is found in the direct object or the subject, either can occur initially. However, this option is not as simply realized in English because its word order is relatively fixed. Because English is less flexible with respect to word order, Hawkins (1985) suggests that English might satisfy pragmatic constraints, such as the given-new principle, by allowing for a wider range of entities in the subject position, including instrumental subjects. By doing so, English can position the given information first regardless of the entity's ability to self-energize.

We suggest that Hawkins' proposals be extended to other languages. In languages in which word order is relatively fixed, due to its role in indicating grammatical relations, the range of entities that can appear in the external argument position is likely to be greater than in languages with more flexible word order, due to their use of a morphological case system to indicate grammatical relations. Our proposal is consistent with Guilfoyle's (2000) division of languages into two types: Type A languages (e.g., Dutch, Irish), which restrict the subject position to entities that can initiate events, that is, mostly intentional agents or natural forces, and Type B languages (e.g., English), which allow for a wider range of entities as long as they participate in the causal chain.[2] We will refer to Guilfoyle's Type A languages as initiator languages and Type B languages as participant languages. Interestingly, the initiator languages cited by Guilfoyle have morphological case whereas the one language that she cites as an example of a Type B language, English, does not. We suggest, then, that Hawkins' and Guilfoyle's proposals can be aligned with each other. According to what we call the *initiator hypothesis*, the kinds of entities that make good external arguments depend on the flexibility of a language's word order, which in turn depends on the presence of morphological case. Languages with morphological case (e.g., Dutch, Irish) will tend to be less restrictive in word order and hence more selective about the kinds of entities that can appear as external arguments; in particular, in initiator languages, external arguments must be able to initiate the causal chain. Languages without morphological case (e.g., English) will tend to be more restrictive about word order and less selective about the kinds of entities that can appear as external arguments; in participant languages, all that may be required is that the entity be an intermediary in the causal chain leading up to the result.

We further propose that the semantics of external arguments in the kinds of languages identified by Guilfoyle and Hawkins can be characterized in terms of force creation. In initiator languages, external arguments may be restricted to entities that can create force only through energy conversion, which is most clearly realized when an entity initiates an event. In contrast, in participant languages, the external argument can be realized by entities that create forces by other means, including physical contact and force redirection.

An analysis of this type offers several theoretical advantages over past accounts of the semantic of external arguments. First, as discussed previously, it offers a more unified account of the difference between intermediary and facilitating instruments in English than previous accounts. Second, the distinctions offered in this account can be extended to explain differences in external argument selection across languages. A third benefit of such a reexpression is that it integrates well with current theories of argument realization; in particular, it complements the notion of a *force recipient* as proposed by Levin (2007), which applies to the direct object of a sentence. As described by Levin (2007), a force recipient is the target of a transmitted force and can be used to distinguish several classes of transitive verbs (e.g., verbs of change of state, surface contact, and perception). Finally, such an account is highly compatible with a relatively well-developed approach to the representation of causation, force dynamics, which has received significant support in the linguistic and psychological literatures (see Copley, 2005; Jackendoff, 2002; Pinker, 2007; Talmy, 1988; Wolff, 2007; Wolff & Song, 2003).

The initiator hypothesis is supported by the acceptability judgments reported in the literature and by our consultants. As predicted, the sentences in (8) and (9), which describe simple instruments, are acceptable in English, a language without morphological case, but unacceptable in languages such as Dutch, German, and (perhaps) Russian, languages that do mark for morphological case.

(8) The rock broke the windshield.
#Het steentje heft de voorruit gebroken. (Dutch; Alexiadou & Schaffer, 2006; van Voorst, 1996)
#Der Stein zerbrach die Windschutzscheibe. (German)
?Kamen′ razbil lobovoe steklo. (Russian)

(9) The key opened the door.
#D′oscail an eochair an doras. (Irish; Alexiadou & Schaffer, 2006; Guilfoyle, 2000)
#Desleutel opende de deur. (Dutch; Alexiadou & Schaffer, 2006; Guilfoyle, 2000)
#Kliuch otkryl dver′. (Russian; Wolff & Ventura, 2009)
?Der Schlüssel öffnete die Tür. (German)

A much stronger test of this hypothesis would be to examine external arguments in languages that have not yet been studied in this respect. In the following section, we examine the predictions of the initiator hypothesis for two such languages, Korean, which has a case system (Song, 1988), and Mandarin Chinese, which does not (LaPolla, 1993). Wolff, Jeon, Klettke, and Li (2009) describe the semantics of causatives in these languages.

EXPERIMENTAL EVIDENCE

External Arguments in English, Chinese, and to Korean

According to our proposal, the range of entities that can appear in the external argument position should be wider for languages without morphological case than for languages with morphological case. We would expect, then, that English and Chinese should have comparable restrictions on what can appear in the external argument position and that both languages should differ from Korean. Furthermore, external arguments in Korean, but not English and Chinese, should be able to generate their own energy. We tested these predictions by having speakers of English, Mandarin Chinese, and Korean rate the acceptability of 30 sentences containing nonagentive external arguments that were either high or low in their ability to generate their own energy (for details, see Wolff et al., 2009). A sample of these sentences is shown in (10).

(10) a. The sunlight dried the towel.
b. The wave flipped the boat.
c. The microwave defrosted the meat.
d. The air conditioner cooled the room.
e. The knife cut the bread.
f. The chopsticks squashed the noodle.
g. The bullet killed the president.
h. The spoon moved the ice cream.

The sentences were developed in consultation with native speakers of Mandarin Chinese and Korean to ensure that the words used had relatively direct translations in each language. All of the sentences described a causal interaction, broadly construed, between the external argument and the object of the sentence; in every sentence, the object underwent a change of state or location as a result of the actions of the external argument. The sentences were divided into two groups. Half of the sentences named external arguments that were able to generate their own energy and the remaining sentences contained external arguments that were unable to generate their own energy. Assignment to these two groups was based on the results from a separate rating task in which English speakers judged "the degree to which the affectors in the sentences were able to generate their own energy/force on a scale of 0 to 100." In the high-energy generation group, roughly half of the external arguments were natural forces and the remaining entities were energy-transforming devices (e.g., microwaves). Of the external arguments naming entities without power sources, the majority were what would be construed of as instruments in the linguistics literature. The English versions of these sentences and associated instructions were translated into Mandarin Chinese and Korean; the translations were then checked by additional native speakers in each of these languages. In the rating task, participants were instructed to "rate how acceptable the sentences are on a scale of 0 to 100 (0 = is completely not acceptable; 100 = is perfectly acceptable)."

The participants were native speakers of English (Atlanta), Mandarin Chinese (Taitung), and Korean (Seoul), with 18 from each language. The Chinese and Korean speakers were tested in their own countries. The English speakers were undergraduates at Emory University.

There were three main results. First, as shown in Figure 5.2, there was an overall effect of language: acceptability ratings were highest for the English speakers, followed by the Chinese speakers, and finally followed by the Korean speakers. Statistical analyses indicated that English and Chinese differed significantly from Korean, but not from each other (see Wolff et al., 2009). Second, there was an overall effect of energy generation. Sentences with external arguments that could generate their own energy had higher acceptability ratings than sentences with external arguments that could not generate energy. The difference between high- and low-energy generation was significant for the English and Korean speakers, and there was a tendency toward significance with the Chinese speakers. The most important finding was a significant interaction between language and energy generation: the difference between high- and low-energy generation for the Korean speakers was greater than the difference between high- and low-energy generation for the Chinese and English speakers. This interaction is consistent with the prediction that external arguments should be acceptable in languages such as Korean only if they are capable of generating their own energy. The ability to self-energize is also important in languages such as Chinese and English, but to a lesser extent. The results from this study support the proposal that external argument selection should be more restricted in languages that have morphological case than in languages that do not. Furthermore, the results suggest that a large proportion of the difference between languages can be attributed to whether the external argument is able to self-energize.

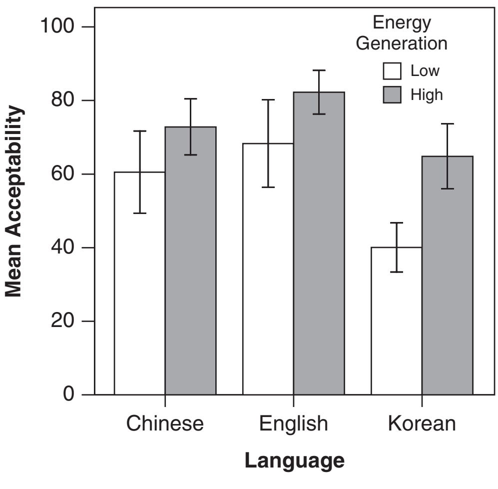


FIGURE 5.2. Acceptability ratings to sentences containing external arguments that were either high or low in their ability to generate their own energy; the error bars indicate 95% confidence intervals.

To the extent that the results reflect constraints on external arguments in general, the effects should apply not only to the main external argument of a sentence, but also to the external argument of embedded clauses within a sentence. According to the initiator hypothesis, speakers of English and Chinese should be willing to use causal chains with periphrastic causative (biclausal) expressions regardless of whether the external argument of the embedded clause is able to self-energize; in contrast, Korean speakers should be willing to use biclausal expressions only if the external argument of the embedded clause is able to self-energize. These predictions were tested in two experiments. In the first, we examined the kinds of scenarios that we expected would be described in essentially the same way by speakers of English, Chinese, and Korean. The point of this experiment was to show that the kinds of constructions being compared across languages have essentially the same meaning for certain kinds of occurrences. In the second experiment, we examined the kinds of scenarios that, according to the initiator hypothesis, should lead to differences in linguistic coding across the languages.

Selection Restrictions on Causees

Causation can be expressed in various ways in English and in other languages (for a review, see Wolff et al., 2005). One such way is by means of a lexical causative. A lexical causative (e.g., *open*, *break*, *melt*) expresses a causal relation in a single clause that includes a causer, a causee, and a change of state. In the lexical causative in (11a), Alison (the causer) causes the door (the causee) to become open (change of state).

(11) a. Alison opened the door. (lexical causative)
b. Alison caused Nathaniel to open the door. (periphrastic causative)
c. Alison caused the key to open the door. (periphrastic causative)

Another way to describe causal relationships is with a periphrastic causative. Periphrastic causatives express causal relationships (broadly construed) with two or more verbs, one associated with the cause and one associated with the result (Baron, 1977; Radford, 1988; Shibatani, 1976; Wolff, 2003; Wolff et al., 2005). For example, in the sentences in (4b) and (4c) the matrix verb, *caused*, expresses the notion of CAUSE and the embedded verb, *open*, expresses a particular endstate or result. Periphrastic causatives are typically analyzed as composed of two clauses, a main clause and an embedded clause (Radford, 1988). Such an analysis is not straightforward since it may initially appear as if the embedded clauses in (11b) and (11c) lack external arguments. Indeed, in most syntactic analyses of the sentences in (11b) and (11c), the noun phrases (NPs) immediately following the matrix verb (*Nathaniel*, *the key*) function as the objects of the matrix verbs, and there exists a covert "empty" external argument in the embedded clause (Jackendoff & Culicover, 2003; Polinsky & Potsdam, 2003; Radford, 1988). In periphrastic causatives, the referential properties of this covert external argument, represented atheoretically by the symbol Δ, can be represented using shared subscripts, as in (12).

(12) Alison caused Nathaniel$_i$ [Δ_i to open the door]

Based on such an analysis, we can say, somewhat inaccurately, that the NP that follows the matrix verb in periphrastic causative structures serves two roles: It functions directly as the object of the matrix verb, and indirectly, through indexing, as the external argument of the embedded verb (Radford, 1988, 1997).

In addition to differing in syntax, lexical and periphrastic causatives differ in their semantics. Whereas periphrastic causatives can express either direct or indirect causation, lexical causatives imply direct causation (e.g. Levin & Rappaport-Hovav, 1995; Pinker, 1989; Shibatani, 1976; Song & Wolff, 2005; Wierzbicka, 1988; Wolff, 2003; among others). For example, the lexical causative in (11a) implies a situation in which Alison made direct physical contact with the door, for example, by turning the doorknob and pushing the door open. The periphrastic causative in (11b) implies a situation in which Alison did something, such as starting to smoke, that indirectly prompted the causee, Nathaniel, to open the door to get fresh air.

The cross-linguistic differences in the selection restrictions on external arguments discussed earlier should have consequences for the way periphrastic causatives are used and understood across languages. In initiator languages, such as Korean, the causee of a periphrastic causative should be restricted to entities that can self-energize. This is predicted because the causee of a periphrastic causative is (via indexing) the external argument of the embedded clause, and just like the external argument in the matrix clause, should be limited to entities that generate their own energy. In effect, then, in initiator languages, both the causer and the causee will tend to be agentive entities, as exemplified in (11b). In participant languages, the causee may also be either agentive or nonagentive, as exemplified in (11b) and (11c).

We tested these predictions by examining how speakers of English, Chinese, and Korean would rate the acceptability of lexical and periphrastic causative descriptions of causal events (Wolff, Jeon, & Yeh, 2006; Wolff et al., 2009). The study had two parts. In the first, participants were shown animations of causal chains in which a human interacted either directly with an inanimate object or indirectly with another human. Figure 5.3A shows a single frame from one of the animations that depicted the first type of scenario. Here a woman closes a door by pushing on it. There is only one agent (the woman), who is able to initiate her own energy, and the causation is direct. We predicted that the speakers of all three languages would give relatively high ratings to single clause, lexical causative descriptions of this event, as in (13).

(13) a. The mother closed the door. (English)
b. māma guān shang le mén. (Chinese)
Mom closed up door.
c. Umma-ga mun-eul dat-atda. (Korean)
Mom-NOM door-ACC close-PST.

Figure 5.3B shows a single frame from an animation that depicted the second type of situation. In these scenarios a human tells another human to do something, that is, these causal chains consisted of two agents, each capable of initiating its own energy. Because the causation in these chains is indirect, we predicted that all three languages would give high ratings to biclausal, periphrastic descriptions of these events, as in (14).

(14) a. The mother caused the son to close the door.
b. māma shǐdé érzi guān shang le mén.
Mom cause son closed door.
c. Umma-ga aadeul-ege mun-eul dat-key haetda.
Mom-NOM son-DAT door-ACC close-ADV CAUSE-PST.

The participants were 48 native speakers of English (Atlanta), Mandarin Chinese (Taitung), and Korean (Seoul), with 16 from each language. The participants were tested in their home countries. Ten pairs of animations were constructed like the one in Figure 5.3, with one member of each pair depicting direct causation and the other depicting indirect causation. For each animation, participants rated the acceptability of single- and biclausal descriptions on a scale of 0 to 100 (0 = not acceptable; 100 = completely acceptable).

As predicted, all language groups gave high ratings to single-clause, lexical causative descriptions when the animations depicted direct causation and high ratings to biclausal, periphrastic causative descriptions when the animations depicted indirect causation (see Fig. 5.4). The results support the hypothesis that self-initiating agents can serve as external arguments of both main and

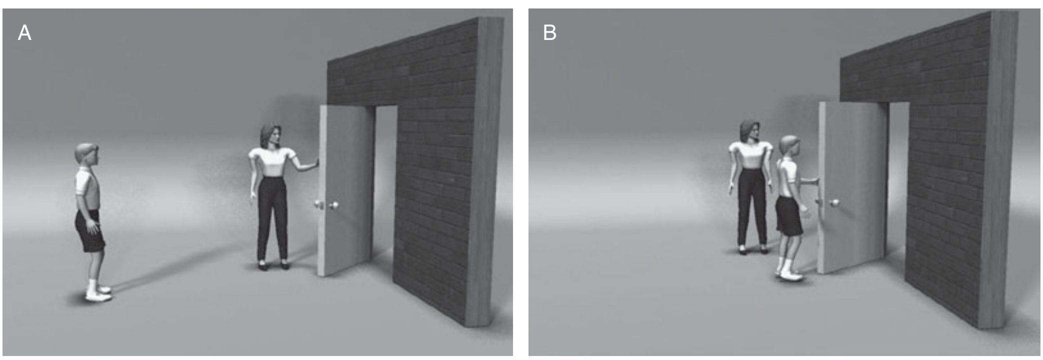


FIGURE 5.3. Frames from animations depicting (A) direct and (B) indirect causal chains used in part 1 of Wolff, Jeon, and Yeh (2006).

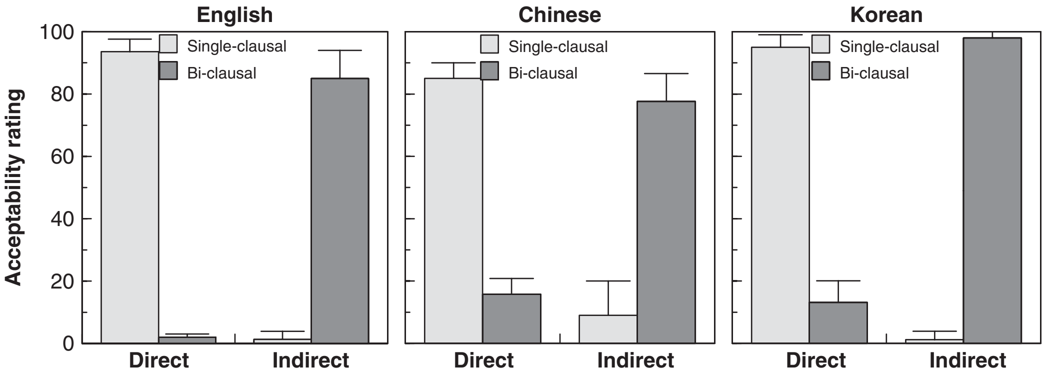


FIGURE 5.4. Acceptability ratings with 95% confidence intervals.

embedded clauses in English, Chinese, and Korean.

In the second part of the study, we examined scenarios for which we predicted the acceptability ratings would differ. In these scenarios, the causal chains included intermediate entities that were inanimate. The animations were again constructed in pairs, this time differing with respect to whether the intermediate entity was fully under the control of the initial human agent. For example, in Figure 5.5A, a girl throws a ball at a vase and breaks it. Because the ball's motion is controlled by the girl, we predicted that the participants, regardless of their language, would view the causation as direct, and hence would give high ratings to single-clause descriptions of the event, as in (15).

(15) a. The girl broke the vase.
b. nǚhái dǎpò le huāping. (Mandarin Chinese)
Girl broke vase.
c. Sonyeo-ga ggotbyoung-eul ggaetda. (Korean)
Girl-NOM vase-ACC broke-PAST.

In Figure 5.5B, in contrast, the girl accidently bounces a ball off her foot and the ball hits the vase, breaking it. Because the ball's actions are not controlled, the causal chain should be viewed as indirect and so bias speakers to prefer biclausal causal descriptions of the event. For the English and Chinese speakers, this is straightforward: The ball acts as a simple machine by redirecting the force acting on it, and hence it should qualify as an external argument in the embedded clause.

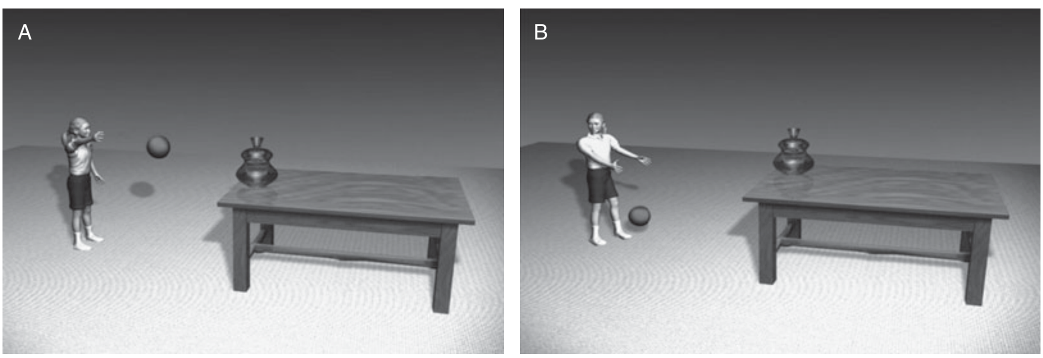


FIGURE 5.5. Frames from two animations depicting direct (A) and indirect (B) causal chains used in part 2 of Wolff, Jeon, and Yeh (2006).

However, for the Korean speakers, a biclausal expression should not be acceptable since, in Korean an external argument must do more than redirect force, it must be able to generate force. As a consequence, Koreans should find it unacceptable to use the ball as an external argument of the embedded clause in a periphrastic causative, and hence should give the biclausal descriptions low ratings, even though the causation is indirect. In sum, we predicted that English and Chinese but not Korean speakers would give relatively high ratings to periphrastic causative descriptions (e.g., 16b) of the animations associated with the kinds of scenes depicted in Figure 5.5B.

(16) a. The girl caused the ball to break the vase.
b. nǚhái shǐdé qiú dǎpò le huāpíng.
Girl caused ball broke vase.
c. *Sonyeo-ga gong-ege ggotbyoung-eul ggae-ge haetda.
Girl-NOM ball-DAT vase-ACC break-CAUSE do-PST-DEC.

The participants in this part of the study were the same as those in the first part. As in the first part, 10 pairs of animations were constructed like the ones shown in Figure 5.5. All of the animations depicted a causal chain that began with a person and was mediated by a nonagentive entity (e.g., ball, stick, wind, remote control). As in the first part, participants rated the acceptability of single- and biclausal descriptions for each animation on a scale of 0 to 100 (0 = not acceptable; 100 = completely acceptable).

As shown in Figure 5.6, all language groups gave high ratings to single-clause descriptions of animations depicting direct causation. However, only the English and Chinese speakers gave high ratings for the biclausal sentences in the case of the indirect causal chains. For Korean speakers, the biclausal expressions, in which the nonagentive entity served as the external argument of the embedded clause, were completely unacceptable. The results from this second study provide further evidence for the proposal that the kinds of entities that can serve as external arguments of causal sentences differ across languages, and that the constraint is specifically associated with external arguments (not simply initial causer in the causal chain).

Cross-Linguistic Differences in the Application of Causal Verbs

In addition to restricting the use of periphrastic causatives to interactions between people (as in Korean), language-specific constraints on possible external arguments might lead the speakers of different languages to describe the same event using different verbs. Such differences might occur because constraints on external arguments might affect which aspects of an event speakers use to establish the presence of certain grammatically relevant properties in a scene, which, in turn, would affect verb preferences. As discussed, for speakers of

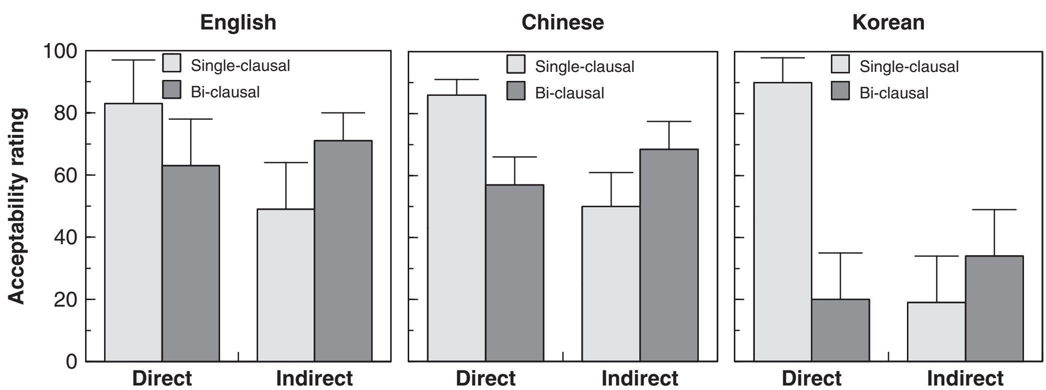


FIGURE 5.6. Acceptability ratings with 95% confidence intervals.

initiator languages, the choice of a causal description will depend on internally generated forces, including intentions, whereas for speakers of participant languages, that choice will depend on both internal and external forces, that is, on both intentions and externally generated physical forces (e.g., friction). For certain events, these two aspects of an event might give rise to conflicting conclusions about the presence of a particular grammatically relevant property. When use of a particular verb depends on such a property, speakers of initiator languages might reach a different conclusion about how to describe an event than speakers of participant languages. To explain exactly how these conclusions might differ, we need to be more specific about the semantics of causal verbs.

According to the force dynamic approach to causal representation (Wolff, 2007; Wolff & Song, 2003; based on Talmy, 1988 and extended in Pinker, 1989), descriptions of causation may be based on one of three major types of periphrastic verbs: CAUSE verbs (*cause, make, force, get*), PREVENT verbs (*prevent, block, keep, protect*), and ENABLE verbs (*enable, allow, permit, help*) (Wolff & Song, 2003; Wolff, Klettke, Ventura, & Song, 2005). In force dynamics, the concept of CAUSE differs from the concept of ENABLE primarily in terms of the causee's inherent tendency for a particular endstate, that is, the causee's physical or intentional inclination for a particular state of affairs. In CAUSE scenarios, the causee does not have a tendency toward a particular state, but is pushed to that state, in opposition to its tendency, by the force associated with the causer. In ENABLE scenarios, the causee does have a tendency for a particular state and reaches that endstate via the force of the causer, which is concordant with the tendency of the causee. The difference between these two concepts, and their associated verbs, can be exemplified by the continuum of scenarios shown in Figure 5.7.

Each panel in Figure 5.7 shows a frame from three different animations. In each of these animations, a man, the causer, holds a rope and pulls another man, the causee, on a dolly, across a line. In the panel on the far left, the causee resists by pushing the dolly backward. According to force dynamics, such a situation should be construed as one of causation because the causee does not have a tendency for the result (crossing the line), but is opposed by the causer and ultimately crosses the line. In the panel on the far right, the causee pushes

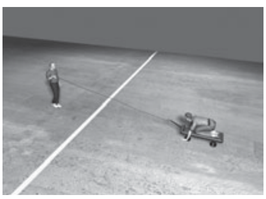
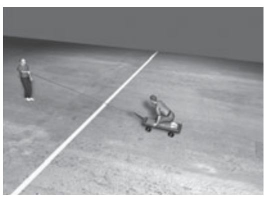
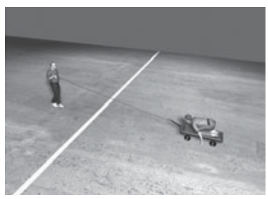
	Causee (man on dolly) opposes the causer	Causee (man on dolly) does nothing	Causee pushes along with the causer
			
English	CAUSE	CAUSE	ENABLE
German	CAUSE	ENABLE	ENABLE
Russian	CAUSE	ENABLE	ENABLE

FIGURE 5.7. Frames from animations depicting three different kinds of events forming a CAUSE–ENABLE continuum and the types of descriptions English, German, and Russian speakers would likely use to describe them.

himself toward the line, but with some difficulty, leading the causer to help pull the causee over the line. In force dynamics, such a situation should be construed as one of enablement because the causee has a tendency for the result, and the causer assists the causee in achieving the result. Given that the scenarios on the far left and right represent clear cases of causal and enabling relations, we should expect a high degree of cross-linguistic agreement in how these situations are described: The situation on the far left should be described with CAUSE verbs and the situation on the far right with ENABLE verbs.

Where we should expect disagreement across languages is for animations such as the one referred to in the middle panel of Figure 5.7. In Wolff and Ventura (2009) and Klettke and Wolff (2003), such animations were used in a series of experiments to investigate the potential impact of grammatical constraints on verb preference, in particular, the use of CAUSE and ENABLE type periphrastic verbs in English, Russian, and German. As discussed earlier, it is expected that English speakers may look for, and take into account, not only the causee's intentions, but also the forces external to the causee, such as resistance to motion due to friction. With respect to the middle panel in Figure 5.7, in particular, because the causee is facing the line, it might appear as if the causee's intention is to cross the line. However, in terms of external forces, the dolly is physically difficult to pull, so there is no physical tendency to cross the line. The apparent intention of the causee and the physical difficulty of pulling him may lead to conflict in terms of the tendency of the causee. This conflict might lead English speakers to describe the scenes with either CAUSE or ENABLE type verbs.

For speakers of German and Russian, in contrast, it is expected that the tendency of the man in the middle panel of Figure 5.7 should be less ambiguous. Both German and Russian have rich case systems, so it is expected that in languages such as these, the external argument will often be agentive or at least self-energetic. As a consequence, the speakers of these languages should focus on the kinds of forces that are self-generated in the causer and causee, including intentions. In the middle scene in Figure 5.7, for example, the causee appears to want to cross the line. As a consequence, Russian and German speakers should prefer descriptions of this scene based on ENABLE type verbs.

These predictions were tested in studies involving 48 native speakers of English (Memphis), Russian (Moscow), and German (Hamburg), with 16 from each language (Wolff & Ventura, 2009; Klettke & Wolff, 2003). Twelve animations were used to create four CAUSE–ENABLE continuums. After watching an animation, participants chose which of two possible periphrastic causative sentences (e.g., *The man in red caused the man in green to cross the line*; *The man in red enabled the man in green to cross the line*), (or the option *"None of the above"*) best described the occurrence. The verbs in the sentences were based on one of several possible CAUSE (*cause, make,* or *force*) or ENABLE type verbs (*enable, let,* or *help*). The English versions of these sentences were translated into German and Russian for the German and Russian versions of this task.

The results supported our predictions. First, English, German, and Russian speakers preferred sentences with CAUSE verbs (88%, 81%, and 89%) over ENABLE verbs (12%, 3%, and 11%) for the animations that depicted clear cases of causation (e.g., far left panel of Fig. 5.7). Second, English, German, and Russian speakers preferred sentences with ENABLE verbs (81%, 81%, and 92%) over CAUSE verbs (14%, 15%, and 5%) for the animations that depicted clear cases of enablement (e.g., the animation on the far right in Fig. 5.7). The responses to the two far ends of the CAUSE–ENABLE continuum indicate that the languages share the same basic framework for distinguishing CAUSE and ENABLE interactions and that the animations depicted relatively clear examples of these types of causation in each language.

Of primary interest was how the speakers of each group would describe the intermediate situations. As shown in Figure 5.8, English speakers preferred CAUSE verbs whereas the

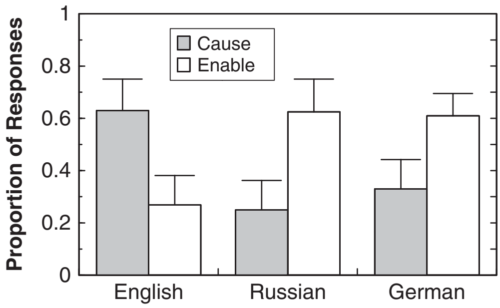


FIGURE 5.8. Proportion of times that English, Russian, and German speakers chose the CAUSE and ENABLE sentences for the intermediate animations in the CAUSE–ENABLE continuum, with associated 95% confidence intervals.

Russian and German speakers preferred ENABLE verbs. The results demonstrate how grammatical constraints on external arguments may have an impact on the range of situations described by a verb, and furthermore, how the range of situations described by rough translations across languages can vary across languages.

It is worth emphasizing that according to our account, these cross-linguistic differences are not due to differences in the meaning of CAUSE and ENABLE verbs in the different languages; rather, these cross-linguistic differences might be due to the semantics governing possible external arguments. Specifically, in languages that restrict the external argument position to entities that are self-energetic (as in German and, arguably, Russian), there may be a tendency to evaluate a situation only with respect to internally generated forces, such as intentions, which may lead the speakers to choose ENABLE verbs. When the entities in the external argument position are not so restricted (as in English) speakers may look for, and take into account, not only forces internal to the patient, but also external forces (e.g., gravity, friction), which may lead speakers to choose CAUSE verbs. In sum, differences at the grammatical level are likely to produce differences in the way related verbs are used in across languages.

SUMMARY

In this chapter we investigated the semantic criteria that determine the range of possible causers in English and other languages, including Mandarin Chinese, Korean, Russian, and German. We first proposed that Hawkins' ideas regarding the relationship between morphological case marking and the range of possible external arguments in English and German be extended: In particular, we proposed that languages with case systems (and relatively free word order) allow for a smaller range of external arguments than languages without case systems (and more fixed word order). Then, for each of these language types, the kinds of external arguments that are acceptable can be characterized in terms of the notion of force creation and the different ways in which force can be created (i.e., through energy conversion, physical contact, and force redirection). So, in languages that are more restrictive, external arguments must be capable of generating their own energy. In less restrictive languages, the external argument position is open to entities that use other methods of force creation, including physical contact and force redirection. Based on this proposal, we predicted that Korean, Russian, and German—because they have case systems—would require self-energetic entities as external arguments. In contrast, we predicted that in English and Chinese—lacking case systems and having relatively fixed word order—external arguments could be entities that create force through means other than strictly generating their own energy, resulting in a wider range of available external arguments.

We then reported three lines of research that provided evidence for the initiator hypothesis and force creation. In the first, we showed that the range of possible external arguments in English and Chinese was greater than in Korean. We further showed that external arguments in Korean had to be capable of generating their own energy, that is, capable of creating their own force through energy conversion. In the second line of research, we showed that when these differences in external arguments were realized in embedded clauses, they resulted in cross-linguistic differences in periphrastic causal expressions. Specifically, we showed that in English and Chinese, intermediary causees can be animate or inanimate, but in Korean, they can only be animate (reflecting the requirement that they be self-energetic). In the third line of research, we examined how the range of acceptable external arguments in a language could affect the kinds of verbs speakers preferred to describe causal situations. This time, English was contrasted with German and Russian. As predicted, we found that because external arguments in German and Russian are typically animate, speakers of these languages tend to focus on internal properties, such as an entity's intentions. As a consequence, for certain kinds of causal situations, speakers of German and Russian may prefer descriptions based on ENABLE rather than CAUSE verbs.

THE LANGUAGE–THOUGHT INTERFACE

In a recent typological study of 408 languages, Song (1996) observed that all had productive methods for the linguistic expression of causation. In a far more modest study, Wolff et al. (2005) observed that several languages (German, Spanish, Russian, and Arabic) have near translations of many of the periphrastic causative verbs in English (e.g., *cause, make, force, get, allow, enable*), implying that these languages make many of the same fine-level semantic distinctions, or at the very least, that they distinguish the causal concepts of CAUSE, ALLOW, and PREVENT. We conclude from these results, and from other evidence, that some form of the concept of CAUSE is universally shared by the speakers of all languages.

According to the dynamics model (Wolff, 2007; Wolff & Song, 2003), as discussed earlier, people may represent causal relations in an analog fashion, namely, in terms of configurations or chains of forces. We propose that such an account suggests how causal relations might be represented on the "thought" side of the language–thought interface. On the language side, word meanings might specify features for determining which linguistic labels can be applied to the analog representations on the thought side. For example, assuming that causal relations are represented in terms of configurations of force, the label "cause" might be licensed when the configurations of forces have the following properties: (1) forces associated with the causer and causee are in opposition, (2) the force associated with the causee is directed away from the endstate, and (3) the causee progresses toward the endstate (see Wolff, 2007). Such criteria might be represented in the brain in the form of lists of features, or otherwise discrete symbols. Hence, the language–thought interface might involve an alignment between discrete, digital units in language with analog structures in thought.

It is possible that the criteria for individual verbs might be much the same across languages, and languages could still clearly differ in how such verbs are used. Our explanation for this variation is that it is caused by differences across languages in the semantics of external arguments. According to the dynamics model, the distinction between CAUSE and ALLOW/ENABLE depends on the tendency of the causee. In this chapter we have proposed that the speakers of different languages may draw on different aspects of a situation in determining the tendency of the causee, with speakers of initiator languages focusing on forces that are produced from energy conversion and speakers of participant languages factoring in a wider range of ways to produce force. One question to be addressed in future research is whether these differences in the determination of tendency are manifested in nonlinguistic activities, such as in the perception of events or in causal reasoning.[3]

CONCLUSIONS

We began this chapter by noting that people have strong expectations about possible causers. We argued that an explanation of this phenomenon was not only a matter of psychology and philosophy, but also a matter of linguistics. As discussed in this chapter, the selection of possible causers depends on the semantics associated with grammatical structures, which appear to differ across languages. These results emphasize how even the most fundamental of concepts, such as CAUSE, are likely to be expressed differently across languages as a consequence of the interaction between word meanings and the rest of the language system.

Notes

1. We thank Beth Levin for making us aware of this idea and pointing us to Hawkins' (1985) work.
2. In Guilfoyle's (2000) account, the two types of languages reflect a syntactic parameter on subjects. She suggests that in Type A languages, subjects raise to the specifier of TP (Tense Phrase) in the underlying syntactic structure, where they can be checked for time (i.e., whether they were the initiator of the event), whereas in Type B languages, the subject raises to the specifier of AgrS (Subject Agreement), where it is checked only for being a participant in the event.
3. See Wolff, Jeon, and Yu (2009) for an initial investigation into this possibility.

References

Achard, M. (2001). Causation, constructions, and language ecology: An example from French. In M. Shibatani (Ed.), *The grammar of causation and interpersonal manipulation* (pp. 127–156). Philadelphia: John Benjamins.

Alexiadou, A., & Schäfer, F. (2006). Instrument subjects are agents or causers. In D. Baumer, D. Montero, & M. Scanlo (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 40–48). Somerville, MA: Cascadilla Proceedings Project.

Asimov, I. (1988). *Understanding physics*. New York: Barnes & Noble.

Baron, N. S. (1977). *Language acquisition and historical change*. New York: North-Holland.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99*,65–382.

Comrie, B. (1989). *Language universals and linguistic typology*, 2nd ed. Chicago: University of Chicago Press.

Copley, B. (2005). Ordering and reasoning. *The MIT working papers in linguistics*. Cambridge, MA: MIT Press.

Cotterell, B., & Kamminga, J. (1990). *Mechanics of pre-industrial technology*. Cambridge, UK: Cambridge University Press.

Craig, C. G. (1977). *The structure of Jacaltec*. Austin: University of Texas Press.

Cruse, D. A. (1973). Some thoughts on agentivity. *Journal of Linguistics, 9*, 11–13.

Delancey, S. (1984). Notes on agentivity and causation. *Studies in Language, 8*, 181–213.

Etzioni-Halevy, E. (2006). *The garden of rush*. New York: Plume.

Fausey, C.M., & Boroditsky, L. (2008). English and Spanish speakers remember causal agents differently. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 64–70). Austin, TX: Cognitive Science Society.

Fillmore, C. J. (1968). The case for case. In E. Back & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York: Holt, Rinehart, & Winston.

Folli, R., & Harley, H. (2007). Teleology and animacy in external arguments. *Lingua, 118*, 190–202.

Grimm, S. (2007). The bounds of subjecthood: Evidence from instruments. *Berkley Linguistics Society*, 33.

Guilfoyle, E. (2000). Tense and N-features in modern Irish. In A. Carnie & E. Guilfoyle (Eds.), *The syntax of verb initial languages*. Oxford: Oxford University Press.

Hawkins, J. A. (1985). *A comparative typology of English and German: Unifying the contrast.* Austin: University of Texas Press.

Jackendoff, R. S. (2002). *Foundations of language*. Oxford: Oxford University Press.

Jackendoff, R. S., & Culicover, P. (2003). The semantic basis of control in English. *Language, 79*, 517–556.

Kamp, H., & Rossdeutscher, A. (1994). Remarks on lexical structure and DRS construction. *Theoretical Linguistics, 20*, 97–164.

Kearns, K. (2000). *Semantics*. London: Macmillan Press.

Klettke, B., & Wolff, P. (2003). Differences in how English and German speakers talk and reason about CAUSE. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the twenty-fifth annual conference of the Cognitive Science Society* (pp. 675–680). Mahwah, NJ: Lawrence Erlbaum.

Lambrecht, K. (1994). Information structure and sentence form: Topic, focus, and the mental representation of discourse referents. In *Cambridge studies in linguistics 71*. Cambridge, UK: Cambridge University Press.

LaPolla, R. J. (1993). Arguments against 'subject' and 'direct object' as viable concepts in Chinese. *Bulletin of the Institute of History and Philology, 63*, 759–813.

Levin, B. (1994). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Levin, B. (2007). The lexical semantics of verbs I: Introduction. 2007 Summer Institute of Linguistics at Stanford University, handout.

Levin, B., & Rappaport-Hovav, M. (1988). Non-event –er nominals: A probe into argument structure, *Linguistics, 26,* 1067–1083.

Levin, B., & Rappaport-Hovav, M. (1995). *Unaccusativity: At the syntax lexical semantics interface*. Cambridge, MA: MIT Press.

Levin, B., & Rappaport-Hovav, M. (2005). *Argument realization*. Cambridge, UK: Cambridge University Press.

Marantz, A. P. (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Nilsen, D. (1973). *The instrumental case in English*. The Hague: Mouton.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking, Penguin Group.

Polinsky, M., & Potsdam, E. (2003). Backward control: Evidence from Malagasy. In A. Rackowski and N. Richards (Eds.), *Proceedings of the eighth meeting of the Austronesian Formal Linguistics Association* (pp. 257–272). Cambridge, MA: MITWPL.

Radford, A. (1988). *Transformational grammar: A first course*. Cambridge, UK: Cambridge University Press.

Radford, A. (1997). *Syntactic theory and the structure of English: A minimalist approach*. Cambridge, UK: Cambridge University Press.

Reinhart, T. (2002). The theta system—an overview. *Theoretical Linguistics, 28,* 229–290.

Schlesinger, I. M. (1979). Cognitive structures and semantic deep structures: The case of the instrumental. *Journal of Linguistics, 15,* 307–324.

Schlesinger, I. M. (1989). Instruments as agents: On the nature of semantic relations. *Journal of Linguistics, 25,* 189–210.

Shibatani, M. (1976). The grammar of causative constructions: A conspectus. In M. Shibatani (Ed.), *Syntax and semantics, Vol. 6: The grammar of causative constructions* (pp. 1–40). New York: Academic Press.

Song, J. J. (1996). *Causatives and causation: A universal-typological perspective*. New York: Longman.

Song, S. C. (1988). *Explorations in Korean syntax and semantics*. Berkeley: University of California, Institute of East Asian Studies.

Soto, V. V. (2001). Some constraints on Cora causative constructions. In M. Shibatani (Ed.), *The grammar of causation and interpersonal manipulation* (pp. 197–244). Philadelphia: John Benjamins.

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

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning and function*. Cambridge, UK: Cambridge University Press.

van Voorst, J. (1996). Some systematic differences between Dutch, French, and English transitive constructions. *Language Sciences, 18,* 227–245.

Wierzbicka, A. (1988). *The semantics of grammar*. Amsterdam: John Benjamins.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition, 88,* 1–48.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General, 136,* 82–111.

Wolff, P., Jeon, G., & Yeh, K. (2006). Causal agents and the individuation of events in English, Chinese, and Korean. *Proceedings of the 5th International Conference on Cognitive Science* (pp. 213–214). Mahwah, NJ: Lawrence Erlbaum.

Wolff, P., Jeon, G., Klettke, B., & Yu, L. (2009). Causal agents in English, Korean and Chinese: The role of internal and external causation. *Language and Cognition, 1,* 165–194.

Wolff, P., Klettke, B., Ventura, T., & Song, G. (2005). Categories of causation across cultures. In W. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside of the lab: Festschrift in honor of Douglas L. Medin* (pp. 29–48). Washington, DC: American Psychological Association.

Wolff, P., & Song, G. (2003). Models of causation and the semantics of causal verbs. *Cognitive Psychology, 47,* 276–332.

Wolff, P., & Ventura, T. (2009). When Russians learn English: How the semantics of causation may change. *Bilingualism: Language and Cognition, 12,* 153–176.

Woodward, J. (2007). Interventionist theories of causation in psychological perspective. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 19–36). Oxford: Oxford University Press.

Young, H. D., & Freedman, R. A. (1999). *University physics,* 10th ed. Reading, MA: Addison-Wesley.

6 THE LANGUAGE-SPECIFICITY OF CONCEPTUAL STRUCTURE
Path, Fictive Motion, and Time Relations

Jürgen Bohnemeyer

This chapter addresses the encoding of spatial semantics at Conceptual Structure (CS) in the framework proposed by Jackendoff (1983, 1987, 1996, 2002). The central question concerns the aspects of the representation of space at CS that are universal and therefore presumably innate.

Jackendoff envisions CS as a language-independent faculty of cognition that generates non-iconic conceptual representations of an algebraic internal structure (a recursive predicate-argument calculus that is syntactically different from both language and predicate logic). Reasoning and any transfer of information between different peripheral systems is divided between CS and another module of higher cognition, Spatial Structure (SpS).[1] SpS encodes geometric properties in an "image-schematic" fashion. SpS representations are primarily the product of high-end visual processing, but receive input in other modalities as well, and are themselves a-modal. Jackendoff assumes that language primarily interfaces with CS. Linguistic meaning is a mapping between the syntactic and phonological representations of utterances and some corresponding CS representations. Lexical meaning components that involve shape, "manner of motion" (Talmy, 2000b), and certain other spatial properties are fully interpreted at SpS (perhaps via some sort of placeholders at CS); but all aspects of syntactic structure map exclusively into CS. The exact division of labor between CS and SpS remains very much an open question within this framework.

My concern here is specifically with the representation of Motion events in language and cognition. Jackendoff (1983, 1990) has advanced a number of arguments to the effect that CS encodes notions of Translational Motion (T-Motion) and Path, based on English data. I argue in the following on the basis of evidence from Yucatec Maya that these arguments do not apply universally, and that Yucatec Motion event descriptions do not involve a semantics based on T-Motion and Path (henceforth, a "Path semantics"), but merely a State-Change semantics. In the account proposed here, cognitive representations of Motion are comparable between English and Yucatec at the level of SpS, but not at CS.

T-Motion involves a homomorphic mapping from the time course of the Motion event into the Path traversed (e.g., Krifka, 1998; Zwarts, 2005), as depicted schematically in Figure 6.1. T-Motion must be encoded on *some* level of cognition—but to what extent is it encoded in language? It has often been assumed that linguistically, Motion is represented as a special case of State-Change—Change of Location (e. g., Miller & Johnson-Laird, 1976; Dowty, 1979).[2] Location-Change

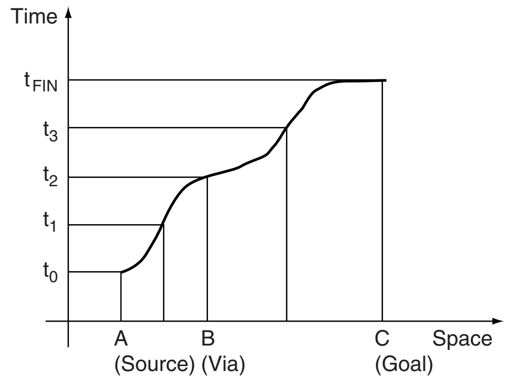


FIGURE 6.1. Space–time diagram of translational motion.

representations decompose Motion events into State-Change event structures and Locative relations that characterize their beginning or end states, rendering, e.g., the meaning of *go under the table* as something like "come to be/ end up under the table" or the meaning of *leave the house* as "cease to be inside the house" or "end up outside the house," etc.

Jackendoff (1983: 170–174; 1990: 91–95) argues against a general reduction of Motion semantics to Location-Change. He proposes that representations of Motion events at CS require a primitive conceptual function of T-Motion (represented by the conceptual function GO) and the set of five basic Path functions TO and FROM (for "Bounded Paths," i. e., Paths defined in terms of their end points), VIA (with "Routes," i.e., Paths defined in terms of Places on them in a nonterminal position), and TOWARD and AWAY-FROM (with "Directions," i.e., Paths defined in terms of their orientation in some Frame of Reference). The alternative is illustrated in (1): Is the meaning of (1a) conceptually encoded as in (1b) or as in (1c)? "INCH" in (1c) stands for the conceptual function of State-Change, represented by "BECOME" in work within or based on the Generative Semantics tradition (e.g., Dowty, 1979). (1c) also captures the meanings of descriptions such as *X came to be at Y* or *X ended up at Y*; so another way of framing the issue at hand is in terms of the question of whether or to what extent (1a) is synonymous with such utterances.

(1)a. *X went to Y*
b. $[_{\text{Event}}\ \text{GO}\ ([_{\text{Thing}}\ X], [_{\text{Path}}\ \text{TO}\ ([_{\text{Place}}\ \text{AT}\ ([_{\text{Thing}}\ Y])])])]$
c. $[_{\text{Event}}\ \text{INCH}\ ([_{\text{Thing}}\ X], [_{\text{State}}\ \text{BE}\ ([X], [_{\text{Place}}\ \text{AT}\ ([_{\text{Thing}}\ Y])])])]$

Henceforth, I refer to representations with the format of (1b) as "Path semantics" and to analyses along the lines of (1c) as "State-Change semantics" or, more specifically, "Location-Change semantics." Jackendoff advances three arguments in favor of a Path semantics for Motion event descriptions. First, T-Motion is clearly a cognitive primitive, so why should CS not encode it as well?

> . . .we can perceive an object as in continuous motion without knowing anything about the endpoints of its motion. It moreover appears (Marr, 1982) that the visual system contains specialized motion detectors that are rather independent of the channels that individuate and localize objects. If motion is a primitive even in elementary aspects of visual cognition, why should conceptual structure be so stingy as to provide no way to encode it? (Jackendoff, 1990: 94)

But this argument can be turned around to buttress the case *against* Path semantics: If T-Motion and Path information are already adequately encoded by other systems of cognition, and there is another way of representing Motion linguistically—namely, in terms of Location Change—then why duplicate the information at CS? Jackendoff's remaining two arguments, however, directly challenge the notion that Motion can be adequately represented as Change of Location in language. Bounded Path functions representing Motion FROM Source and/or TO Goal are straightforwardly decomposed enough along the lines of (1c). But such an analysis seems much less natural for Route Path functions as in (2), where location at the Ground defines neither the Source nor the end state of the event, but some state of the Figure in between:

(2)a. *The eagle soared across the canyon*
b. *The train went through the tunnel*
c. *The expedition crossed the river*
d. *The horse jumped over the fence*

I would like to add a similar problem, which arises with complex Motion descriptions in which multiple Path functions are combined in a single verb phrase, as in (3). State-Change descriptions do not appear to specify both the source and the target state (rather than to treat one as the negation of the other), unless they involve Motion metaphors, as in (4).

(3) *The supporters went from the meet-up to the rally*
(4) *The lights went from green to red*

Jackendoff's third argument concerns the use of Path functions in what Talmy (1996, 2000a) has called Fictive Motion metaphors: state descriptions that do not encode, and therefore cannot be reduced to, Location-Change:

(5)a. *The highway extends from Denver to Indianapolis*
b. *The house faces away from the mountains*
c. *The firehouse is across the street from the library* (Jackendoff, 1983: 167–172)

My working assumption is that the phenomena illustrated in (2)–(5) robustly support the case for Path semantics in English Motion event descriptions. The question I wish to address in the following is to what extent these arguments extend to other, and perhaps all, languages. As my test case, I choose Yucatec Maya. The evidence to be examined includes Location-Change descriptions that are true both of Motion events and of events involving, for instance, objects emerging into or disappearing from spatial configurations.

In "satellite-framed" (Talmy, 2000b) languages such as English, a Location-Change verb phrase can be constituted by combining a manner-of-motion verb such as *walk* or *slide* with a Path-denoting satellite or prepositional phrase [*walk in*(*to the room*); *slide down/off the table*]. Yucatec behaves like a "verb-framed" language in this respect: Only verb phrases projected from Location-Change verbs—verbs corresponding to the English "Path verbs" (Talmy, 2000b) *come, go, enter, exit, ascend, descend,* and *pass*—can be used in reference to Location-Change events. In fact, as discussed in detail in the next section, in contrast to better-studied verb-framing languages such as Japanese, Spanish, and Turkish, in Yucatec, Ground phrases [the expressions of the Place with respect to which Location (Change) of the Figure is described] do not encode Locative or Path relations, but merely specify spatial regions that may serve as "landing sites" for such relations. If Path relations are lexicalized in Yucatec, they must be lexicalized in the Location-Change verbs—just as Path relations are lexicalized, on Talmy's analysis, in the English and Spanish equivalents of these verbs. But do Yucatec Location-Change verbs have Path semantics? Evidence that they do not comes from the fact that Motion event descriptions formed with the Location-Change verbs can be used in reference to events involving not only Figure Motion, but also Ground Motion or emergence/disappearance of Figure or Ground, discussed later. Such uses of Location-Change descriptions were first documented by Kita (1999) for Japanese. Consider Figure 6.2. The circle moves and ends up enclosing the square. Example (6), but not its literal English translation, can be used to describe the scenario in Figure 6.2:[3]

(6)	Shikaku-ga	en-ni	hai-ta.
JPN	square-NOM	circle-LOC	enter-PAST

'The square entered the circle.' (Kita, 1999: 344)

Kita concludes that the verb *hairu* really means "become inside," rather than "enter." As will be shown later, similar phenomena occur in Yucatec on a broader scale, involving not just "enter" and "exit" verbs, but also verbs corresponding to *ascend, descend,* and *pass*. This provides direct evidence against Path semantics in Motion event descriptions formed with these verbs. As far as descriptions formed with these Location-Change verbs are concerned, a Yucatec speaker and an English speaker looking at the same Motion event in extralinguistic reality must form different CS representations to talk about it, if we assume, as Jackendoff does, that linguistic meaning is a direct mapping from syntax into CS. If the CS representations that "interpret" Yucatec Motion descriptions encoded Translational Motion of the Figure along a Path, the Yucatec description would be incompatible

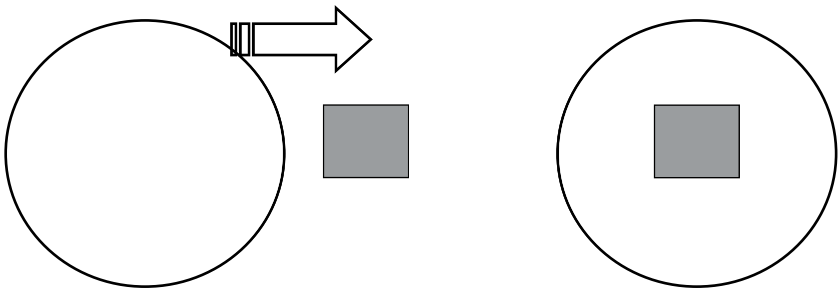


FIGURE 6.2. A scenario for (6).

with non-Figure-motion scenarios, just like their English expressions.

In addition to presenting direct counterevidence against Path semantics in Yucatec Motion event descriptions, I also show that the arguments that favor a Path semantics for English do not apply to Yucatec. Descriptions of Motion events involving Route Paths are generally vague, since they all employ the same Location-Change verb, *máan* "pass" (4.1). Because there are no verbs that lexicalize Location-Change with respect to multiple Grounds (in Yucatec or, as far as I am aware, any other language), combinations of multiple Path functions ina single verb phrase are impossible. Consequently, a journey from Source A to Goal B is described by a multiclause sequence along the lines of "She left A, and eventually she arrived at/on/in B" (4.2). And there is no evidence of Fictive Motion metaphors in Yucatec. There are metaphoric uses of Location-Change expressions, but these have much more restricted domains of use that do not support an analysis in terms of Path meanings (4.3). The case against Path semantics in Yucatec is further buttressed with indirect evidence from spatio-temporal metaphors. As discussed later, Yucatec lacks temporal connectives with meanings such as "after" and "before," which on localist accounts draw on Motion metaphors (e.g., Clark, 1973; Traugott, 1978). To round out the picture, L2-Spanish data from Yucatec native speakers are briefly considered later. I conclude that there is no *linguistic* evidence for the encoding of Path semantics in Yucatec.

Do Yucatecans require CS representations of Translational Motion and Path at CS to *reason* about Motion? Although this question cannot be answered conclusively in this article, I will argue that SpS may well be able to afford the requisite functions. I also briefly examine the typological conditions of the framing of Motion as State-Change—making it clear that the case of Yucatec is probably not exotic. Finally, I discuss possible implications of the language-specificity of Motion semantics for Jackendoff's framework, drawing in particular on the Thematic Relations Hypothesis, which accords Path semantics a special role built into the very architecture of CS.

THE GRAMMAR OF MOTION EVENT DESCRIPTIONS IN YUCATEC

Yucatec is a Mayan language spoken by over 800,000 people on the Yucatan peninsula in Mexico and Belize. Like all Mayan languages, Yucatec is a polysynthetic language, i.e., a language in which grammatical functions are predominantly expressed by the structure of word forms rather than or in addition to combinations of words or phrases. It is exclusively head-marking (i.e., to the extent that the relation between the head of a phrase and a dependent is morphologically marked, it is marked on the head), shows productive incorporation of nouns and adverbs into the verbal complex and productive verb compounding, and has rich valence changing and voice morphology (i.e., morphological derivations that change the argument structure of verbs and inflections that change their linking properties, such as a passive). Yucatec is verb-initial and almost exclusively head-initial. The language has a typologically unusual argument marking split in intransitive clauses governed by aspect-mood marking (see Bohnemeyer, 2004 and references therein).

Most of the work reported on here was conducted in annual field trips between 1995 and 2004. The main consultants were six adult native speakers, one woman (age 30 in 2004) and five men (between age 27 and 56 in 2004), in the municipal district of Felipe Carrillo Puerto in the state of Quintana Roo, Mexico.

The following two subsections provide background information on the structure of the verbal core[4] (2.1) and the Ground phrase (2.2) in Motion event descriptions. Two facts are introduced that are critical prerequisites to the discussion of the framing of Motion as Change of Location (CoL) in Yucatec: Verbal cores of Motion event descriptions must be headed by CoL verbs and Ground phrases are strictly Path neutral.[5]

The Structure of the Verbal Core in Motion Event Descriptions

In terms of Talmy's (1985, 2000b) lexicalization typology, Yucatec may be considered in first approximation (but see below!) a "verb-framed" language. For a clause to be able to describe events of Motion, its main verb must encode CoL. The verbs that are most commonly used in this role are listed in Table 6.1. The English glosses used in Table 6.1 and throughout this chapter do not adequately capture the hypothesized CoL semantics of the verbs; they merely serve to facilitate reading here. Evidence in support of the absence of the Path functions in the semantics of the verbs comes primarily from their applicability to events that do not involve Figure Motion, discussed later. All verbs are base-intransitive, but produce derived causative stems. The spatial semantics of the verbs is captured by a Place function, denoting a spatial region projected from the Ground. On the analysis presented in this chapter, the output of this Place function is mapped into an event representation, not by a Path function, as in (1b), but by a Locative state function, which characterizes the source state, target state, or a transitional phase in between in a State-Change event description, as in (1c). The corresponding Path functions are added in parentheses for ease of processing. *Tàal* "come" and *u'l* "return" (and their causative counterparts) assign the role of Ground to the deictic center.[6] *Bin* "go" (and *bis* "bring") can be interpreted with respect to either the deictic center or some Place specified in context as Ground. The remaining verbs combine with Ground phrases or track Places anaphorically from context in the absence of a Ground phrase in the clause.

TABLE 6.1. The Basic CoL Verbs of Yucatec

CoL Root	Causative Stem	Place Function of Ground	Locative Description Characterizes	Ground Encoding
tàal "come";*u'l* "return"	*tàas* "bring";*u's* "return"	AT	Target state ("TO")	Inherently deictic
k'uch "arrive"	*k'uhs* "cause to arrive"			Lexical
bin "go"	*bis* "take"		Source state ("FROM")	Inherently indexical
luk' "leave"	*lu's* "remove"			Lexical
lúub "fall"	*lu's* "fell," "drop"	ON/ABOVE	Target state ("TO")	
na'k "ascend"	*na'ks* "lift"			
em "descend"	*èens* "pluck," "lower"		Source state ("FROM")	
lík' "rise"	*li's* "lift"			
òok "enter"	*òoks* "insert"	IN	Target state ("TO")	
hóok' "exit"	*ho's* "extract"		Source state ("FROM")	
máan "pass"	*máans* "pass"	Underspecified	N.A.[a]	

[a]Compare the section on The Treatment of Routes.

The further discussion is limited to the verbs in Table 6.1 because of the systematic character of the set. There are, however, other verbs that may occur in verbal cores denoting CoL. These include *náak* "reach," "extend up to," which is sometimes used as an alternative to *k'uch* "arrive." *Náachtal* "become distant" may be used in some contexts instead of *bin* "go" or *luk'* "leave." *Sùut* "turn," "spin," "return," the antipassive form of the transitive root *sut* "turn," is basically an activity verb, but is recruited by metaphoric extension for the purpose of expressing return to a Place not necessarily identical with the deictic center. It thus fills a gap in the system of Table 6.1, given the deictic specialization of *u'l*. There are also transitive roots of caused CoL, in particular in the domain of insertion and extraction and in the ballistic Motion domain. One example is *pul* "throw." For the interaction between verbs and Ground-denoting adjuncts in CoL-denoting verbal cores, it makes no difference whether the cores are headed by such transitive verbs or by the CoL verbs in Table 6.1; hence further discussion is restricted to the latter.

The roots in Table 6.1 belong to two different inflectional classes, both of which host exclusively (noncausative) State-Change verbs (cf. Bohnemeyer, 2002: 153–215; 2004 and references cited there). State-Change is attested on the basis of criteria such as the one illustrated in (7)–(9): combinations of CoL verbs such as *bin* "go" (7), *òok* "enter" (8), and *hóok'* "exit" (9) with the progressive aspect marker *táan* [fused with the third-person cross-reference marker *u*- in (7) and (9)] allow only for prospective (prestate reference) interpretations, not for imperfective interpretations, as they would if the verbal core had process semantics. The diagnostic of prospective reference in (7)–(9) is paraphrased with the prospective aspect marker *mukah*.[7,8]

(7) Túun bin Juan Carrillo=e',
PROG:A3 go Juan Carrillo=TOP
káa=h-k'àas-chah u=kòombi.
CON=PRV-bad-INCH.CMP(B3SG) A3=van
Káa=t-y=a'l-ah=o', mukah bin.
CON=PRV-A3=say-CMP(B3SG)=D2 PROSP(B3SG) go(INC)
'Juan was going to Carrillo, (when/and then) the van broke down. At this moment (lit. when it said that), he was going to go.'

(8) Pedro=e' táan y=òok-ol t-u=nah-il=e',
Pedro=TOP PROG A3=enter-INC PREP-A3=house-REL=TOP
káa=t-y=il-ah=e', hach sùusyo u=nah-il.
CON=PRV-A3=see-CMP(B3SG)=D3 really dirty(B3SG) A3=house-REL
Káa=t-y=a'l-ah=o', ma' òok-ok=i'.
CON=PRV-A3=say-CMP(B3SG)=D2 NEG enter-SUBJ(B3SG)=D4
Mukah òok-ol.
PROSP(B3SG) enter-INC
'Pedro, he was entering his house, (when/and then) he saw it, his house was very dirty. At that moment (lit. when it said that), he hadn't entered yet. He was going to enter.'

(9) Hun-túul uy=alak' wakax don Valen=e',
one-CL.AN A3=CL.domestic.animal cow don Valen=TOP
túun hóok'-ol te=koràal=o',
PROG:A.3 exit-INC PREP:DET= corral=D2
káa=h-k'uch u=yúum-il.
CON=PRV-arrive(CMP)(B3SG) A3=master-REL
Káa=t-y=a'l-ah=o', mukah hóok'-ol.
CON=PRV-A3=say-CMP(B3SG)=D2 PROSP(B3SG) exit-INC
'One of don Valen's cows, it was exiting the corral, (when/and then) its owner arrived. At that moment (lit. when it said that), it was going to exit.'

Other diagnostics of State-Change semantics include compatibility with the stative resultative derivation in *–a'n* and incorporation of the universal quantifier to encode complete affectedness of the theme.

Process verbs are employed in Motion event descriptions to denote "manners of motion" (Talmy, 2000b). An example is *xîiknal* "flutter," "fly (in the manner of birds)" in (10)–(12):

(10) Le=ch'íich'=o' túun xíiknal y=óok'ol le=che'=o'.
DET=bird=D2 PROG:A3 fly A3-top DET=wood=D2
'The bird is flying (i.e., circling!) above the tree.'

(11) Le=ch'íich'=o' xíiknal-il h-úuch uy=em-el
DET=bird=D2 fly=REL PRV-happen(B3SG)A3=descend-INC
te=che'=o'.
PREP:DET=wood=D2
'The bird flew down from the tree [lit. flyingly (is how) it happened to descend with respect to the tree].'

(12) Le=ch'íich'=o' h-em u=xíiknal te=che'=o'.
DET=bird=D2 PRV-descend(B3SG) A3=fly PREP:DET=wood=D2
'The bird flew down from the tree (lit. it descended flying with respect to the tree).'

In clauses formed with a Manner verb as the only verb, as in (10), Ground phrases merely refer to the Location of the event; CoL is neither entailed nor implicated. There are two constructions that are regularly used to integrate Manner information: the Manner focus construction (Bohnemeyer, 2002: 123–125) exemplified in (11), in which the CoL-denoting verbal core is subordinate to the Manner predicate in a cleft-like structure, and the gerundial construction (Bohnemeyer,

TABLE 6.2. Some Common Manner-of-Motion Verbs of Yucatec

Propulsiveness Property	Selective Restrictions	
	Figure Must Be Animate	Figure Need Not Be Animate
Propulsive	*áalkab* "run"; *bàab* "swim"; *xîimbal* "walk"; . . .	*balak'* "roll"; *háarax* "slide"; . . .
Nonpropulsive	*sîit'* "jump"; *xîiknal* "flutter," "fly"; *òokot* "dance"; . . .	*mosòon* "whirl," "revolve"; *péek* "move"; *pi'k'* "shake," "twirl"; *úumbal* "swing," "rock," *walak'* "turn," "revolve"; . . .

2002: 100–101) illustrated in (12), in which the Manner-denoting core1 is embedded as an adjunct. Table 6.2 provides an overview of the Yucatec Manner-of-Motion verbs, sorting them in terms of selectional restrictions regarding the Figure's animacy and the property of "propulsiveness"—propulsive Manners may cause CoL, whereas nonpropulsive ones involve Motion with respect to some axis of the Figure.

The facts reviewed so far establish a broad similarity between Yucatec and better-studied languages of Talmy's verb-framed type such as Japanese, Spanish, and Turkish, in that verbs that appear to be translational equivalents of "path-conflating" verbs such as *enter*, *exit*, *ascend*, and so on are required to form Motion descriptions. However, there are two important differences: First, as discussed in the next section, Path distinctions are not reflected outside the verb; so Yucatec at the very least exhibits a more radical kind of verb framing. But second, evidence is presented later suggesting that the Yucatec CoL verbs do not, in fact, encode Path functions either—and that these notions are therefore not lexicalized in Yucatec.

The Structure of the Ground Phrase

Ground phrases denote Places with respect to which Location and Motion (or Location-Change) of the Figure are described. If the Ground object is denoted by a common noun (as opposed to a toponym), the Ground phrase is headed by a preposition or relational noun. The prepositions that occur in Ground phrases are the generic *ti'* and *ich(-il)* "in" (cf. Bohnemeyer & Stolz, 2006; Levinson, Meira, & The Language and Cognition Group, 2003). The relational nouns found most commonly in Ground phrases are listed in Table 6.3.[9]

In better-studied exemplars of both the satellite-framed and the verb-framed language type, the Ground phrase denotes a Path or Locative function. Thus, in (1), repeated below for convenience, the PP *to Y* maps the Ground object denoted by *Y* into the Place denoted by *at Y* and the latter into a Path that has that Place as its end point.

(13) a. *X went to Y*

b. $[_{Event}\ \text{GO}\ ([_{Thing}\ \text{X}], [_{Path}\ \text{TO}\ ([_{Place}\ \text{AT}\ ([_{Thing}\ \text{Y}])])])]$

For verb-framed languages such as Japanese, Spanish, or Turkish, this has the consequence of actual "double-marking" of Path in both the verb and the Ground phrase (cf. Bohnemeyer, Enfield, Essegbey, Kita, Ibarretxe-Antuñano, Lüpke, et al., 2007). Consider the Spanish paradigm illustrated in (14):

(14)a.	El	carro	de	juguete	esta-ba	en	la	caja
SPA	DEF	cart	of	toy	be.at-PAST.IMPF3SG	in	DEF	box

"The toy car was in thebox"

b.	El	carro	de	juguete	entr-ó	en	la	caja
	DEF	cart	of	toy	enter-PAST.PRV3SG	in	DEF	box

"The toy car entered (lit. in) the box"

c.	El	carro	de	juguete	sali-ó	de (/*en)	la	caja
	DEF	cart	of	toy	exit-PAST.PRV3SG	of in	DEF	box

'The toy car exited (lit. from) the box'

The PP *en la caja* "in(to) the box" conflates Locative ("in"; 14a) and Goal ("into"; 14b) functions —a pattern of syncretism common across languages according to Clark (1973)—but is incompatible with the Source function ("out of") in (14c). Compare this to the Yucatec equivalents in (15):

(15)a.	Le=kàaro=o'	ti'=yàan	ich/ti'	le=kàaha=o'
	DET=cart=D2	PREP=EXIST(B3SG)	in/ PREP	DET=box=D2

'The cart, it is in the box'

b.	Le=kàaro=o'	h-òok	ich / ti'	le=kàaha=o'
	DET=cart=D2	PRV-enter(B3SG)	in / PREP	DET=box=D2

'The cart, it entered (lit. in) the box'

c.	Le=kàaro=o'	h-hóok'	ich /ti'	le=kàaha=o'
	DET=cart=D2	PRV-exit(B3SG)	in / PREP	DET=box=D2

'The cart, it exited [lit. in] the box'

TABLE 6.3. Frequent Relational Nouns in Yucatec Ground-Denoting Phrases[a]

Construction	Relational Noun	Gloss
$[\text{SetA}_i - N_{rel}\ NP_i]_{GroundP}$	*àanal*	Under
	iknal	At
	óok'ol	On/over
$[$*ti'* $[\text{SetA}_i - N_{rel}\ NP_i]]_{GroundP}$ or $[N_{rel}($-*il*$)$ *ti'* $NP]_{GroundP}$	*chúumuk*	Center
	háal	Edge
	nak'	Belly
	(ba')pàach	Back/outside
	(ak)táan	Front
	tséel	Side
	ts'u'	Core
	xno'h	Right
	xts'i'k	Left
	xùul	End
	yáam	Interstice

[a]GroundP, Ground phrase; NP, Ground-denoting nominal; $N_{rel,}$ relational noun; SetA, cross-reference marker "Set A."

In (15), *ich(-il)* "in" alternates with the generic preposition *ti'*. Neither *ich(-il)* nor *ti'* distinguishes among Locative (15a), Goal (15b), or Source (15c) functions; they are compatible with Route (VIA) and Direction (TOWARD/AWAY-FROM) functions as well. *Ich(-il)* is compatible with all of these interpretations because it does not encode any of them—it is Path neutral. The Ground phrases in (15) merely denote a Place projected from the Ground object, the box. *Ich(-il)* specifies the inside of the box as this Place; *ti'* is semantically compatible with any spatial region projected from the box. Either way, it is the verb that determines the role the Place has in the CoL description. This analysis generalizes to all Yucatec Ground phrases. The examples in (16) illustrate the point for Ground phrases headed by the relational noun *óok'ol* "on," "above." The Ground phrase in (16a) refers to the location of the rolling event, the one in (16b) to the Goal of a CoL event, and the one in (16c) denotes a Source.

(16) a. . . . h-tàal | u=balak' | y=óok'ol le=pak'=o'
PRV-come(B3SG) | A3=roll | A3=on DET=brickwork=D2
'. . . it came rolling on the wall'

b. H-na'k | y=óok'ol | le=che'=o'
PRV-ascend(B3SG) | A3=on | DET=wood=D2
'It went onto the piece of wood'

c. Káa=h-em | y=óok'ol | le=che'=o'. . .
CON=PRV-descend(B3SG) | A.3=on | DET=wood=D2
'It went down from the piece of wood. . .'

Yucatec Ground phrases do not encode Path functions. The role of the Ground in the CoL event is encoded by the predicate entailing a Locative relation that characterizes the source or target state of the CoL event (an exception is *máan* "pass"; cf. The Treatment of Routes section for a treatment). Also compatible with the facts presented so far is an analysis of the Ground phrase as invariably encoding event Locations, i.e., Locations at which the CoL event described by the main verb takes place, rather than Places at which the Figure is located at the beginning or end of the event. Under an event Location analysis, the role of the event Locations in the semantic composition of the event description is left to pragmatic inferences. Based on this account, both (17a) ("The Figure entered the circle") and (17b) ("The Figure entered the square") should be fine as descriptions of the scenario in Figures 6.3–6.4, in which a ball enters a circle and both the ball and the circle are located inside a square throughout the event. This, however, is not the case.

(17) a. . . . káa=h-òok | (le=bòola) | ich-il | le=sìirkulo=o'.
CON=PRV-enter(B3SG) | DET=ball | in-REL | DET=circle=D2
'. . . it entered (lit. in) the circle.'

b. #H-òok | (le=bòola) | ich-il | le=kwàadro=o'.[10]
PRV-enter(B3SG) | DET=ball | in-REL | DET=square=D2
'. . .it entered (lit. in) the square.'

My consultants reject (17b) as a description of Figures 6.3 and 6.4, despite the fact that the

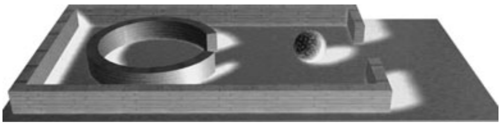


FIGURE 6.3. First frame of ENTER_EXIT 10.

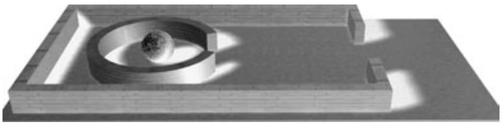


FIGURE 6.4. Last frame of ENTER_EXIT 10.

CoL event takes place inside the square. For (17b) to be true, the ball would have to be located outside the square in the source state of the CoL event and inside in the end state. An event Location interpretation of the Ground phrase in (17b) is unavailable. The Ground phrase invariably encodes a Place function, not a Locative relation. Using the notational conventions of Jackendoff (2002), the semantic composition in (17b) can be represented as in Figure 6.5, either in terms of a Jackendoffian Path semantics (CS I) or in terms of a State-Change semantics (CS II); both analyses are compatible with the facts presented in this section. Double lines indicate the projection of phrases from their heads in the syntactic representation and the determination of ontological types from conceptual functions in CS; the parallelism is intended as a reminder that conceptual functions tend to be encoded by syntactic heads. Dashed lines encircle the domains of the contribution of "Lexical Conceptual Structures" (LCS) as encoded in particular by the verb *òok* "enter" (or "become inside") and the preposition *ich (-il)* "in." Functional categories and the internal structure of the noun phrases are ignored in Figure 6.5. Indices encode the mapping between CS and syntax. The LCSs of *òok* and *ich*(*-il*) are combined through the process of "argument fusion" as discussed in Jackendoff (1990). The PP headed by *ich(-il)* encodes a Place function, which is mapped either into a Path function (CS I) or a Locative state, which in turn maps into a state change (INCH "inchoative") function (CS II), depending on whether *òok* has a Path (CS I) or CoL (CS II) semantics. The evidence presented in particular in the following section suggests that the latter analysis (CS II) is correct.

The absence of Path encoding in Ground phrases sets Yucatec apart from better-studied verb-framed languages such as those previously mentioned, although, as argued in Bohnemeyer et al. (2007), it does not appear to be a rare phenomenon in the languages of the world. For present purposes, the absence of Path specifications in the Ground phrase combined with the requirement of CoL verbs as heads of verbal cores in CoL-denoting clauses sets the stage for the hypothesis, pursued in the following sections, that Path is not encoded in Yucatec and that Motion is systematically cast as CoL in Yucatec semantics.

LOCATION CHANGE WITHOUT FIGURE MOTION

The previous section has shown that Path functions are not encoded *outside* verb roots in Yucatec, and that the verbal core of a Motion event description must be headed by a verb of "inherently directed motion" (Levin, 1993: 263), which aspectual tests identify as a State-Change verb, i.e., a CoL verb. This and the following sections make the case that Path functions are not lexicalized in Yucatec CoL verbs either, and are therefore not expressed in Yucatec—put differently, the case for a consistent framing of Motion as CoL, not T-Motion, in this language. The most direct source of evidence is presented in this section: the applicability of verbal cores and clauses projected from CoL verbs to scenarios that involve CoL, but not T-Motion, of the Figure with respect to the Ground, along the lines of Kita's (1999) work on Japanese *hairu* and *deru*, as discussed in the introduction. The following sections examine additional evidence of a more indirect nature. The impossibility of composing complex Path functions, the underspecification of CoL with respect to Route Paths, and the lack of "Fictive Motion" and

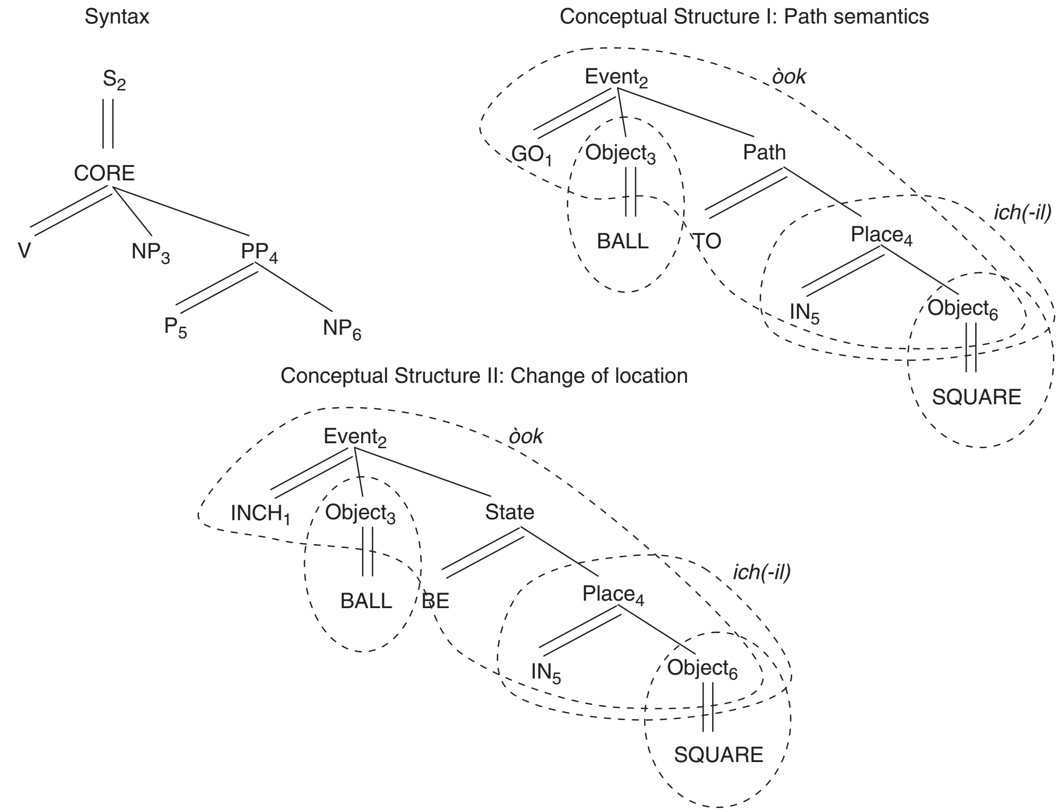


FIGURE 6.5. Semantic composition in (17b).

spatiotemporal metaphors involving Path functions are all readily understood as consequences of the absence of Path encoding.

The data presented in this section were collected with five adult native speakers in 2001, using the "Motion verb stimulus" (MoVerbs) designed and produced by Stephen Levinson (Levinson, 2001). MoVerbs comprises 96 computer-animated video clips featuring a variety of CoL scenarios varied according to the spatial relation between Figure and Ground at the source or target state or in between, the involvement of Figure Motion, and perspective (toward/away from the observer vs. lateral to the observer's viewing axis). Additional data collected with improvised stimuli are reported in Bohnemeyer (1997).

Three types of scenarios are discussed in the following subsections: scenarios in which the Ground moves instead of the Figure ("Ground Motion"; 3.1) and scenarios in which the Figure emerges in or disappears from a configuration with the Ground or, conversely, the Ground emerges in or disappears from a configuration with the Figure (3.2). It is not claimed that such scenes are significantly more natural to Yucatec speakers than they are to English speakers. These scenes are merely used here as analytical tools to probe the semantics of CoL-encoding constructions, since they effectively divorce CoL from T-Motion. As it so happens, the results suggest that Path semantics plays less of a role in such constructions in Yucatec than it does in English.

Ground Motion

Consider Figures 6.6 and 6.7: The enclosure moves such that the ball ends up inside. Out

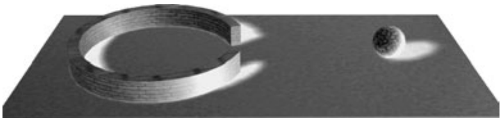


FIGURE 6.6. First frame of ENTER_EXIT 03.

FIGURE 6.7. Last frame of ENTER_EXIT 03.

of context, most speakers consider (18) misleading as a description of this scenario:

(18) #Le=bòola=o′ h-òok te=sìirkulo=o′.
DET=ball=D2 PRV-enter(B3SG) PREP:DET=circle=D2
'The ball, it entered the circle.'

However, unlike its English translation, (18) is not semantically in contradiction with Figures 6.6 and 6.7 for most of my consultants. Example (18) merely invites a strong implicature to the effect that the theme of *òok* "enter," "become inside," the ball, moves. If this implicature is blocked or cancelled in context, application of (18) to Figures 6.6 and 6.7 is fine for most speakers:

(19) H=tàal le=àaro y=iknal le=bòola=o′;
PRV=come(B3SG) DET=ring A3=at DET=ball=D2
le=bòola=o′ h=òok-ih.
DET=ball=D2 PRV=enter-B3SG
'The ring came to the ball; the ball, it entered.'

And even consultants who reject (19) generally accept (20), in which a derived stative form of the verb is used to ascribe the result state of having entered to the ball:

(20) T-u=huts′-ah u=báah=e′,
PRV-A3=approach-CMP(B.3.SG) A3=self=D3
káa=t-u=k′al-ah le=bòola=o′,
CON=PRV-A3=close-CMP(B3SG) DET=ball=D2
káa=h=ts′o′k=e′, le=bòola=o′, òok-a′n, (...)
CON=PRV=end(B3SG)=TOP DET=ball=D2 enter-RES(B3SG)
'[The ring] approached, and it enclosed the ball, and then, the ball, it was entered, (...)'

It appears that (20) is even more widely accepted than (19) in reference to Figures 6.6 and 6.7 because the Path semantics implicature is weaker with the resultative form, as the resultative form gives more prominence to the target state of the ball being inside the enclosure than to the event that brought about that state.

Essentially the same distribution found with *òok* "enter" is found with *na′k* "ascend" in relation to the scenario in Figures 6.8 and 6.9, in which a slope slides under a ball. Most consultants find the description in (21) perfectly acceptable for this scenario:

(21) Le=chan tàabla=o′ h=péek-nah-ih,
DET=DIM plank=D2 PRV=move-CMP-B3SG
káa=h-na′k le=chan kanìika
CON=PRV-ascend(B3SG) DET=DIM marble
y=éetel che′ te′l y=óokol=o′.
A.3=with wood there A3=on=D2
'The little plank, it moved, and the little marble and the tree ascended there on top.'

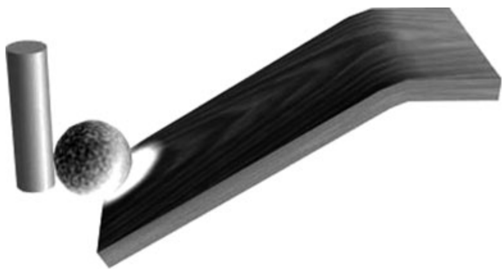


FIGURE 6.8. First frame of FIGURE_GROUND 14.

FIGURE 6.9. Last frame of FIGURE_GROUND 14.

And again, the result state of *na'k* "ascend" is considered even more applicable to the ball:

(22) Le=tàabla=o' káa=h-háarax-nah=e',
DET=plank=D2 CON=PRV-slide-CMP(B3SG)=D3
káa=h-em kàabal. Káa=h-p'áat
CON=PRV-descend low CON=PRV-quit\ACAUS(B3SG)
le=bòola y=óokol na'k-a'n.
DET=ball A.3=on ascend-RES(B3SG)
'The plank, it slid, it went down. The ball ended up on top of it ascended.'

However, not all CoL verbs/scenarios are compatible with Ground Motion. Consider the scenario in Figures 6.10 and 6.11, in which a stick moves to a ball. In this case, the verb *k'uch* "arrive" is completely unacceptable with the ball as theme to all consultants, even if it is stated in context that it is the stick that moves. Even the result state of *k'uch* is considered applicable to the ball by only one out of five speakers. The description is quoted in (23).

(23) Káa=h-bin u=háarax=e'; káa=h-ts'o'k=e',
CON=PRV-go(B3SG) A3-slide=D3 CON=PRV-end(B3SG)=D3
k'uch-a'n le=bòola y=iknal=o'.
arrive-RES(B3SG) DET=ball A3=at=D2
'(The stick) went sliding; [when/and then] that became over, the ball was in the state of having arrived next to it.'

It appears that there is a hierarchy of CoL verb roots in terms of acceptability with Ground Motion; cf. Figure 6.12.[11]

By hypothesis, the verbs in the column on the right in Figure 6.12 are most strongly associated with and those in the left column are least strongly associated with Path semantics. But the explanation for the existence of this hierarchy is not entirely clear. It is of course possible that the roots in the right column in fact lexicalize Path functions. But given that they pattern with the other CoL verbs in all those aspects discussed in the previous section and the sections to follow, such a radical semantic difference would itself call for an explanation that is nowhere in sight at present. In contrast, a hypothetical explanation of Figure 6.12 in line with the CoL analysis can at least be outlined. The three columns of Figure 6.12 differ neatly in terms of the Place function of the Ground: IN (containment) in the left column, ON (support) or ABOVE (superposition) in the middle column, and AT (proximity or contact) in the column on the right (see Table 6.1). Now, at least in English and related languages, it is perfectly natural to linguistically locate a Figure IN, ON, or ABOVE a *moving* Ground (e.g., a moving vehicle); but to do so AT a moving Ground seems impossible. Thus, the car in (24a) may be in Motion or stasis, but (24b) is acceptable only if the car is not moving at the time.

FIGURE 6.10. First frame of FIGURE_GROUND 11.

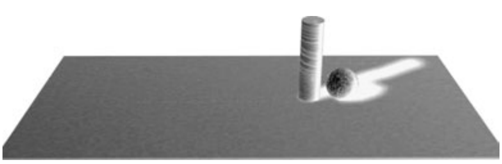


FIGURE 6.11. Last frame of FIGURE_GROUND 11.

(24) a. *Floyd was in the car*
b. *Floyd was at the car*

The generalization seems to be that AT Place functions can be assigned only to static objects. Future research will have to establish whether this generalization holds for Yucatec as well. If it does, that would explain why the verbs in the column on the right in Figure 6.12 are not applicable to Ground Motion scenarios. Indirect confirmation of this hypothesis comes from the fact, reported in the next subsection, that the verbs in the right column are in fact more compatible with events involving the emergence or disappearance of the Figure in or from a configuration with the Ground.

Figure/Ground Emerging/Disappearing

Another test of CoL semantics is CoL coming about as a result of the Figure emerging in or disappearing from a configuration with the Ground. Compatibility with such "beaming" scenarios—just as compatibility with Ground motion—shows that Yucatec CoL descriptions

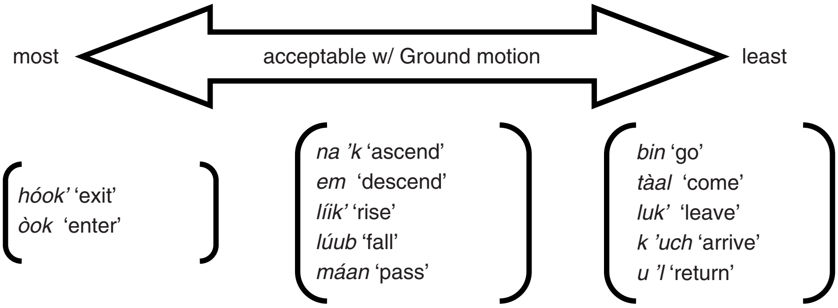


FIGURE 6.12. Acceptability of CoL roots with Ground Motion.

do not entail T-Motion of the Figure. The stimuli employed in the present study instantiate this type of scenario with teleportation of the Figure or Ground, as in science fiction movies. Examples (25a)—(25b) feature *òok* "enter" in descriptions of a scene in which a ball "beams" into an enclosure; cf. Figures 6.13 and 6.14.

(25) a. Le=chan bòola=o', káa=h-sáat=e',
DET=DIM ball=D2 CON=PRV-lose\ACAUS(B3SG)=TOP
káa=h-chíik-pah ka'=téen=e', ich le=chan àaro
CON=PRV-appear-SPONT(B3SG) two=CL.times=TOP in DET=DIM ring
yàan=o'; h=òok chíik-pah-al.
EXIST(B3SG)=D2 PRV=enter(B3SG) appear-SPONT-INC
'The little ball, [when/and then] it vanished, [when/and then] it appeared again, it was in the ring ring; it entered emerging.'

b. Káa=h-sáat=e',
CON=PRV=lose/ACAUS(B3SG)=TOP
káa=h-chíik-pah=e',
CON=PRV=appear-SPONT(B3SG)=TOP
ich-il le=sìirkulo yàan=i'; òok-a'n.
in-REL DET=circle EXIST(B3SG)=D4 enter-RES(B3SG)
'[When/and then] [the ball] disappears; [when/and then] it appears [again], it's inside the circle; it has entered.'

FIGURE 6.13. First frame of MoVerbs ENTER_EXIT 07.

FIGURE 6.14. Last frame of MoVerbs ENTER_EXIT 07.

The applicability of CoL verbs under teleportation Motion of the Figure is similar to that under Ground Motion as in (25a). And likewise, just as illustrated with Ground Motion scenarios, acceptability of uses of CoL verbs in reference to teleportation scenarios generally increases when some form of the verb is chosen that focuses on the result state of the CoL event, such as the resultative derivation in *–a'n* in (25b).

Applicability of CoL verbs to teleportation events seems to vary across verbs along a scale similar to the one for Ground Motion depicted in Figure 6.12. Only scenarios of teleportation into or out of some kind of containment configuration elicit dynamic CoL descriptions with the Figure as theme in perfective aspect; the remaining types of scenes are merely amenable to descriptions featuring resultative forms of the CoL verbs with the Figure as the sole argument. Example (26) illustrates this type of response with *máan* "pass" in reference to the result state of an event of "beaming" across a dyke, as depicted in Figures 6.15 and 6.16:

(26) Káa=h-sáat=e',
CON=PRV-lose/ACAUS(B3SG)=TOP
káa=h-ka'=chíik-pah=e' tu=láahun-tséel
CON=PRV-REP=appear-SPONT(B3SG)=TOP PREP:A3=other:one-side
le=pak' màaha'n yàan=o'.
DET=wall pass:RES(B3SG) EXIST(B3SG)=D2
'[When/and then] [the ball] vanished, [when/and then] it reappeared, it had passed [to] the other side of the wall.'

Unlike in Ground Motion scenarios, the verbs in the right column of Figure 6.12 are acceptable with Figure teleportation, as shown in (27), a description of the scenario

FIGURE 6.15. First frame of MoVerbs PATHS 06.

FIGURE 6.16. Last frame of MoVerbs PATHS 06.

FIGURE 6.17. First frame of MoVerbs PATHS 11.

in Figures 6.17 and 6.18 (a ball "beaming" from a tree to a hill).

(27) Káa=h-sáat=e′,
CON=PRV=lose/ACAUS(B3SG)=TOP
káa=h-chíik-pah=e′
CON=PRV= appear-SPONT(B3SG)=TOP
sáam k′uch-uk y=iknal le=mùul=o′.
REC arrive-SUBJ(B3SG) A3=at DET=hill=D2
'[When/and then] [the ball] vanished, [when/and then] it appeared, it had already/just arrived at the hill.'

FIGURE 6.18. Last frame of MoVerbs PATHS 11.

In (27), the verb *k′uch* "arrive" appears with a recent past marker, a construction sometime used as a pragmatic alternative to the resultative and various other constructions denoting post-state reference (Bohnemeyer, 2002: 328–342). Another speaker described the same clip using a resultative form of *tàal* "come":

(28) Káa=h-sáat t-u=chùun le=che′=o′,
CON=PRV-lose/ACAUS(B3SG) PREP-A3=begin/ATP DET=wood=D2
káa=h-tàal chíik-pah-al,
CON=PRV-come(B3SG) appear-SPONT-INC
náats′ t-inw=iknal tàaha′n.
near PREP-A1SG=at come:RES(B3SG)
'[When/and then] [the ball] vanished at the trunk of the tree, [when/and then] it came appearing, it was come close to me.'

Three out of five speakers accept descriptions such as (27) or (28) in reference to the "beaming" scenario in Figures 6.17 and 6.18. This supports the hypothesis that the blocking of the verbs in the right column of Figure 6.12 with Ground Motion is due to AT-Place functions operating on static objects only. Under this hypothesis, the same verbs should be acceptable in reference to emerging/disappearing Figures, and (27)–(28) confirm this. Unfortunately, the scenario in Figures 6.17 and 6.18 is the only one of this kind in the set; more evidence is clearly needed here.

A spatial configuration may also change due to the Ground emerging or disappearing. There are relatively natural instances of this (at least compared to scenarios of the Figure emerging or disappearing); e.g., if an enclosure is built around some object, can it be said that

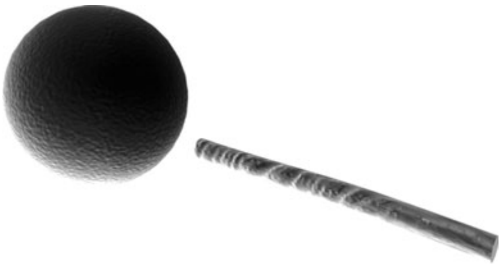


FIGURE 6.19. First frame of MoVerbs FIGURE_GROUND 20.

FIGURE 6.20. Last frame of MoVerbs FIGURE_GROUND 20.

the object has entered the enclosure? And does the object exit when the enclosure is torn down? This has been tested only with ENTER, EXIT, and ASCEND scenarios (and, once again, with animations of teleportation). The results suggest a strong preference for result state reference with inactive CoL verbs. Example (29) shows a description of a stimulus clip in which a stick pierces a ball by the latter "beaming" onto it, depicted in Figures 6.19 and 6.20. The description uses the resultative form of *òok* "enter."

(29) Káa=h-chíik-pah le=bòola=o',
CON=PRV-appear-SPONT(B3SG) DET=ball=D2
òok-a'n che' ti'.
enter-RES(B3SG) wood PREP(B3SG)
'[When/and then] the ball appeared, [a] stick had entered it.'

Summary

CoL verbs are used in a wide range of scenarios that do not involve Motion of the Figure/theme, namely under Ground Motion and with the Figure or the Ground emerging or disappearing. Generally, consultants are much more likely to produce or accept CoL verb constructions under lack of Figure Motion in case the context makes it clear that the Figure does not move. This suggests that the CoL verbs do not entail Translational Motion of the Figure, but carry generalized conversational implicatures to its effect. A plausible source for such implicatures would be Grice's (1975) second maxim of Quantity, "Do not make your contribution more informative than is required," or Levinson's (2000) corresponding I(nformativeness) Heuristic ("What is expressed simply is stereotypically exemplified"). Furthermore, aspectual reference has an impact on acceptability of CoL verb constructions under lack of Figure Motion. Perfect or resultative predications, focusing on the result state of the CoL event instead of the event itself, are accepted across the board (with the exception of verbs encoding AT-Place functions, as these arguably require static Ground objects). In contrast, the acceptance of perfective-aspect clauses in reference to the CoL events themselves is always equal to or less than that of result state constructions.

INDIRECT EVIDENCE FOR THE ABSENCE OF PATH LEXICALIZATION

The previous section presented direct evidence for the framing of Motion as CoL in Yucatec: the semantic compatibility of the same clauses used to describe Motion of a Figure with scenarios in which a Figure undergoes CoL through Ground Motion or the emergence/disappearance of the Figure or the Ground. In the present section, I examine additional indirect evidence, in the form of consequences arising from the absence of the lexicalization of Path functions. In so doing, I address the two linguistic arguments Jackendoff (1990) advances against a representation of Motion in terms of CoL in CS: the difficulty of encoding Motion with respect to Route Paths in this way and the occurrence of Path functions in what Talmy (1996, 2000a) has called "Fictive Motion" metaphors. I add an argument of my own: the difficulty of encoding events involving complex Path functions as CoL. I show that none of these arguments applies to Yucatec in a convincing fashion. Furthermore, I discuss the absence of Path metaphors for temporal relations and interference effects in Yucatecan L2-Spanish that may be viewed as reflexes of a lack of Path encoding in Yucatec CS.

The Treatment of Routes

The framing of Motion as CoL leads, probably inexorably, to a certain amount of loss of information in the case of Route Grounds, which define neither the beginning nor the endpoint of the Path, but some point in between. Conceptually, CoL is composed of a Locative relation plus information about a particular part of the event during which this relation applies. Routes cannot be reduced to Locative relations without "oversimplification." My walking *across* the road is characterized only inadequately by saying that at some point during the "nucleus" of the event, I am *on* the road (cf. also Jackendoff, 1983: 174; 1990: 93–94). Thus in a language in which Motion is construed purely in terms of CoL we should expect a drastic amount of underspecification in the encoding of CoL VIA Route Grounds. And this is exactly what is found in Yucatec, in which a single verb, *máan* "pass," is used to encode all CoL events involving Route Grounds. Consider (30):

(30) Túun bin u=balak'=e',
PROG:A3 go A3=roll=D3
káa=h-máan tu=bèel le=trèen=o',
CON=PRV-pass(B3SG) PREP:A3=way:REL DET=train=D2
káa=h-òok ich le=che'-o'b=o' . . .
CON=PRV-enter(B3SG) in DET=wood-PL=D2
'[The ball] was going rolling, [and then] it passed across/along/on the railroad tracks, and it entered the group of trees. . .'

Example (30) was originally elicited as a description of a scene in which a ball rolls across railroad tracks. The clause *káah máan tu bèel le trèeno'* "it passed across/along/on the railroad tracks" was also elicited in response to a scene in which a ball rolled along a set of railroad tracks, and my consultants confirm that the entire description in (30) can be understood to the effect that the ball crosses the tracks, moves along them, or follows the tracks rolling on them. However, the drastic vagueness of (30) is to some extent a function of the one-dimensional structure of railroad tracks. Both spatial prepositions of Yucatec, the generic *ti'* and *ich(il)* "in," and all the relational nouns listed in Table 6.3 are compatible with *máan* "pass." Enriched through application of Gricean implicatures, these combinations accurately represent most scenarios.

There are two residual questions. First, does *máan* "pass" itself encode a Path function? Because it is compatible with scenarios in which a Figure "beams" through/over/across a barrier (cf. 26), I tentatively conclude that this is not the case. And second, what might a plausible CoL semantics for *máan* look like? The verb is obviously not amenable to an ordinary CoL decomposition in terms of a Locative function that characterizes either the source or the target state. An alternative might be an underspecified Place function (as *máan* is compatible with any Place function encoded by the Ground phrase) combined with change from the state of this Place not having been passed by the Figure to it having been passed, along the lines of (31):

(31) $[_{\text{Event}}\ \text{INCH}\ ([_{\text{Thing}}\ ], [_{\text{State}}\ \text{BE}\ ([\ ],[_{\text{Place}}\ \text{PAST}\ ([_{\text{Place}}\ ([_{\text{Thing}}\ ])])])])]$

The Place function immediately projected from the Ground is left unspecified in (31). The state of having passed this ground is represented in terms of a secondary place function PAST. Of course, PAST must in turn derive its meaning from the mental representation of a Path. But as I will argue, such a representation may be afforded by the SpS system instead of CS. The "Lexical Conceptual Structure" of *máan* would tap into this SpS in a manner similar to how, for example, Manner-of-Motion verbs and shape expressions derive part of their meanings via SpS encoding (cf. Jackendoff, 1996; 2002: 345–350).

Complex Path Functions

The composition of multi-Ground Paths represents another challenge to framing Motion in terms of CoL. Consider (32):

(32) *The supporters went from the meet-up to the rally*

A State-Change analysis of (32) would have to rely on both source and target state of the CoL event being characterized by Locative functions—the state of being at the meet-up and the state of being at the rally. The format of the INCH function does not afford this; it allows for specification of a single state only—usually the target state. Tellingly, English often relies on Motion metaphors to express complex state changes:

(33) a. *The lights went/changed from green to red*
b. *Floyd's mood went/changed from exuberant to gloomy in a flash*

As shown in Bohnemeyer (2003b, 2007) and Bohnemeyer et al. (2007), Yucatec clauses do not express CoL with respect to more than one Ground. Complex CoL events are broken down into sequences of single-Ground CoL events each of which is encoded by a separate clause. For illustration, (34) is a description of a video clip in which a ball rolls from a tree past a dip to a hill (the setting is the same as in Figures 6.17 and 6.18, but the ball rolls through the landscape instead of "beaming"):

(34)

H-luk'	y=iknal	le=che'=o',	
PRV-leave(B3SG)	A3=at	DET =wood=D2	
káa=h-tàal	u=ba'+pàach-t-ik	le=àaktúun=o',	
CON=PRV-come(B3SG)	A3=?+back-APP-INC(B3SG)	DET=hole=D2	
káa=h-k'uch	he'l-el	y=iknal	le=búut'un=o'.
CON=PRV-arrive(B3SG)	rest-INC	A3=at	DET=hill=D2

'[The ball] left at the tree, [and then] came going around (lit. surrounding) the dip, [and then] it arrived to rest at the hill.'

The restriction to one CoL Ground *per verbal core* is a consequence of the fact that Ground phrases denote Place functions, which are mapped into Locative functions by the verb, as per the semantic composition illustrated in Figure 6.5. For multiple Ground phrases to be licensed in a single core, the verb would have to lexicalize multiple Locative functions and assign these to the different Grounds. Such verbs are unattested in Yucatec or any other language. The restriction to one CoL Ground *per clause* is a consequence of the restriction to one Ground per core and the lack of constructions of an appropriate kind that combine multiple CoL-denoting cores into clauses. Examples of such constructions are "serial verb" or "multi-verb" constructions in Ewe (Kwa/Gbe; Ghana and Togo) and Lao (Tai-Kadai; Laos), as discussed in Bohnemeyer et al. (2007).[12] Thus, in line with the hypothesis of a systematic construal of Motion as CoL in Yucatec, there is no evidence of semantic composition of complex Path functions in the language.

Fictive Motion Metaphors

One of the arguments Jackendoff (1983, 1990) gives in defense of Path semantics is the occurrence of Path relations outside the Motion domain, for instance in expressions of extent (cf. 35a), orientation (cf. 35b), or as "reference paths" ["Access Paths" in Talmy's (2000a: 136–137) parlance] in Locative predications (35c):

(35) a. *The highway extends from Denver to Indianapolis.*
b. *The house faces away from the mountains.*
c. *The firehouse is across the street from the library.* (Jackendoff, 1983: 167–172)

If Path functions occur independently of CoL, they should be primitives of CS, and this status should extend to the event functions that occur uniquely with them, i.e., event functions of T-Motion (encoded at CS by the primitive GO). The event functions in cases such as (35) are extensions of GO along the lines of Talmy's (1996, 2000a) "Fictive Motion."

The following observations are based on the elicitation of Yucatec renditions for instances of all types of English Fictive Motion metaphors[13] discussed in Talmy (2000a: 105–138) with five adult native speakers. The CoL verbs of Table 6.1, the prepositions *ti'* (generic) and *ich(il)* "in," the relational nouns listed in Table 6.3, and the constructions that combine these expressions all can be used metaphorically. But such metaphors are subject to the constraints on framing Motion as CoL discussed above. Thus, CoL verbs can be used in descriptions of “coextension paths” (Talmy, 2000a: 138) such as (35a); however, due to the restriction to one CoL Ground per clause, these have to be broken down into sequences of clauses denoting “Fictive CoL” with respect to a single Ground. Example (36) is a rendition of “This road here goes from Señor via Tixcacal to Yaxley”:

(36) Le=bèeh he′l=a′, k-u=hóok′-ol Señor,
DET=way PRSV=D1 IMPF-A3=exit-INC Señor
k-u=ts′o′k-ol=e′, k-u=máan Tixcacal,
IMPF-A3=end-INC=TOP IMPF-A3=pass(INC) Tixcacal
k-u=ts′o′k-ol=e′, k-u=k′uch-ul Yaxley,
IMPF-A3=end-INC=TOP IMPF-A3=arrive-INC Yaxley
‘This road here, it exits Señor; then [lit. that having ended] it passes [through]
Tixcacal; then [lit. that having ended] it arrives [in] Yaxley.’

There is no evidence that metaphorical uses of CoL expressions as in (36) involve Path semantics. Thus, they are more properly considered instances of “Fictive CoL” rather than Fictive Motion.[14]

Among the various types of metaphors distinguished by Talmy, only the “Co-extension Paths” and “Frame Motion” (virtual Motion effects; e.g., trees seen as passing by a car) types have equivalents in Yucatec that employ CoL expressions. Meanings corresponding to those of the remaining types are expressed nonmetaphorically in Yucatec. Consider, first, the family of “Orientation Path” metaphors (Talmy, 2000a: 106–111), all of which involve the Direction Paths functions TOWARD and AWAY-FROM. Because these do not entail CoL, they are not morphologically encoded in Yucatec. The same Ground phrases that, depending on the verb they combine with, may have AT, FROM, TO, or VIA readings are also compatible with TOWARD and AWAY-FROM readings. But because there are no verbs that lexicalize Direction, Direction specifications are never unambiguous, except in combination with the indexical verbs *bin* “go,” *tàal* “come,” and *u′l* “return (to deictic center).” These verbs entail CoL with respect to the deictic center or an anaphorically traced Place; so when they occur with Ground phrases, these are unambiguously interpreted as Directional adjuncts. However, none of these verbs can be used in Orientation Path metaphors such as (35b). Although the reason is not entirely clear, the finding meshes with the fact that all verbs in Talmy’s (2000a: 108–111) examples of the various types of orientation Paths are either stative (as in 35b) or involve change of Direction (as in *I looked down into the well*). “Demonstrative Path” (Talmy, 2000a: 109), which describe a person or object pointing in a certain direction, are expressed using stative verbs such *e′s* “show” or *chîikult* and *túuchul*, both “sign,” “signify.” The “Directional Goal” (corresponding to the TOWARD Ground in English Directional expressions) is expressed by the object of these verbs. Thus, (37) is a Yucatec equivalent of “The broom is pointing toward don Modesto’s house”:

(37) Uy=òok le=mìis=o′, k-uy=e′s-ik
A3=foot DET=broom=D2 IMPF-A3=show-INC(B3SG)
u=nah-il don Modesto.
A3=house-REL don Modesto
‘The broom stick (lit. the leg of the broom) is showing don Modesto’s house.’

It is impossible to semantically encode Directions AWAY-FROM a Ground in this fashion. Example (38) is a typical response to persistent attempts at eliciting a rendition of “The broom is pointing away from the bucket”:

(38) Le=mìis=a′, y=áanal+tu′x súut-ul uy=òok, ma′
DET=broom=D1 A3=other+where turn\ACAUS-INC A3=foot NEG
t-u=toh-il le=kùubo=o′. Pero u=mìis-il=e′,
PREP-A3=straight-REL DET=bucket=D2 but A3=broom-REL=TOP
ti′=yàan náats iknal le=kùubo=o′.
there=EXIST(B3SG) near at DET=bucket=D2
‘This broom, its stick is turned elsewhere, not in the line of the bucket. But its bristles (lit. its broom), they are close to the bucket.’

The orientation of an object with a designated front part is often described in English with a “Prospect Path” metaphor (Talmy, 2000a: 108) such as (35b). Yucatec has nonmetaphorical expressions for such configurations. An example is the relational noun *aktáan* “front-to-front” in (39), which indicates that Figure and Ground face each other:

(39) U=nah-il Pablo=e′, ti′=yàan t-u=láak
A3=house-REL Pablo=TOP there=EXIST PREP-A3=other
hun-p′éel tséel le=bèeh=o′; ak+táan ti′ u=nah-il Pedro.
one-CL.IN side DET=way=D2 ?+front PREP A3=house-REL Pedro
‘Pablo’s house, it is on the other side of the road, front to front with Pedro’s house.’

Example (39) also illustrates how Yucatec speakers convey the meanings expressed by "Access Path" metaphors (Talmy, 2000a: 136) such as *across the street from the library* in (35c) in English. Again, the Yucatec expression is non-metaphorical ("on the other side of the road").

Finally, consider "Line of Sight" and "Sensory Path" metaphors, which describe perception as fictive motion (e.g., *look into the well*; *see the enemy from the hill*; Talmy, 2000a: 110–111, 115–116). Most Yucatec perception verbs link the stimulus of perception to the undergoer argument, which may remain implicit with this class of verbs. In addition or alternatively, some combine with Ground phrases that denote the Place on which perception is focused. How, then, do we convey the idea of looking *through* a window or the like? Example (40) shows one solution: The window and the stimulus seen through it are referred to in different clauses with different perception verbs; the spatial relation between them is left to inferences.

(40) Káa=t-a=pakat-ah te=béentanah=o',
CON=PRV-A2=look.at-CMP(B3SG) PREP:DET=window=D2
káa=t-aw=il-ah ba'x yàan ich le=nah=o'.
CON=PRV-A2=see-CMP(B3SG) what EXIST(B3SG) in DET=house=D2
'[When/and then] you looked (lit. at it) at the window, [when/and then] you saw what was in the house.'

The findings presented here generalize to all known types of Fictive Motion metaphors. These meanings are expressed either as "Fictive CoL" or nonfiguratively in Yucatec.

Spatiotemporal Metaphors

Many natural languages use spatiotemporal metaphors to express ordering relations between time intervals. Temporal connectives such as *after*, *before*, and *while* often etymologically derive from metaphors involving Path semantics, and have been argued to be always based on such metaphors conceptually (e.g., Clark, 1973; Fillmore, 1971; Miller & Johnson-Laird, 1976: 462–464; Traugott, 1978). These are "localist" analyses, i.e., analyses that accord a prominent role to spatial relations as models in the conceptualization of nonspatial domains. The domain mapping in spatiotemporal metaphors is made possible by an isomorphism between the conceptual structures of time and Paths.[15] Bohnemeyer (1998, 2000, 2002, 2003a) has shown that Yucatec lacks expressions of temporal ordering relations, with a few systematic exceptions such as deictic calendrical adverbs ("yesterday," "tomorrow"), adverbs meaning "now" and "formerly," and idioms used as generic temporal anaphors ("when"). There are no connectives that encode a specific order between two time intervals such as *after*, *before*, *while*, *during*, *since*, or *until*. Temporal ordering in discourse is conveyed through the encoding of fine aspectual and modal distinctions in combination with Gricean implicatures. Consider, for example, the aspectual verb *ts'o'k* "end," used in (often reduced) topicalized clauses as a kind of aspectual connective (e.g., 20, 23, 36, and 41). Semantically, the construction [S_1 [. . .ts'o'k. . .]$_{Topic}$ S_2] encodes sequential order (nonoverlap) between the events referred to by S_1 and S_2. Which of the two events happens first is inferred from the order of clauses; antiiconic ordering, as is possible and quite natural with *after* (*Sally finished her report after talking to Floyd*), cannot be expressed in this construction.

Several of the prepositions and relational nouns previously discussed can in fact be used in spatiotemporal metaphors. However, these metaphors do not represent two-place ordering relations, and that seems to be a direct consequence of the fact that the source expressions do not lexicalize Locative or Path relations. For example, the relational nouns *táan* "front" and *pàach* "back" can be used to refer to the first or last Place in a sequence of events (as well as occurring in compound verb stems with the meaning "do something prematurely/belatedly"). This is illustrated for *táan* in (41)–(42):

(41) Yáax táan-il=e', Pedro h-síih-ih.
first front-REL=TOP Pedro PRV-be.born-B3SG
Káa=h-ts'o'k=e' káa=h-síih Pablo.
CON=PRV-end=TOP CON=PRV-be.born(B3SG) Pablo
'First, Pedro was born. Then (lit. it having ended), Pablo was born.'

(42) Pedro=e', h-síih táan-il ti' Pablo;
Pedro=TOP PRV-be.born(B3SG) front-REL PREP Pablo
Pablo=e', h-síih táan-il ti' José.
Pablo=TOP RV-be.born(B3SG) front-REL PREP José
'Pedro, he was born first with respect to Pablo; Pablo, he was born first with respect to José.'

Táan(il ti') "first (with respect to)" cannot take a verbal core or clause as a complement. Thus, the interpretation of what it is that happened to Pedro before Pablo in the first clause of (42) has to come from the verb [*sîih* "be born" in (42)]. This is in direct parallel to the role of a Ground phrase headed by *táan(il ti')* in the semantic composition of Locative or CoL descriptions. Just as this Ground phrase describes a Place whose role in the event is determined by the verb instead of a Locative or Path function, so the PP in the temporal use describes a metaphorical Place in a sequence of events whose interpretation is determined by the verbal core instead of a temporal ordering relation. For one more illustration, consider the preposition *ich* "in." *Ich* is used with both duration ("for X time") and time span ("in X time") adverbials, as well as expressions of temporal distance as in (43)–(44). In (43), distance is projected into the past of reference time, due to the fact that the verbal core appears in the bare subjunctive, where in (44), the *ich* phrase is understood to measure distance with respect to an event in the future of reference time, due to the presence of the irrealis subordinator *kéen* (see Bohnemeyer, 2002: 411–413, 421–426 for discussion). *Ich* remains neutral with respect to the temporal relation, just as it does not distinguish Locative or Path functions in spatial usage.

(43) Pwes to'n =e', ich ts'e'ts'ek k'ìin hóok'-ok-o'n.
well us=TOP in a.few sun exit-SUBJ-B1PL
'Well, as for us, it was a few days ago that we left.'

(44) Pwes to'n =e', ich ts'e'ts'ek k'ìin keen hóok'-ok-o'n.
well us=TOP in a.few sun SR.IRR exit-SUBJ-B1PL
'Well, as for us, it is in a few days that we will leave.'

The absence of Locative/Path distinctions in the source expressions seems to preclude spatiotemporal metaphors in Yucatec from picking up temporal ordering relations. This supports localist assumptions about Motion and Path as the conceptual basis of expressions of temporal relations, albeit in an unexpected fashion, as localists might not expect expressions of Motion and Path to be language specific to the extent argued for here. Discussion of the point will be resumed later.

Path in L2-Spanish

If Path functions are universal primitives of CS, it follows that they are primitives in the CS of Yucatec speakers as much as they are primitives in the CS of English speakers. If Yucatec speakers entertain CS representations of Path functions, there is no reason to expect that learning the meanings of Path expressions in a contact language should pose a particular problem for them, even if their native language does not express Path functions. Lehmann (1992) quotes anecdotal evidence indicating that this prediction might fail. The second-language Spanish utterances in (45a)–(48a) were produced by speakers whose L1 is Yucatec. In contrast to L1-Spanish usage (given in the b examples), the Ground phrases in these sentences are interpreted as Place denoting, suggesting straightforward calquing from Yucatec.

(45) a. ¿Donde vienes?
L2SPA where come:PRS:2SG
'Where do you come?' [intended: 'where from?']
b. ¿De donde vienes?
L1SPA from where come:PRS:2SG
'Where do you come from?'

(46) a. El ratón salió en su agujero.
L2SPA the rat exit:PAST:3SG in its hole
'The rat exited in its hole.' [intended: 'from its hole']
b. El ratón salió de su agujero.
L1SPA the rat exit:PAST:3SG from its hole
'The rat exited from its hole.'

(47) a. El ratón pasó en su agujero.
L2SPA the rat pass:PAST:3SG in its hole
'The rat passed in its hole.' [intended: 'through its hole']
b. El ratón pasó por su agujero.
L1SPA the rat pass:PAST:3SG via its hole
'The rat passed through its hole.'

(48) a. Saqué el venado sobre el camino.
L2SPA sack:PAST:1SG the deer on DEF way
'I took the deer on the road.' [intended: 'from the road']
b. Saqué el venado del camino.
L1SPA sack:PAST:1SG the deer from:DEF way
'I took the deer from the road.' (Lehmann 1992: 626)

A contrastive quantitative study is needed to assess how widespread such interference phenomena are. If they turn out to be representative of learner varieties among Yucatec native speakers, this would support the hypothesis that Yucatec speakers neither map Path functions from CS into syntax nor do they encode them at CS.[16] However, this support would still be quite weak, as it rests on the unproven assumption that language-specificity at CS may foster L1-transfer. There is currently no empirical evidence for or against this assumption that I am aware of (although it is certainly clear that L1-transfer occurs independently of variation at CS); there simply has not been much research into the language-specificity of CS to date.

Summary

Corroborating evidence against a Path semantics for Yucatec Motion descriptions comes, first, from the semantic underspecification of CoL involving Route Grounds (see the section on The Treatment of Routes). All events of this type are described with the verb *máan* "pass," regardless of whether they involve, from an English perspective, Motion past, along, across, over, under, or through a Ground object. The chunking of complex Motion events into sequences of single-Ground CoL events, each encoded by a separate clause, replaces the composition of complex Path functions, which is unavailable under the framing of Motion as CoL (see the section on Complex Path Functions). Instead of "Fictive Motion" metaphors, which extend Path functions to non-Motion spatial domains, Yucatec employs a limited amount of "Fictive CoL" metaphors, which are subject to the same constraints as all CoL expressions, and otherwise uses nonmetaphoric expressions of these meanings. The evidence from descriptions involving Route Grounds, multi-Ground change, and metaphoric uses of CoL descriptions shows that the arguments for a Path semantics in English previously introduced do not apply to Yucatec. In addition, temporal ordering relations, which have been hypothesized to be conceptualized as metaphorical extensions of Path functions, are largely not encoded at all; the semantics of temporal metaphors that employ spatial prepositions or relational nouns are constrained by the Place functions denoted by their sources.

THE LANGUAGE-SPECIFICITY OF PATH FUNCTIONS AT CS

Let us now consider the implications of the evidence previously assembled for the question of language-specificity in Conceptual Structure (CS). Jackendoff (1992: Ch. 2 and 3; 2002: 334–339) has endorsed the view that the bulk of "lexical concepts"—more or less, word meanings—must be learned, but can be decomposed into (or, from the learner's perspective, built up from) conceptual primitives, a core set of which is innate. Following common practice, I assume that innate concepts are universal, whereas acquired concepts may (but need not) vary with language and culture. Which concepts are innate and which are acquired is an empirical question. Answers to this question can be provided by developmental psychology and (directly or indirectly) by the study of semantic acquisition and cross-linguistic variation in semantics (or "semantic typology"; cf. Bohnemeyer et al., 2007; Levinson, Meira, & The Language and Cognition Group, 2003). The last-mentioned angle is, of course, the one from which this study aims to make a contribution.

The question is, then, whether the T-Motion event function GO and the ontological type of Path functions are innate and therefore universal primitives of CS. The relevance of this question derives from the "Thematic Relations Hypothesis" (TRH), which proposes (following Gruber, 1965) an organization of CS in terms of different "semantic fields." Each field applies a subset of the same inherently domain-neutral and thus highly abstract conceptual functions and ontological types (Jackendoff, 1983: Ch. 10; 1992: Ch. 2 and 3; 2002: 356–373).[17] I understand Jackendoff's hypothesis to be that these abstract functions and types are unlearnable—they are a part of the innate organization of CS itself. Jackendoff has always maintained that the Path type and the function GO are among the domain-neutral categories. The spatial senses of Motion and Path expressions are generated by applying these abstract functions to the spatial field; other applications are found, for example, in the field of possession, where donors/givers are assigned the Path function FROM and recipients the Path function TO.[18]

The evidence previously presented suggests that Path semantics is not encoded in Yucatec. To be more precise, it suggests that T-Motion and Path functions are not mapped into syntactic representations—that they are neither lexicalized nor grammaticalized. This result does not, however, directly bear on the question of the language-specificity of CS itself. If both T-Motion and Path functions as well as State-Change functions are part of the abstract innate core of CS, then speakers of all languages have the same conceptual resources at their disposal, but English speakers use the Path system to linguistically describe Motion events, whereas Yucatec speakers achieve the same relying on the State-Change system. This outcome is prima facie an eminently reasonable one, as CS is assumed to not only encode linguistic meaning, but at the same time support reasoning—and there is at present no evidence that Yucatec speakers reason about Motion events in any way other than how English speakers reason about them.

There are, however, several sources of indirect evidence that can be brought to bear on the question of the accessibility of Path-semantic functions in Yucatec. First, if CS encodes both conceptual and semantic representations—as Jackendoff argues—then the Yucatec speakers who produced the descriptions of the non-Figure-Motion scenarios previously discussed must have mentally represented these events in terms of CoL. Had they "thought" about the events, *for the purposes of linguistic encoding*, in Path-semantic terms, and stored these CS representations in memory, their descriptions would have been truth-conditionally incompatible with the scenes in question. However, we cannot conclude from this observation that Yucatecans represent Motion as CoL in CS for the purposes of linguistic encoding *outside this task*. This caveat carries some weight because of the observation that Ground Motion and teleportation scenarios seem just as "unnatural" to Yucatecans as to English speakers.

Independent evidence against the availability of Path-semantic functions in Yucatec comes from the L2-Spanish data previously presented. These indicate that Yucatec native speakers transfer the Path-neutral semantics of Yucatec Ground phrases to Spanish. If Path functions were readily available in the CS of Yucatec native speakers, we would expect the Spanish Path prepositions to be able to pick them up easily. Future research will have to assess to what extent these anecdotal data are representative of learner varieties among Yucatec L1 speakers. To the extent that they are, the support for innate Path-semantic primitives is beginning to look somewhat thin.

On the other side of the equation, Jackendoff (1990: 93–94) argued that T-Motion and Path should be primitives of CS in view of experimental evidence suggesting that they are primitives of spatial cognition. But this argument seems to rely on the original version of the "Conceptual Structure Hypothesis":

> There is a single level of mental representation, *conceptual structure*, at which linguistic, sensory, and motor information are compatible. (Jackendoff, 1983: 17)

In Jackendoff (1987), however, this single level was complemented by a second, independent representational system, SpS. SpS encodes object geometry as axial structure and spatial relationships across objects in a way that is neutral regarding sensory modality. It is an iconic and "image-schematic," but not "imagistic," representation. Jackendoff (2002: 347) characterizes the division of labor between CS and SpS as follows:

> The work of understanding the conceptualized world is divided between CS and SpS . . . Judgments and inferences having to do with predicate-argument relations, category membership, the type-token distinction, quantification, and so forth can be formulated only in terms of CS. Judgments and inferences having to do with exact shapes, locations, and forces can be formulated only in terms of SpS. On the other hand, there is overlap between the two levels, in that the notion of physical objects, part-whole relationships, locations, force, and causation have reflexes in both systems.

It is perfectly evident that SpS *must* encode Motion, as well as the Locations of any Ground objects with respect to which the Path is conceptualized. It thus seems that SpS is sufficient to fully support nonlinguistic reasoning about Motion events. I am not aware of any evidence that would motivate a duplication of the information at CS, except for the sole purpose of linguistic encoding. And that motivation does not appear to hold for Yucatec. This in turn calls into question the universality and innateness of Path semantics from an evolutionary perspective. Why would a particular subsystem of CS become encoded in the human genome, if it exists for the sole purpose of representing certain types of linguistic meanings, yet these meanings are not even expressed in all languages?

At this point, the question becomes relevant of just how exotic or widespread the systematic framing of Motion as CoL, Yucatec-style, is in the languages of the world. Two critical typological boundary conditions for representing Motion as CoL seem to be strict Path-neutrality of Ground phrases and the absence of "multiverb" constructions that permit the semantic composition of complex Path functions. Among the 18 genetically and typologically diverse languages surveyed in Bohnemeyer et al. (2007), these conditions are simultaneously met by seven languages: the Mayan languages Tzeltal and Yucatec and the Oto-Manguean language Zoogocho Zapotec (all spoken in Mexico); the Western Oceanic languages Kilivila and Saliba, spoken in Papua New Guinea; the West-Papuan language Tidore of Indonesia; and the East-Papuan language Yélî Dnye of Papua New Guinea. To this I would tentatively add many if not most members of the Bantu language family, which happened to not be represented in the sample of Bohnemeyer et al. (2007). Although none of the other languages has been examined for the phenomena previously discussed to the extent Yucatec has been, I see no reason at present to assume that Yucatec is an isolated case, or that the linguistic framing of Motion as CoL is restricted to a particular family or group of languages.

If it can be confirmed that there are languages all over the world (at least outside Eurasia and Australia) that systematically encode Motion as CoL, and if it can be confirmed that reasoning about Motion events is afforded by SpS alone, then the case for the innateness of Path semantics collapses. What are the implications of this hypothetical outcome? Jackendoff makes a convincing case for the spatial manifestations of Path semantics being just special instances of more abstract conceptual functions built into the very core of CS. We might have to seriously consider, then, the possibility that aspects of the very core of CS may be language specific. This extent of language-specificity of CS would be made possible by a much greater degree of universality and language-independence in the SpS system. This in turn would call into question the position, advocated in Jackendoff (2002), that CS predates language considerably in evolution, being shared at least among primates and possibly other higher animals, and that language has evolved as an external representation for CS. The alternative picture more in line with the evidence for language-specificity discussed here is one according to which the known facts of animal cognition are attributable to SpS, and CS has evolved as a cognitive support system to enable translation between SpS and language.

CONCLUSIONS

Converging evidence from a variety of sources suggests that Motion is consistently framed as state change—CoL—in Yucatec. Verbal cores that describe Motion must be projected from State-Change verbs. Ground phrases denote Place functions and are strictly Path-neutral. Yucatec thus exhibits a more radical type of "verb framing" than the languages considered in Talmy (2000b). CoL-denoting clauses implicate, but do not entail, Motion, as evident from the fact that they are acceptable as descriptions of scenarios in which CoL comes about by the Ground moving or Figure or Ground emerging or disappearing. Such phenomena, first attested in Japanese by Kita (1999), occur on a larger scale in Yucatec. Exempt from application to non-Figure-Motion scenarios are verbs selecting AT-Place functions, presumably because such Functions can be projected only from stationary Grounds. Assuming identity of semantic and CS representations, the compatibility of CoL descriptions with teleportation and Ground Motion events suggests that Yucatec speakers conceptualize and memorize such events in terms of CoL. The case for a possible absence of Path functions from the CS of Yucatec native speakers is further bolstered by the lack of spatiotemporal metaphors expressing two-place temporal ordering relations; these are assumed on localist accounts to be grounded in Path functions. Furthermore, anecdotal evidence points to transfer of Place semantics onto L2-Spanish Ground phrases. If Yucatec encoded Path functions in CS (even without directly expressing them syntactically), such apparent difficulty in the acquisition of L2-Path expressions would be unexpected.

Arguments that may be advanced in defense of Path semantics in English do not apply to Yucatec. Thus, in line with the construal of Motion as CoL, Motion with respect to Route Grounds is semantically underspecified—all events of this type are described with a single verb, *máan* "pass." Complex Motion involving multiple Grounds is broken down into sequences of single-Ground CoL events, each encoded by an independent clause. CoL-denoting clauses can be used metaphorically to describe the extension of spatial objects; but such metaphors are subject to the one-Ground-per-clause rule as well. The meanings conveyed by other "Fictive Motion" metaphors in English are described nonmetaphorically.

Lack of Path semantics may not be rare among the lesser-studied languages of the world. Language-specificity in the representation of Motion at CS may be afforded by the Spatial Structures system of cognition.

ACKNOWLEDGMENTS

This chapter has grown out of discussions among the members of various research projects at the Max Planck Institute for Psycholinguistics, going back as far as the short-lived *Change of State Project* of 1994–1995. I would like to single out Sotaro Kita, who first introduced me and the other members of *Change of State* to the idea of a language-specific framing of Motion as state change, and Steve Levinson, who created the *Motion Verbs* stimulus, which elicited the data presented in the section on Location Change without Figure Motion. I have had many insightful discussions of this topic with both of them over the years, and I am certain that neither of the two will completely agree with the take I develop here. Earlier versions of this chapter have been presented at the University at Buffalo, the University of Rochester, and the 2004 Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas in Boston. I thank the participants of these presentations for comments and suggestions, and I am indebted to the editors of this volume, Barbara Malt and Phillip Wolff, and to Carolyn O'Meara for insightful comments and general help with the final version of the chapter. The research presented here was fully supported by the Max Planck Institute for Psycholinguistics and the University at Buffalo.

Notes

1. SpS was added to the framework in Jackendoff (1987) under the label "Spatial Representations"; the term "Spatial Structure" was introduced in Jackendoff (2002).
2. It is only Motion of a "Figure" with respect to some external reference point—a "Ground"—that can be argued to be represented in terms of Location Change in language, not Motion with respect to some internal axis, such as rotation, spinning, or wobbling. Such nontranslational Motion is represented in language mostly as Manner of Motion (Talmy, 2000b) and ignored in the following.
3. Abbreviations used in morpheme glosses: 3, third person; A, Cross-reference "Set-A" (actor/possessor); ACAUS, Anticausative/middle voice; AN, Animate; B, Cross-reference "Set-B" (undergoer/theme); CL, Classifier; CMP, Completive; CON, Perfective connective; D1, Proximal deictic particle; D2, Distal/anaphoric particle; D3, Textual deictic particle; D4, Negation/anaphoric Place particle; DEF, Definite article (Spanish); DET, Determiner; DIM, Diminutive particle; IMPF, Imperfective; IN, Inanimate (classifier); INC, Incompletive;

INCH, Inchoative; IRR, Irrealis (subordinator); LOC, Locative (Japanese); NEG, Negation; NOM, Nominative (Japanese); PAST, Past tense (Japanese, Spanish); PREP, Generic preposition; PROG, Progressive; PROSP, Prospective; PRS, Present (Spanish); PRSV, Presentative; PRV, Perfective; REC, Recent past marker; REL, Relational derivation; REP, Repetitive particle; RES, Resultative derivation; SG, Singular; SPONT, Anticausative derivation; SUBJ, Subjunctive; TOP, Topic.

4. The term "verbal core," adapted from Van Valin and LaPolla (1997: 25–52), is used here for the maximal syntactic projection of verb stems in Yucatec. There is no evidence of a verb phrase in the customary sense in this language. See Bohnemeyer (2002: 81–129) for a discussion.
5. Cf. also Bohnemeyer (1997, 2007) and Bohnemeyer and Stolz (2006).
6. These are *deictic* or *indexical* verbs just like English *come* and *go* (Fillmore, 1971; Wilkins & Hill, 1995). For instance, the verb *come*, when used without a ground phrase, will be interpreted with the deictic center—the location of the speaker and/or addressee—as the goal. However, in the case of *come*, it is possible to replace the intrinsic deictic goal with one encoded by a ground phrase (e.g., *come to the bookstore*); in the case of Yucatec *tàal* "come," *bin* "go," and *u'l* "return," this is not possible.
7. The verbal cores in (7)–(9) are achievements, i.e., they describe instantaneous events. The same verbs produce accomplishments—events whose completion may be noninstantaneous—if Figure, Ground, or both are conceptualized as spatially extended. In this case, the progressive yields imperfective reference. Imperfective interpretations also occur when the Figure and/or Ground argument are nonquantized (cf. Krifka, 1998)—e.g., when having the reference of bare plurals in English. In this case, the verbal core is atelic. All verbs in Table 6.1 follow this pattern, except for *lúub* "fall," *na'k* "ascend," *em* "descend," *lîik'* "rise," and their causative stems. When occurring without Ground phrases, these verbs can be used as "degree achievements" (Dowty, 1979: 88–91), i.e., as encoding gradual change without a discrete end state or specific degree of change.
8. Note that "Path verbs" [in Talmy's (2000b) parlance] of English, such as ascend and enter, behave like State-Change verbs according to similar aspectual diagnostics. This by itself does not mean that these verbs do not lexicalize Path functions. Direct evidence against the expression of Path functions in the Yucatec CoL verb roots is presented in the section on Location Change without Figure Motion. That the verb roots in Table 6.1 have State-Change event structures is a necessary, but not sufficient, condition for the validity of the claim that Motion is framed as Location Change in Yucatec.
9. Table 6.3 sorts the relational nouns into two sets: those that can be possessed by the Ground-denoting nominal and those that require combination with *ti'* to head a Ground phrase. *Chúumuk* "center" is special in that it permits optional dropping of *ti'*.
10. The hatch mark (#) is employed here for forms or constructions that are structurally well formed, but cannot be used in reference to a particular scenario.
11. Note that the placement of *em* "descend," *lîik'* "rise," *lúub* "fall," and *u'l* "return (to deictic center)" in Figure 6.12 is, by conjecture, based on their semantic relations to the other verb roots; these have not actually been tested for applicability under Ground Motion.
12. Yucatec does in fact have multicore constructions that permit the integration of multiple CoL verbs in a single clause. However, in these constructions, the first core must be projected from *bin* "go" or *tàal* "come" and the second core is an oblique that stands in a purposive relation to the first, such that realization of the event encoded by the second core is not entailed. The pragmatic function of such structures seems to be to add a deictic perspective, as expressed by the first core, to the CoL event denoted by the second core. Multi-CoL sequences such as in (34) cannot be expressed in this way. In other Mayan languages, structures of this kind often grammaticalize, yielding directional particles (Zavala, 1993).
13. Jackendoff (1983: 209–211; 356–360) rejects the analysis of such expressions as metaphors in the context of the Thematic Relations Hypothesis discussed in the section on The Language-Specificity of Path Functions in CS. This question is, however, irrelevant to present matters.
14. Matsumoto (1996) finds differences between Fictive Motion metaphors in English and Japanese that are likewise attributable to lexical and syntactic differences between the two languages in the source domain.

15. Briefly, the subinterval and subpath relations define linear partial orders over time intervals and subpaths such that any two time intervals and any two subpaths either overlap, are adjacent to one another, or are connected by exactly one subinterval/subpath that is adjacent to both; cf. Krifka (1998) and Zwarts (2005).
16. Why is it that the representation of Path information in SpS may not be sufficient to support Path encoding in L2-Spanish? There are two conceivable answers. First, Jackendoff has argued that all information relevant to syntax *must* be encoded in CS. Second, SpS presumably encodes much richer representations of the trajectory of moving entities. In CS, these are reduced to abstract Path functions determined in terms of topological relations with respect to one or more reference entities. Much continuous information about curvature, angles, and distances is lost. This abstraction may not be obvious to speakers of a language such as Yucatec, which does not express it.
17. It appears that it has been an unstated goal of the Conceptual Semantics enterprise to push decomposition of verb meanings in terms of these generalized conceptual categories to its limits. Whereas the set of ontological types is probably indeed small and wholly domain-neutral, the set of domain-neutral conceptual functions may need to be supplemented by an indefinite number of domain-specific functions. This will depend in part on the division of labor between CS and SpS addressed in the following.
18. In Yucatec, the donor/giver of *k'am* "receive" and the recipient of *ts'a'* "give," "put," "provide" are encoded by PPs headed by the same generic preposition *ti'*, thus confirming once more its Path neutrality.

References

Bohnemeyer, J. (1997). Yucatec Mayan lexicalization patterns in time and space.In M. Biemans & J. v. d. Weijer (Eds.), *Proceedings of the CLS Opening Academic Year 1997–1998* (pp. 73–106). Tilburg: Center for Language Studies.

Bohnemeyer, J. (1998). Temporal reference from a Radical Pragmatics perspective: Why Yucatec does not need to express 'after' and 'before.' *Cognitive Linguistics*, *9*, 239–282.

Bohnemeyer, J. (2000). Event order in language and cognition.In H. de Hoop & T. van der Wouden (Eds.), *Linguistics in the Netherlands 17* (pp. 1–16). Amsterdam: John Benjamins.

Bohnemeyer, J. (2002). *The grammar of time reference in Yucatec Maya*.Munich: LINCOM.

Bohnemeyer, J. (2003a). Invisible time lines in the fabric of events: Temporal coherence in Yucatec narratives. *Journal of Linguistic Anthropology*, *13*, 139–162.

Bohnemeyer, J. (2003b). The unique vector constraint. In E. van der Zee & J. Slack (Eds.), *Representing direction in language and space* (pp. 86–110). Oxford: Oxford University Press.

Bohnemeyer, J. (2004). Split intransitivity, linking, and lexical representation: The case of Yucatec Maya.*Linguistics*, *42*, 67–107.

Bohnemeyer, J. (2007). The pitfalls of getting from here to there: Bootstrapping the syntax and semantics of motion event expressions in Yucatec Maya.In M. Bowerman & P. Brown (Eds.), *Cross-linguistic perspectives on argument structure: Implications for learnability* (49-68). Mahwah, NJ: Lawrence Erlbaum.

Bohnemeyer, J., Enfield, N. J., Essegbey, J., Ibarretxe-Antuñano, I., Kita, S., Lüpke, F. and Ameka, F. K. (2007). Principles of event segmentation in language: The case of motion events.Language 83(3): 495-532.Bohnemeyer, J. & Stolz, C. (2006). Spatial reference in Yucatec Maya: A survey.In S. C. Levinson & D. P. Wilkins (Eds.), *Grammars of Space* (pp. 273-310). Cambridge: Cambridge University Press.

Clark, H. H. (1973). Space, time, semantics, and the child.In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York: Academic Press.

Dowty, D. R. (1979). *Word meaning and Montague grammar*. Dordrecht, Netherlands: Reidel.

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

Grice, H. P. (1975). Logic and conversation.In P. Cole & J. L. Morgan (Eds.), *Speech acts* (pp. 41–58). New York: Academic Press.

Gruber, J. S. (1965). *Studies in lexical relations*. Doctoral dissertation, Massachusetts Institute of Technology.

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

Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Jackendoff, R. (1992). *Languages of the mind*. Cambridge, MA: MIT Press.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface.In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 1–30). Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.

Kita, S. (1999). Japanese ENTER/EXIT verbs without motion semantics. *Studies in Language, 23*, 307–330.

Krifka, M. (1998). The origins of telicity.In S. Rothstein (Ed.), *Events and grammar* (pp. 197–235). Dordrecht: Kluwer.

Lehmann, C. (1992). Yucatecische lokale Relatoren in typologischer Perspektive [Yucatec spatial relators in typological perspective]. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, *45*, 626–641.

Levin, B. (1993). *English verb classes and alternations*. Chicago, IL: University of Chicago Press.

Levinson, S. C. (2000). *Presumptive meanings.* Cambridge, MA: MIT Press.

Levinson, S. C. (2001). Motion verb stimulus, version 2.In S. C. Levinson & N. J. Enfield (Eds.), *'Manual' for the field season 2001* (pp. 9–11). Nijmegen: Max Planck Institute for Psycholinguistics.

Levinson, S. C., Meira, S., & The Language and Cognition Group (2003). 'Natural concepts' in the spatial topological domain—adposition meanings in crosslinguistic perspective. *Language*, *79*, 485–516.

Marr, D. (1982). *Vision*. New York: Freeman.

Matsumoto, Y. (1996). Subjective motion and English and Japanese verbs. *Cognitive Linguistics*, *7*, 183–226.

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

Talmy, L. (1985). Lexicalization patterns. In T. Shopen (Ed.), *Language typology and syntactic description. Vol. 3: Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge University Press.

Talmy, L. (1996). Fictive motion in language and "ception." In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 211–276). Cambridge, MA: MIT Press.

Talmy, L. (2000a). *Toward a cognitive semantics. Vol. I: Concept structuring systems*. Cambridge, MA: MIT Press.

Talmy, L. (2000b). *Toward a cognitive semantics. Vol. II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Traugott, E. C. (1978). Spatio-temporal relations. In J. H. Greenberg (Ed.), *Universals of human language. Vol. 3: Word structure* (pp. 369–400). Palo Alto, CA: Stanford University Press.

Van Valin, R. D., Jr. & LaPolla, R. J. (1997). *Syntax*. Cambridge, UK: Cambridge University Press.

Wilkins, D. P., & Hill, D. (1995). When GO means COME: Questioning the basicness of basic motion verbs. *Cognitive Linguistics*, *6*, 209–259.

Zavala, R. (1993). *Clause integration with verbs of motion in Mayan languages*. MA Thesis, University of Oregon.

Zwarts, J. (2005). Prepositional aspect and the algebra of paths.*Linguistics and Philosophy*, *28*, 739–779.

7 CATEGORIES IN MIND AND CATEGORIES IN LANGUAGE

Do Classifier Categories Influence Conceptual Structures?

Mutsumi Imai & Henrik Saalbach

In this chapter, we explore the relation between classifier grammar systems—grammatical systems that categorize objects/entities into over 100 grammatical categories—and how people think about objects. This is of course a question concerning linguistic relativity, or the Whorfian hypothesis (Whorf, 1956). Many researchers have asked whether linguistic categories, be they lexical or grammatical, influence people's concepts and cognitive processes (for reviews see Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996; Hunt & Agnoli, 1991; Lucy, 1992). However, researchers differ in how the impact of language should be evaluated. Thus, we first review different views of how the role of language (linguistic categories) should be evaluated in thought. We then specify our own position concerning this issue, reporting empirical results of two series of studies on the impact of classifier systems. We found some influence of classifiers in some cognitive tasks but not others in Chinese speakers. We found no influence of classifiers in any of the cognitive tasks in Japanese speakers, including the tasks in which Chinese speakers exhibited the influence of classifiers. Based on these results, we attempt to clarify the nature of the influence of classifiers, specifying how large the classifier effect would be relative to other major conceptual relations such as taxonomic or thematic relations, in what cognitive contexts the language-specific classifier effect is observed (and in what contexts it is not observed), and how the language-specific classifier effect might arise. We argue that it is time for us to go beyond a pro-Whorf or anti-Whorf conclusion. Instead of seeking a cross-linguistic difference in one task for the purpose of establishing evidence for (or against) the linguistic relativity hypothesis, we must investigate the relation between language and thought to reveal complex interactions between the semantic as well as structural nature of the grammatical system and the type of cognitive activities and to evaluate how pervasive and important the influence of a target linguistic categorization system is in a full range of cognitive processes.

INFLUENCE OF LANGUAGE ON THOUGHT WITHIN THE REALM OF LANGUAGE USE OR BEYOND

Some researchers (e.g., Lucy, 1992, this volume; Gennari, Sloman, Malt, & Fitch, 2002; Vigliocco, Vinson, Paganelli, & Dworzynski, 2005) emphasize the distinction between two versions of linguistic relativity: In one version, the influence of language is observed within the realm of language use (e.g., paying attention to and encoding a particular aspect of the world to talk about events and objects); in the other, the influence goes beyond the realm of language use and penetrates into the realm of nonlinguistic cognition. The first version is often referred to as "thinking for speaking" (Slobin, 1987), and is considered a weaker version of linguistic relativity, whereas the second version is stronger and is "true" linguistic relativity (Lucy, 1992).

However, determining the boundary between "linguistic" and "nonlinguistic" cognition is not simple. It appears that almost every researcher in the field has a different idea about it (e.g., Gennari et al., 2002; Imai & Mazuka, 2003, 2007; Lucy, 1992). One way to distinguish "linguistic" and "nonlinguistic" effects may be to determine whether the effect is obtained with or without the explicit invocation of the target linguistic categories. For example, Vigliocco and colleagues (Vigliocco, Vinson, Indefrey, Levelt, & Hellwig, 2004) found that in their semantic substitution paradigm, German speakers tended to produce substitution errors within the same gender category when speakers produced phrases with determiners marked for gender, but that this gender preservation effect disappeared when they produced bare nouns or phrases with determiners not marked for gender (a phrase with an indefinite determiner plus noun).

Another way of distinguishing "linguistic" and "nonlinguistic" effects has been the usage of different kinds of stimulus material, verbal and pictorial, respectively. Underlying this is the assumption that pictures but not the names have direct access to the conceptual (i.e., nonlinguistic) representations of the objects. For example, Vigliocco and colleagues (Vigliocco et al., 2005) determined whether grammatical gender categories would have a direct impact on semantic representation of Italian, using an odd-one-out categorization task. They found that Italian speakers chose two verbally presented objects of the same gender category (in Italian) at a rate higher than English speakers (although this gender effect was found only for animal stimuli but not for artifact stimuli). However, the gender effect was not observed when objects were presented in pictures instead of words. On the basis of these results, Vigliocco and colleagues (2005) concluded that gender categories in Italian exert their influence at the lexicosemantic level but not at the conceptual level, and hence argued that their results are consistent with the "thinking for speaking" hypothesis but are evidence against the linguistic relativity hypothesis.

In our view, however, the two effects cannot be so simply distinguished just by the use of different types of stimulus material, as pictorial presentation has its own limitations for assessing people's "nonlinguistic" representations. First, even if the objects were presented in pictures, the participants may have unconsciously named the objects. Second, a picture may invoke a strong visual image of the particular instance of the object depicted in it. Thus, visual similarity among pictures may affect judgments of similarity more strongly than when objects were presented in words without specific visual images. In fact, one possible account for the disappearance of the gender effect with the pictorial stimuli in the study of Vigliocco et al. (2005) is that concrete visual images of objects wiped out the subtle conceptual similarity between objects arising from gender category membership that had shown up in the similarity judgment task with words.

EVALUATING THE INFLUENCE OF LANGUAGE IN LIGHT OF MEANINGFULNESS AND MAGNITUDE

What seems to be more important than characterizing the target effect along the dimension of being "linguistic" or "nonlinguistic" is to specify when (i.e., in what cognitive contexts) and how a given linguistic categorization system affects cognitive processes and representations, as well as to specify how large and how meaningful the influence is. Again, take the effect of gender-marking grammar for example. Provided that speakers of a language with a gender-marking grammar rate two objects from the same gender category more similar (even when the grammatical gender was not evoked) compared to speakers of a different language that does not have gender-marking grammar at a statistically significant level, how should we evaluate this effect in a global picture of human concepts and cognition? Should we interpret this data to suggest that the speakers of the two languages "think differently" and have "different conceptual structures"?

Again, the answer depends on how we define "thought" or "conceptual structures." Here, we would like to operationally define "conceptual structures" as how people organize their knowledge of each object into a coherent body of knowledge so that knowledge of each object is related and can be used to infer properties of other objects. In this light, to make a strong claim that a linguistic categorization system "shapes thought," we may want to see the system serve as a basis for how we relate things in the world. If a given conceptual relation is an important and meaningful one, we would not only use it as a basis of categorization or similarity judgments, but also use it for inductive inference about unseen properties of novel objects. One important function of categories is to promote inductive inferences, as they enlarge the scope of knowledge and allow predictions about novel items (Medin, 1989; Murphy, 2002). Also, strong conceptual relations are expected to be accessed automatically as shown in many studies using the semantic priming method (e.g., Meyer & Schvaneveldt, 1971; Neely, 1977).

Two conceptual relations have been particularly noted to function as connecting knowledge of individual objects. One is, of course, taxonomic relations. Taxonomic categories are denoted by nouns, and include items of the same kind. They are differentiated into levels of varying specificity (e.g., animal, dog, collie) related by class inclusion (e.g., a collie is a dog, a dog is an animal, a collie is an animal). Numerous studies have shown that taxonomic relations organize concepts and provide a basis for categorization, similarity judgments, and inductive reasoning (e.g., Markman, 1989; Osherson, Smith, Wilkie, López, & Shafir, 1990; Gelman & Markman, 1986, Waxman & Gelman, 1986) both in children and adults, and are included in the semantic network that is automatically activated in fast-speed processing (e.g., Yokosawa & Imai, 1997).

However, recently, researchers have noted that thematic relations are also an integral and important part of the conceptual structure not only for children (e.g., Imai, Gentner &Uchida, 1994; Markman, 1989; Smiley & Brown, 1979) but also for adults (e.g., Lin & Murphy, 2001; Wisniewski & Bassok, 1999; see also Bassok & Medin, 1997). Lin and Murphy (2001, see also Markman, 1989) suggest that many human concepts include knowledge about nontaxonomic relations, with thematic relations being most important among them. They define thematic relations as external relations that arise through objects cooccurring or interacting together in space or time, or objects being linked by functional or causal relationships (e.g., table and chair, morning and newspaper, scissors and paper). Through a series of experiments with varying paradigms, they demonstrated that thematic relations play a role not only in similarity judgments but also in inductive reasoning about properties. They have also shown that the thematic relations are evoked not only in conscious, strategic cognitive processes but also in fast-speed, automatic processes (see also Yokosawa & Imai, 1997). Based on these results, Lin and Murphy (2001) suggested that thematic relations play a prominent role in conceptual structure in well-educated young American adults.

One important point in evaluating the effect of linguistic relativity might be, then, to compare the size of a potential language-specific effect of the classifier categorization system to that of other major conceptual relations such as taxonomic and thematic relations. If the speakers of the language rely on the language-specific conceptual relation in question equally or more strongly than the taxonomic and/or thematic relations, a very strong case is made for the linguistic relativity hypothesis, and it could be concluded that the target linguistic categorization system indeed "structures" or "organizes" the speakers' concepts. On the other hand, even if a statistically significant crosslinguistic difference is found (e.g., similarity between pairs of objects from the same grammatical classifier category is rated higher by speakers of the language having that grammatical system than by speakers of a language that does not have such a linguistic system), this effect may not be strong enough grounds for a claim that the target linguistic categorization system "structures" (or "shapes," even) the speakers' concepts if the effect is much weaker than other conceptual relations.

We therefore suggest that we should be more cautious about advancing a strong and attractive conclusion that language "shapes" thought or language "structures" concepts. To advance such a strong conclusion, the following points should be considered: (1) whether the conceptual relation underlying the target linguistic categories serves as a basis not only for similarity judgments but also for inductive reasoning, (2) whether the conceptual relation is evoked automatically, and (3) whether the effect of the conceptual relation is comparable to that of other major conceptual relations such as taxonomic and thematic relations.

DOES A CLASSIFIER CATEGORIZATION SYSTEM INFLUENCE THOUGHT?

In this chapter, we examine the effect of a classifier categorization system on thought, along with the theoretical points we laid out above. In addition to nouns, objects are further categorized by grammar. Different languages have developed a broad variety of nominal classification systems including the classifier grammar system, the count/mass grammar system, and the grammatical gender system (e.g., Aikhenvald, 2000; Senft, 2000). The way language subcategorizes nouns varies across different languages, and it can be characterized as a continuum, with a fully obligatory grammatical classification on one end and a fully "lexical" classification on the other end (Grinevald, 2000). Examples of the grammatical type are count-mass grammar and the gender grammar systems. For example, languages with count-mass grammar (e.g., English, German) classify nouns in two semantic categories: the category of individuated things (e.g., people, animals, machines) and the category of non-individuated things (e.g., water, sand, butter) and this classification must be done for all nouns, including concepts denoting nonphysical entities. Likewise, languages with gender grammar systems (e.g., German, French, Spanish) classify all nouns into a small number (usually two or three) of gender classes. An example of a lexical system of nominal classification is quantifiers that express quantities of nouns (e.g., a glass of water, a spoon of flour) and that classify nouns only temporarily but not inherently.

The classifier system is located somewhere in between the lexical and the grammatical extreme of the continuum. It differs from the count-mass grammar system or the gender grammar system in that classifiers classify nouns into over 100 classes according to the noun's inherent semantic features.[1] Classifiers seem to have two important semantic functions. First, they serve to individuate the referent by providing a unit of counting to the noun just as quantifiers in language with count-mass grammar do [yi (one) wan (bowl) mi (rice) "a bowl of rice" in Chinese]. However, different from quantifiers that are used only for quantifying mass nouns, classifiers are applied to all nouns when quantifying them, including what seems to be clearly individuated objects. Second and more importantly, classifiers classify nouns and provide additional semantic information to nouns that are classified (Senft, 2000).

Many researchers have attempted to specify semantic criteria for classifier systems in different languages (e.g., Craig, 1986; Denny, 1986; Downing, 1996; Gomez-Imbert, 1996; Senft,1996) and have identified several universal semantic features that serve as criteria for dividing entities into classifier categories (e.g., Adams & Conklin, 1973; Allan, 1977; Croft, 1994; Denny, 1979). The system of classification by classifiers is complimentary to the system of classification by nouns and, hence, categories created by the classifier largely cross-cut categories created by nouns. In particular, whereas the noun lexicon is structured hierarchically around taxonomic relations, the system of classification by classifiers is usually organized around semantic features such as animacy, shape, function, size, rigidity, and social importance, and it does not have hierarchical structures as the noun lexicon does. Let us look at some examples from Chinese classifiers. *Tou* is a classifier for big animals such as cow, elephant, and rhino. *Tiao* is used for objects that are long and curved or flexible, including both animals and inanimates, such as road, jumping rope, snake, or fish. *Ba* is used for objects with a handle or objects that can be grasped by the hand (e.g., umbrella, screw driver, broom, key, or comb).

Given that a classifier system carves up the world in a very different way from taxonomic categories, what cognitive consequences should we expect with respect to representation and cognitive processes, if there are any? One possibility is that classifier systems provide an alternative organization of object concepts that results in categories that are nonexistent for speakers of nonclassifier languages. If this is the case, we expect that speakers of a classifier language and those of a nonclassifier language will behave very differently in almost all cognitive activities including category formation, similarity judgments, and, most importantly, inductive reasoning, as discussed earlier. Furthermore, if the classifier system provides a way of organizing object concepts, we may expect that classifier relations are accessed automatically in on-line processing. Of further interest here is to see the magnitude of the effect due to classifier categories relative to that of taxonomic or thematic relations, as discussed earlier. If speakers of a classifier language utilize classifier category membership in a range of cognitive activities to a greater degree than they utilize taxonomic and thematic relations, that would suggest that classifiers truly serve as the most dominant organizer of the speakers' conceptual structure. Even if the classifier relations are utilized to an equal or smaller degree than the taxonomic or thematic relations, as long as we see the effect in multiple cognitive contexts, especially in inductive reasoning, we would conclude that classifier categories serve as an organizer of the speakers' conceptual structure, along with (and in parallel to) taxonomic and thematic relations.

A second possibility is that the classifier system is not qualified to be considered a major organizer of the speakers' concepts in light of the criteria discussed earlier, but the experience of linguistically categorizing objects by the use of classifiers may heighten the sense of similarity, and as a consequence, similarity among objects that are members of the same classifier category may be magnified in speakers of a classifier language. If this is the case, the difference between speakers of a classifier language and those of a nonclassifier language may be observed in similarity judgments, but similarity due to classifier relations would not exceed similarity due to taxonomic or thematic relations. Furthermore, even though the same-classifier relations may influence inductive reasoning in the context in which similarity is the only available source for the inference, the classifier influence may vanish when other sources such as background knowledge is available.

A third possibility is that classifiers are "frozen," linguistic conventions, and do not have any cognitive impact on speakers of classifier languages. In this case, we should not see any difference between speakers of a classifier language and those of a nonclassifier language, although the two groups may differ due to factors other than the classifier categorization system (e.g., culture).

To our knowledge, there are only a few studies in the literature that directly addressed these possibilities. One such study is research by Zhang and Schmitt (1998), who asked whether classifiers influence perception of similarity between two objects. They had speakers of Chinese and English rate the similarity of pairs of everyday objects. Half of the pairs consisted of objects that share the same classifier in Chinese and half of the pairs consisted of objects from different classifier categories. Zhang and Schmitt (1988) found that Chinese speakers rated the same classifier pairs more highly than the native English speakers do, whereas ratings of the different classifier pairs did not differ cross-culturally. On the basis of these results, they concluded that classifier categories strongly affect the speakers' conceptual organization, saying that "The results obtained were a strong indication that objects sharing the classifier are grouped into schematic organizations in Chinese speakers' mental representations. That is, although English speakers may group these objects on the basis of their conceptual similarity, Chinese speakers seem to add a linguistic categorization to the classification of objects" (p.381).

Zhang and Schmitt's results can indeed be taken to be "some" support for the linguistic relativity hypothesis. However, it is not clear whether their results suggest that Chinese speakers' organization of object concepts is significantly different from that of English speakers due to classifier categories, because their experimental design does not allow us to determine how we should interpret the effect they found in light of the criteria we suggested earlier. If Chinese speakers draw inductive inferences on the basis of classifier category membership and the classifier category membership is automatically evoked even when a classifier is not explicitly expressed, we would agree that the classifier system indeed provides Chinese speakers with a way of organizing objects that English speakers do not possess. However, even if this is the case, we would like to know the magnitude of the impact of the classifier system relative to the impact of other major conceptual relations in order to evaluate how important the classifier system is as one of the organizers of concept in the mind of the speakers. It is also possible that the classifier effect found by Zhang and Schmitt (1998) was limited to similarity ratings, and that the classifier relations are not utilized in inductive reasoning, or not activated in automatic processing. If so, claiming that the classifier system adds a new way of organizing concepts might be an overstatement, though it may still be taken as a weak form of linguistic relativity.

To test these possibilities, we conducted two series of cross-linguistic studies, each of which consisted of multiple cognitive tasks. Study 1 compared Chinese, a classifier language, and German, a nonclassifier language. Study 2 included a second classifier language, Japanese, in addition to Chinese and German to see whether the effect of classifiers is observed across different classifier languages, given the structural difference between Chinese and Japanese. Specifically, as we describe in more detail, classifiers accompany nouns much more systematically and frequently in Chinese than in Japanese. We explore whether this linguistic difference affects the magnitude and nature of the classifier effect for Chinese and Japanese speakers.

STUDY 1: EXAMINATION OF CLASSIFIER CATEGORIES IN CHINESE SPEAKERS' CONCEPTUAL STRUCTURES

Study 1 examined whether the Chinese classifier categories are utilized in categorization, similarity judgments, and inductive reasoning of a novel property, and whether classifier relations are automatically activated in online processing (Saalbach & Imai, 2007). For this purpose, we compared Chinese and German adults on forced-choice categorization, similarity judgments (on a rating scale), inductive inference of novel properties, and fast-speed word–picture matching tasks. We designed a stimulus set of everyday objects in a way that allowed us to examine four types of relations around the same target (e.g., FLOWER) using four test objects, each of which represented one of the four conceptual relations: (1) taxonomic (e.g., TREE), (2) thematic (e.g., VASE), (3) classifier (e.g., CLOUD), and (4) no-relation (SHOE). The four relations were orthogonally crossed, so that the object serving as the same-classifier item was not related to the target taxonomically or thematically. Likewise, neither the taxonomic item nor thematic item belonged to the same-classifier category in Chinese. Table 7.1 shows the stimulus set of our first study. The objects were presented verbally in the categorization, similarity judgments, and property induction tasks (i.e., in words).

TABLE 7.1. Stimuli Items Used for Study 1

Standard	Classifier	SameClassifier	Taxonomically Related	Thematically Related	Control
Comb	*Ba*	Key	Hair dryer	Hair	Ticket
Pistol	*Ba*	Umbrella	Canon	Bullet	Stamp
Scissors	*Ba*	Fan	Cutter	Paper	TV
Chain	*Tiao*	Carp	Rope	Lock	Poster
Necklace	*Tiao*	Blanket	Ring	Dress	Book
Towel	*Tiao*	Eel	Handkerchief	Shower	Potato
Mountain	*Zuo*	Tower	Hill	Snow	Necklace
Bell	*Zuo*	Building	Buzzer	Temple/church	Bike
Piano	*Jia*	Ladder	Violin	Music book	Scarf
Plane	*Jia*	Swing	Boat	Airport	Chain
Flower	*Duo*	Cloud	Tree	Vase	Cup
Newspaper	*Zhang*	Bed	Book	Morning	Tube
Drum	*Mian*	Wall	Trumpet	Sticks	Scissors
Tent	*Ding*	Hat	Sleeping bag	Campfire	Table

Culture versus Language

The design of the stimuli also provided us with a unique opportunity to examine a specific hypothesis concerning the extent to which the structure of object concepts is universally shared and to which it is malleable by environmental factors such as culture and language. Nisbett and colleagues have put forward a bold proposal (Nisbett, 2003, Nisbett, Peng, Choi, & Norenzayan, 2001) concerning the role of culture. They argue that philosophy, values, and customs that have been nursed in a culture throughout its history lead to a "culturally specific" style of cognition (Nisbett, 2003, Nisbett et al., 2001). In his empirical work, Nisbett focused on the comparison between East Asians and Westerners. Characterizing the former as "holistic" and the latter as "analytic," Nisbett and colleagues argued that East Asians tend to view the environment as a unified whole and pay much attention to relations that tie elements in the environment. Westerners tend to focus individual elements of the environment separately. Based on this scheme, they have made a specific prediction regarding the conceptual structure of East Asians and Westerners: East Asians, with their predisposition to see a scene or event as a whole, are expected to categorize the world around thematic relations; Westerners, with their focus on properties of individual objects, are expected to categorize the world by taxonomic relations. Ji, Zhang, and Nisbett (2004) in fact reported that monolingual Chinese people showed a preference for "relational" groupings whereas European Americans tended to group things "categorically."

However, as in the case for the classifier effect in the study of Zhang and Schmitt (1998), it is not clear whether the data of Ji et al. (2004) warrant a strong conclusion that object concepts are organized differently for Easterners and Westerners. Ji et al. (2004) showed that Chinese college undergraduates have a relatively stronger preference for thematic-based groupings than American undergraduates, and conversely, the American undergraduates showed a relatively stronger preference for the taxonomic-based groupings than the Chinese undergraduates. However, again, we do not know whether this cross-cultural difference in the relative preference between taxonomic-based and thematic-based groups holds for inductive reasoning or if it is seen in the automatically activated semantic network. Furthermore, as we reviewed earlier, Lin and Murphy (2001) demonstrated that even educated European American young adults sometimes show a preference for categorizing objects based on thematic relations over taxonomic relations. Thus, it is important to determine whether the East–West differences found in the study of Ji et al. (2004) warrant a strong claim that Easterners and Westerners have different ways of organizing objects concepts.

In our paradigm, although the influence of language (the impact of a classifier system) and culture (culture-specific biases of Easterners and Westerners) on conceptual structures was examined simultaneously, we do not need to evaluate the effect of culture and language in a mutually exclusive, black-and-white fashion. For example, it is possible for us to find that Chinese speakers show stronger sensitivity both to classifier relations and to thematic relations than German speakers. Our paradigm also allows us to evaluate the relative importance of taxonomic relations, thematic relations, and classifier relations within the culture. Thus, it is possible that we find that the people from the two culture/language groups show reliance on the three types of relations in the same order but nonetheless find that the two groups rely on the three relations to different degrees. In this way, we can place the effects of culture and language, if we find any, in a global picture of the conceptual structure of everyday objects.

Procedure and Results of the Categorization Task The Chinese (undergraduates of Peking University in China) and German (undergraduates of Berlin Institute of Technology) participants were shown a triad of objects, one of which served as the standard and the other two of which served as test items. They were asked to determine which of the two test items best matched the target item. As stated earlier, our stimuli included a taxonomic item, a thematic item, a classifier item, and a control item around the same target object. We thus constructed six types of contrasts around the same target item by making pairwise combinations of the four relations: (1) classifier vs. taxonomic, (2) classifier vs. thematic, (3) classifier vs. control, (4) taxonomic vs. thematic, (5) taxonomic vs. control, and (6) thematic vs. control.

The results are shown in Table 7.2. They indicate that classifier categories were not used as the basis for categorization by Chinese speakers. When the same-classifier item was pitted against the taxonomic or the thematic item, the Chinese as well as the German speakers made categories exclusively on the basis of the taxonomic or the thematic relations. When the same-classifier item was contrasted with the object that was not related to the target object (control), both the Chinese and German participants judged the same-classifier item to be the better match to the target. This finding suggests that there is an inherent similarity among objects belonging to the same-classifier category even when they do not share any taxonomic or thematic relations, and this inherent similarity is detectable by

TABLE 7.2. Mean Percentages of Classifier, Taxonomic, and Thematic as Well as Control Item Choices across Conditions of the Categorization Task of Study 1

Relation	Chinese ($N = 23$)	German ($N = 24$)
Classifier vs. Taxonomic		
Classifier	17.1	11.0
Taxonomic	82.9	89.0
Classifier vs. Thematic		
Classifier	17.1	15.2
Thematic	82.9	84.8
Classifier vs. Control		
Classifier	76.1	71.1
Control	23.9	28.9
Classifier vs. Thematic		
Taxonomic	36.0	34.5
Thematic	64.0	65.5

speakers of a nonclassifier language and is used when there is no other kind of similarity to resort to in forming categories. However, it is not the kind of similarity even speakers of a classifier language rely on for categorization, when taxonomic or thematic relations are present.

In addition, we did not find evidence for the proposal put forward by Ji et al. (2004) that Westerners organize object concepts around taxonomic relations and Easterners organize them around thematic relations. Unexpectedly, not only the Chinese but also the German participants preferred the thematic match over the taxonomic match (see the converging results in American undergraduates in Lin & Murphy, 2001; for more detailed discussion of this results, see Saalbach & Imai, 2007).

In summary, we did not find evidence that classifier categories are used as bases for categorization in the face of taxonomic and thematic relations, when classifiers are not explicitly mentioned. The performance of Chinese and German speakers in the categorization task was strikingly similar. However, this does not preclude the possibility that the classifier system affects the speakers' cognition in a subtler way (e.g., heightening attention to semantic features underlying classifier categories) on more sensitive tasks. We thus conducted similarity judgment and inductive reasoning tasks using a rating scale.

Similarity Judgments Chinese and German speakers were presented with pairs of objects and asked to judge similarity between the two objects on a rating of scale of 1 (very dissimilar) to 7 (very similar). The pairs were drawn from the 14 sets of quintuplets that were used in the categorization task. Around the same target object, four pairs were constructed representing taxonomic, thematic, same-classifier, and unrelated relations (see above). As in the forced choice categorization tasks, the objects were presented in words rather than in pictures (see Saalbach & Imai, 2007, for methodological details for this task as well as for other tasks).

The overall pattern of rated similarity was very similar across the two language groups. Figure 7.1 shows the mean adjusted similarity scores for each of the taxonomic, thematic, and same-classifier relation. Participants in both language groups gave the highest ratings for the taxonomic pairs, followed by the thematic pairs, followed by the classifier pairs. Consistent with the results of the categorization task, both Chinese and German participants rated the same-classifier pairs as more similar than the control pairs. This result again suggests that even speakers of a nonclassifier language can detect an inherent similarity between objects belonging to the same-classifier category. However, this inherent similarity may be magnified for speakers of the classifier language, as Chinese participants' similarity judgments for pairs drawn from the same-classifier classes were higher than those of the German participants, even with the correction for overall higher ratings.

Here, we also found some evidence for the proposal of Ji et al. in that Chinese speakers gave higher similarity ratings for thematically related object pairs than the German speakers, and when we tested the effect of culture in the contrast between the effect of the taxonomic item and that of the thematic item by a preplanned contrast after an overall ANOVA, a significant effect of culture was detected. These results suggest that language and culture could influence people's cognitive processes simultaneously, and warrants reconsideration of the traditional approach that assumes the influence of language and culture to be contrastive and asks *which* is *the* factor to shape thought.

Property Induction To assess whether Chinese speakers utilize classifier category membership as a basis of inductive reasoning, we had Chinese and German participants rate the likelihood that the two objects in the pair share an unknown property. They were instructed as follows: "Suppose that property X is an important property for [Object 1]. If [Object 1] has property X, how likely is it that [Object 2] also has property X?" They were

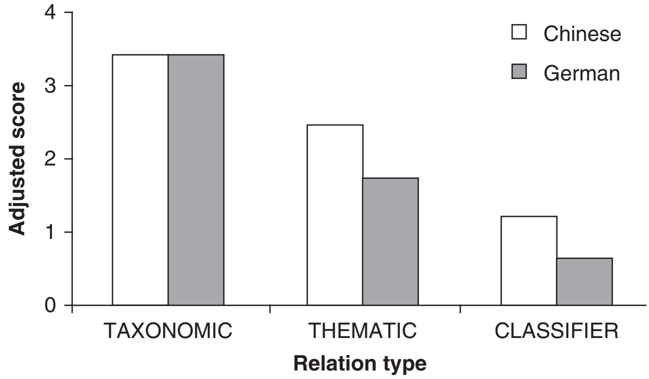


FIGURE 7.1. Adjusted scores for each relation type in each language in Similarity Judgments of Study 1.

Note: The adjusted score was obtained by subtracting the raw score for the control item from each target relation in each set within each language group. We did this manipulation to adjust the baseline differences across the different language groups. For example, Chinese participants tended to give higher ratings than Germans and Japanese for all stimuli types, including the control (unrelated) pairs; and Japanese tended to response more quickly than Germans and Chinese in the speeded word-picture matching task across all conditions.

asked to judge the likelihood on a rating scale of 1 (not likely at all) to 7 (very likely). The object pairs were the same as those used in the similarity judgment task.

As shown in Figure 7.2, the pattern of the results of this study was strikingly similar to the pattern observed for the similarity judgments. Participants in both language groups rated the likelihood in the order of the taxonomic, thematic, same-classifier, and control (unrelated) items. As in similarity judgments, not only Chinese but also German speakers judged it more likely that the same-classifier items shared the same unknown property X with the target than the control items did. At the same time, parallel to the results of similarity judgments, likelihood ratings for the same-classifier items as well as for the thematic items were magnified by Chinese speakers compared to German speakers. This result again shows that the influence of language and culture may not necessarily be mutually exclusive.

The result suggests that classifier category membership not only heightens similarity, but also carries some inductive potential. However, inference of a blank property did not allow the participants to recruit any specific knowledge; thus they had nothing to resort to but similarity (Osherson et al., 1990). It is interesting to see whether people utilize classifier relations in property inference even in a context in which they are able to recruit some background knowledge. We thus conducted a second inductive reasoning task. We used the property "carry the same bacteria," which

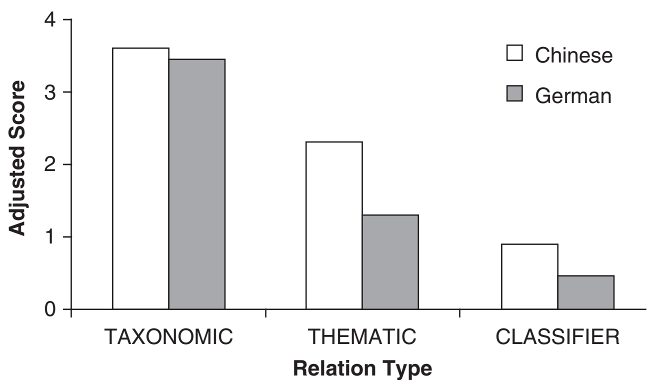


FIGURE 7.2. Adjusted scores for property induction (about blank property) for each target type in each language in Study 1.

was used by Lin and Murphy (2001). Participants were asked "How likely is it that [Object 1] and [Object 2] carry the same bacteria?" and judged the likelihood on a rating scale of 1 (not likely at all) to 7 (very likely).

This time, neither Chinese nor German participants rated the same-classifier item as having a higher probability than the control item in carrying the same bacteria as the target object, as shown in Figure 7.3. Although both Chinese and German speakers noted the similarity underlying classifier categories in the similarity judgment task, neither utilized this similarity in inductive reasoning in this context. Furthermore, the language-specific classifier effect observed in the inference of a blank property was no longer found here. The results from the two inductive inference tasks thus suggest that when people make an inductive inference from a completely unknown property, people use similarity as a basis for inductive reasoning. As classifier relations influence Chinese speakers' construal of similarity, classifier relations influence Chinese speakers' inductive inference in this context. However, when they determined the likelihood of the two objects carrying the same bacteria, they engaged in causal reasoning by utilizing existing knowledge about the conditions in which same bacteria were likely to be found. Specifically, as noted by Lin and Murphy (2001), people may have decided that taxonomically related objects were likely to carry the same bacteria because things of the same kind may have similar living conditions for a kind of bacteria. Likewise, they probably decided that thematically related objects were likely to carry the same bacteria because the transmission of bacteria depends on external contact among items that cooccur in space and time. In this case, Chinese as well as German participants clearly decided that the kind of relation underlying classifier category membership (e.g., shape similarity, size, rigidity, functionality) would not heighten the likelihood of the two objects having the same bacteria.

Speeded Word–Picture Matching We also examined whether the language- and culture-specific differences that we observed in the similarity judgment and inductive inference of a blank property tasks are observed in a task that accesses fast and automatic processes. For this purpose we used a version of a semantic priming paradigm. It is widely known that recognition of a word involves activation of its corresponding node in a semantic network, and a priming effect is observed when two objects that are presented sequentially are conceptually related (Anderson, 1983; Joordens & Becker, 1997; Tulving & Schacter, 1990). Yokosawa and Imai (1997) have demonstrated that the conceptual priming effect is observed in picture recognition as well. In their study, participants (Japanese adults) saw a cue, which was presented either orthographically or pictorially, followed by a target picture. The participants were to judge whether the target matched the cue. Yokosawa and Imai (1997) found that regardless of whether the cue was presented as a word or a picture, when the cue was

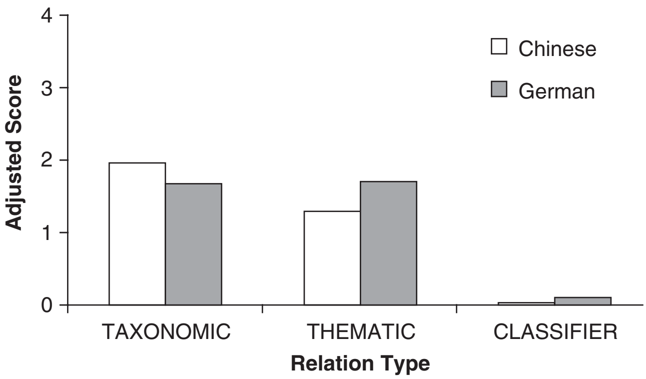


FIGURE 7.3. Adjusted scores for property induction (on the "carrying the same bacteria" question) for each target type in each language in Study 1.

taxonomically (e.g., dog) or thematically (e.g., carrot) related to the target (e.g., rabbit), the participants took a longer time to judge that the cue and the target were different objects (at the basic level) than when the cue was unrelated to the target (e.g., hammer). Here, they demonstrated that conceptual relatedness delays (interferes) with participants' judgments that two objects (the cue and the target) are indeed different.

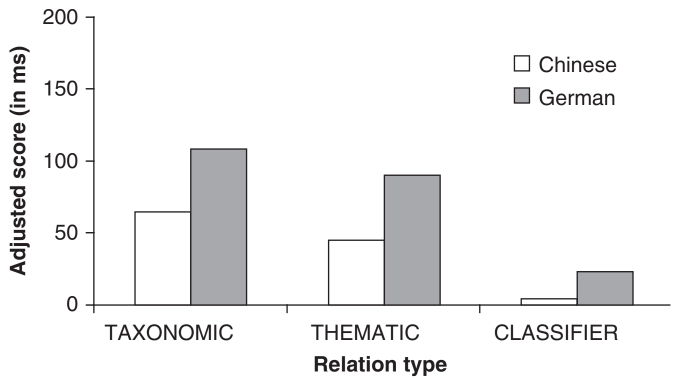
We thus borrowed this paradigm to test whether the influence of classifier relations is observed in Chinese but not in German speakers. Of additional interest was to see whether the thematic effect was larger for Chinese than for Germans and the taxonomic effect would be larger for German than for Chinese participants, to test Nisbett and colleagues' culture-specific mode of cognition proposal (Nisbett, 2003; Ji et al., 2004). The same 14 quintuplets were used for this study. The target object was presented pictorially, and the cue representing one of the four relations (taxonomic, thematic, same-classifier, unrelated) was presented as a written word. The participants were instructed to verify whether the picture matched the word they had just seen (see Saalbach & Imai, 2007, for the methodological details).

Figure 7.4 shows the mean adjusted response time for each relation type for Chinese and German speakers. Adjusted response times were obtained by subtracting the response time for the control item in each set from the taxonomic, thematic, and classifier item. As in the categorization, similarity judgment, and inductive reasoning tasks, the taxonomic and thematic relations strongly affected the participants' response latencies in both language groups. In contrast, the classifier relation did not influence picture recognition of the target object in the Chinese group.

Taken together, it appears that the language-specific influence of the classifier classification system obtained for the unspeeded similarity judgments does not hold for a task that requires fast, automatic cognitive processes. The lack of the priming effect due to classifier category membership in Chinese speakers suggests that objects belonging to the same classifier category are not automatically activated when the target object is accessed, whereas taxonomically and thematically related objects are both activated.

It should also be noted that even though thematic items were perceptually very dissimilar to the target objects, the magnitude of delay due to thematic relations did not differ across the two groups. This result again supports the proposal that thematic relations are a universally important and integral part of the conceptual structure (Lin & Murphy, 2001; Wisniewski & Bassok, 1999), and that, unlike classifier relations, thematically related objects are automatically activated with the target object. However, the thematic effect was no larger for the Chinese group than for the German group nor was the taxonomic effect any larger for the German group than the Chinese group. Here, thus, support for Nisbett and colleagues' (Ji et al., 2004;

FIGURE 7.4. Adjusted scores for response latencies for each relation type in each language in the Word–Picture Matching task of Study 1.



Nisbett, 2003) culture-specific mode of cognition proposal was not obtained.

Interim Conclusions: Language/ Culture-Specific Cognitive Processes versus Task-Specific Processes

The results overall show striking similarity across Chinese and German speakers. In both culture/language groups, taxonomic and thematic relations both proved to be important conceptual relations in the structure of object concepts. The results also show that German speakers are sensitive to similarity due to semantic features underlying classifier categories (presumably because they are often based on shape or simple semantic properties), but the magnitude of this effect was larger for Chinese speakers, which provides support for linguistic relativity. However, it would be an overstatement, given this effect, to state that the classifier categorization system in Chinese provides the speakers with a new way of organizing objects, for several reasons. First, the effect of classifier relations was much weaker than the effect of taxonomic or thematic relations when it was found (in similarity judgments and induction of a blank property). Second, the classifier effect found in the blank property induction easily diminished when participants were able to access some background knowledge. Third, classifier relations do not seem to be accessed in fast-speed, on-line cognitive processing. Taken together, a plausible conclusion seems to be that the classifier categorization system does not serve as a major organizer of the conceptual structure nor does it play a major role in the cognitive process in Chinese speakers. The language-specific classifier effect found among Chinese speakers is perhaps best characterized as a magnified sense of similarity through the habitual use of classifiers in association with the names of objects. This magnified similarity may be witnessed in cognitive activities that directly involve similarity, but does not extend to the entire range of cognition.

Concerning the issue of the taxonomic vs. thematic preference across Easterners and Westerners (Nisbett, 2003; Ji et al., 2004), conclusions were similar to those for the classifier effect. We did find that the Chinese participants gave thematic relations higher similarity ratings as well as higher likelihood judgment in inductive inferences of an unknown property than the Germans, which is consistent with the findings of Ji et al. However, this culture-specific preference toward the taxonomic or thematic relations was not observed in the categorization, inductive inference of a known property, and speeded word–picture matching task. It is important to note that thematic relations are important for Germans (Westerners) just as taxonomic relations are important for Chinese (Easterners). Thus, even though we found stronger preference for thematic relations in Chinese participants than in German participants in the similarity judgment task using a rating scale, this effect should at best be characterized as a quantitative rather than a qualitative difference.

In any case, the results emphasize the importance of examining the effect of given language-specific categories (as well as the effect of culture) in a range of cognitive tasks and systematically comparing the size of the effect to that of other conceptual relations, as the effect may be observed in one type of cognitive activity but not in others, and the effect may be limited in magnitude compared to other conceptual relations. The fact that we obtained the classifier effect in the inductive reasoning task with a blank property but not in the same task with a concrete property (sharing the same bacteria) should be particularly noted in this respect, as it suggests that the influence of linguistic categories deeply interacts with task-specific constraints, such as type of knowledge and cognitive processes required for the task and type of conceptual relations relevant for the task (e.g., Smith, Shafir, & Osherson, 1993).

Can the Classifier Effect in Chinese be Generalized to Other Classifier Languages?

Another aspect that should be considered in examining the effect of language is whether the effect of a target grammatical categorization holds across all languages having that grammatical function. In the realm of grammatical gender, Vigliocco and colleagues (Vigliocco et al., 2005) found an effect of grammatical gender in Italian speakers'categorization behavior, but this effect was not found in German speakers. Likewise, Sera et al. (2002) found an influence of grammatical gender in categorization in Spanish and French but not in German. Thus, we have grounds to expect that the relation between grammatical categorization system and cognition is not one-fold; whether and how a given grammatical system affects cognition depend on the structural and semantic nature of the system in a given language. In our case, it is important to determine whether the classifier effect we found in Chinese is also found in speakers of other classifier languages, in which the semantic function as well as the grammatical function of the classifier categories are not identical to those of Chinese.

For this purpose, we conducted a second series of studies comparing Chinese and Japanese speakers, with German speakers serving as a control group again. Before reporting this study, however, we briefly discuss how the classifier system differs across Japanese and Chinese.

Differences between the Japanese and Chinese Classifier Systems

Although Japanese and Chinese are both classifier languages, there are substantial differences between the two languages, especially with respect to the grammatical functions the classifiers play. Chinese classifiers must be used not only in numeral phrases (e.g., [numeral + classifier] table) but also in phrases with demonstratives (e.g., this [numeral + classifier] table). (The numeral after the demonstrative is often dropped, however. Thus, the most often heard demonstrative construction consists of the combination of "demonstrative + classifier + noun.") In contrast, Japanese classifiers are used only with numerals, and are not used in the construction with demonstratives. Furthermore, in Japanese, the classifier is used only when the mention of number is pragmatically important in the discourse. For example, when a Japanese speaker talks about her cat, unless the fact that she has one cat (i.e., one and not more than one) is important in the discourse, we would not say "Watashi (I) wa (Topic-marking particle) ichi (1) hiki (Classifier for small animals) no (Genetive) neko (cat) wo (Accusative-marking particle) katte (have)-imasu (Present Progressive/ State Aspect)" ("I have one cat"). Rather, she would simply say "watashi wa neko wo katte-imasu" ("I have cat") without specifying number. In contrast, in Chinese, the [numeral + classifier + noun] construction, especially with the numeral *yi* (one), is very often used even when the specification of number is not pragmatically important, and this [yi + classifier + noun] phrase serves like the English indefinite article. Thus, a Chinese speaker will say "wo (I) yang (grow) yi (1)- zhi (Classifier for small animals)-mao (cat)," meaning "I have *a* cat," in the context in which the specification of the number is not important in the discourse.

It is expected that this structural difference results in a much higher frequency of classifier use in Chinese than in Japanese. We confirmed this by comparing Chinese–Japanese translation texts using two sources. In the first source, Lammare (2009) compared the frequency of the classifier construction between Chinese and Japanese using the translation of the same original text (Chapter 4 of *Harry Potter and the Chamber of Secrets,* Rowling, 1999). She reported that the classifier construction appeared four times as frequently in the Chinese than in the Japanese translation (82 vs. 19 tokens).

In the second source, we compared a classic Japanese novel "Bocchan" (Master Daring) by Soseki Natsume (1964) using the Chinese–Japanese parallel corpus (Beijing Center for Japanese Studies, 2003). Here, if our linguistic analysis is correct, classifiers must be added in the process of translation from the original Japanese text to Chinese. In the original Japanese text, there were 111 classifier counts. In the Chinese translation, there were 405 counts. Thus, 294 classifier tokens were added through the translation process from Japanese to Chinese. On closer examination, there were 58 cases in which a classifier was used with "one" ("ichi") in the Japanese original. In the Chinese translation, there were 156 cases of "one" ("yi") with a classifier construction. When the number was "two" or "three," there were 21 classifier counts in Japanese and 53 in Chinese. In the Chinese translation, classifier counts were 175 in the "demonstrative + classifier + noun" construction [e.g., "Zhe (this) zhang (classifier) weirenzhuang (document)"]. However, in the original Japanese text, these were simple "demonstrative + noun" constructions without a classifier in all the cases. Here, classifiers are used roughly four times as frequently in Chinese as in Japanese.

STUDY 2: FURTHER EXAMINATION OF THE INFLUENCE OF CLASSIFIER CATEGORIES: COMPARISON OF CHINESE, JAPANESE, AND GERMAN SPEAKERS

Given that there was a language-specific classifier effect for Chinese speakers on similarity judgments and blank property induction in Study 1, we examined whether the classifier effect would be replicated for a larger set of stimuli in Chinese speakers and if it would also be found for speakers of another classifier language, Japanese. We compared Chinese, Japanese, and German speakers on similarity judgments and the two types of property induction tasks as well as on the speeded word–picture matching task.

In this study, we designed the stimuli in such a way that the classifier effect could be examined more finely than in Study 1 so that we could see whether the magnified similarity effect due to classifier relations is observed only for object pairs belonging to the same classifier category in the speakers' own language. We thus tested the classifier effect in three situations: (1) object pairs belonging to the same classifier class both in Chinese and Japanese, (2) object pairs belonging to the same classifier class in Chinese but not in Japanese, and (3) objects pairs belonging to the same classifier class in Japanese but not in Chinese. In addition, to determine if there would be a classifier effect over and above any effect of taxonomic category membership, we contrasted pairs in which two objects share both taxonomic category membership and classifier membership in both Chinese and Japanese (e.g., bed and table) to pairs in which the two objects shared only taxonomic category membership but not classifier category membership (e.g., bed and chair).

The first type of pairs was particularly important as it allowed us to test the influence of classifier categories in Japanese and Chinese speakers on the very same objects. Even if a stronger awareness of similarity was observed in Chinese speakers than in Japanese speakers for object pairs sharing the same classifier category membership in Chinese (but not in Japanese) but no stronger awareness of similarity was detected in Japanese speakers than in Chinese speakers for pairs sharing the same classifier membership in Japanese (but not in Chinese), it would be difficult to rule out the possibility that this difference could be due to some inherent properties of the stimuli. However, if the amplified similarity due to classifier category membership was observed in Chinese speakers but not in Japanese speakers for the same stimuli, the difference in the classifier effect could be attributed to the structural difference between Chinese and Japanese.

Stimulus set types 1–4 in Table 7.3 were used. Type 1 sets were used to contrast the object pairs from the same classifier class in both Chinese and Japanese (e.g., bone and tube: CH/JP CLS) with the pairs sharing no relation in either (e.g., bone and platter: Control). Type 2 sets, which were also used for Experiment 1, contrasted object pairs sharing classifier membership only in Chinese but not in Japanese (CH CLS) with pairs sharing no relation in either. Type 3 sets were used to contrast object pairs from the same classifier class in Japanese but not in Chinese (JP CL) with pairs sharing no relation in either language. Type 4 sets contrasted pairs in which the two objects shared both taxonomic category membership and classifier class membership in both languages (e.g., bed and table: CH/JP CLS + TAX) with pairs in which the two objects belonged to the same taxonomic category but to different classifier classes in both (e.g., bed and chair: TAX).

TABLE 7.3. Structure of the Stimuli Used for Study 2 with a Sample Set for Each Contrast Type

Contrasts		Same Classifier (CLS)					Control (CON)
		−TAX			+TAX		−TAX
	STANDARD	CH + JP	CH	JP	CH + JP		
Type 1 CH/JP CLS	Bone	Tube	—	—	—	—	Platter
Type 2 CHCLS	Flower	—	Cloud		—		Cup
Type 3 JPCLS	Bus	—	—	TV	—	—	Hat
Type 4 CH/JPCLS + Tax	Bed		—	—	Table	Chair	

The procedures for the similarity rating tasks and the two property induction tasks (induction about a blank property and induction about bacteria) were the same as those in Study 1. To test the language-specific classifier effect in fast-speed, automatic processing, we conducted two versions of the priming task. In Study 1, when the cue noun was presented as a bare word, no language-specific classifier effect was found. In fact, within Chinese speakers, the response latencies for the same-classifier items were not different from the control items. However, it is possible that when the noun is accompanied by the classifier, some online activation of classifier relations is evoked.

In a study examining the on-line influence of grammatical gender using a semantic substitution paradigm, Vigliocco, Vinson, Indefrey, Levelt, and Hellwig (2004) found that German speakers tended to produce substitution errors within the same gender category when producing phrases with determiners marked for gender, but that this gender preservation effect disappeared when they produced bare nouns or phrases with determiners not marked for gender (i.e., a phrase with a indefinite determiner plus noun). Similarly, classifier relations may be activated on-line when a noun is presented with a classifier but not when it is presented in isolation (cf. Gao and Malt, 2009). In Study 2, we thus conducted the word–picture matching task in a version in which the cue word was presented in the classifier phrase [e.g., "yi (one) ge (classifier) pingguo (apple)"], in addition to the version identical to that in Study 1 (i.e., the bare noun version). The same stimuli sets as in Experiments 1–3 were used for this study. Furthermore, additional control items were included: In Type 1 through 3 sets, for each target object, an object that was taxonomically related but belonged to a different classifier category was added; in Type 4, an object having neither a taxonomic nor a classifier relation to the target was added. These manipulations were done to ascertain whether the semantic priming procedure would work throughout Type 1–4 sets. If this was the case, delayed rejection of taxonomically related pairs should occur as compared to unrelated control pairs. Thus, in case no delay due to the same classifier relation should be found, we would be able to determine whether the null result was due to the absence of the classifier effect or to methodological problems.

Given the results of previous research, we might expect a classifier priming effect to occur in the phrase condition but not in the bare noun condition. Here, we tested only Chinese and Japanese speakers, as German has no classifiers. We compared the Chinese and Japanese data here with the German data in the bare noun version.

Results of Study 2

Similarity Judgments As in Study 1, we present data as adjusted (difference) scores so that readers could directly see the effect of classifiers compared to control (see Fig. 7.5). Replicating the results of Study 1, object pairs belonging to the same classifier category, either in Chinese or Japanese, or both, were rated more similar than the control pairs (object pairs sharing no relation) by not only Chinese or Japanese speakers but also by German speakers. We thus provide additional evidence that there is inherent similarity among objects belonging to the same classifier category, and this similarity is detectable even by nonspeakers of a classifier language. However, when the objects in a pair shared a taxonomic relation, an additional classifier relation did not increase similarity, as there was no statistically reliable difference between the CH/JP CLS + TAX pairs and the Tax (-CH/JP CLS) pairs in the rated similarity.

The language-specific classifier effect observed among Chinese speakers in Study 1 was replicated, but this effect was not found in Japanese speakers. As in Study 1, we first compared the size of the effect of classifier relations with that of taxonomic relations, and whether there was any cross-cultural difference in this respect. In all three language groups, object pairs sharing taxonomic relations received much higher similarity ratings than object pairs sharing classifier relations (but did not share taxonomic relations) in the same degree.

We then conducted a preplaned contrast on the means of the object pairs belonging to the same classifier category in Chinese (i.e., the average of the CHJP CLS and CH CLS) and the means of the corresponding controls (i.e., the average of the Control items in Type 2 and Type 4 sets) and tested if Chinese speakers showed a larger effect than German speakers on this contrast[2]; there was a highly significant effect for Language on this contrast, revealing that Chinese speakers rated the pairs belonging to the same classifier category in Chinese higher than German speakers. In contrast, when we contrasted the mean similarity ratings for the object pairs belonging to the same Japanese classifier categories (i.e., the average of the CHJP CLS and JP CLS) and the corresponding controls, no effect for Language was found on this contrast, suggesting that Japanese speakers do not show a language-specific classifier effect as compared to German speakers. Importantly, when we compared Chinese

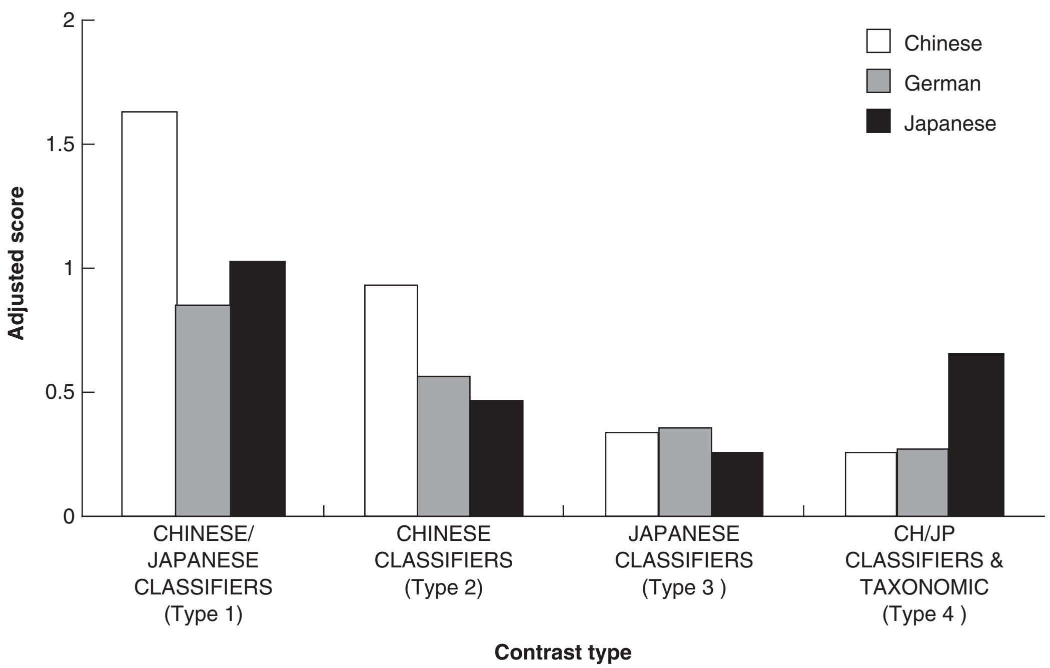


FIGURE 7.5. Adjusted scores for each contrast type in each language in Similarity Judgment of Study 2.

and Japanese speakers on the object pairs belonging to the same-classifier categories on both languages, Chinese speakers gave higher ratings for the same-classifier pairs than Japanese speakers. When object pairs were taxonomically related, sharing the classifier categories in addition did not increase similarity over and above the similarity due to the taxonomic relations in any of the three language groups.

To recapitulate the results, we found the amplified classifier similarity effect in Chinese speakers (in contrast to German speakers), but we did not find the same effect for Japanese here. Furthermore, when we directly compared Chinese speakers and Japanese speakers on the object pairs that shared classifier membership in both languages, Chinese speakers rated these items significantly higher than Japanese speakers, again suggesting that the language-specific amplified classifier effect is found only in Chinese but not in Japanese speakers.

Induction of a Blank Property As can be seen in Figure 7.6, all three groups gave significantly higher ratings for object pairs belonging to the same classifier category (only in Chinese, only in Japanese, or in both Chinese and Japanese) than for their corresponding control pairs. However, when both the same-classifier item and the control item belonged to the same taxonomic class (Stimuli Set Type 4), no effect of classifier membership was obtained over and above the effect due to taxonomic relations.

When the language-specific classifier effect was examined, it was again found for Chinese speakers but not for Japanese. In other words, Chinese speakers rated the likelihood of two objects sharing an unspecified property higher than speakers of a nonclassifier language (German) when the objects belonged to the same classifier category in Chinese (CHJP CLS and CH CLS). In contrast, Japanese speakers did not show a language-specific magnified classifier effect for the object pairs from the same Japanese classifier category (CHJP CLS and JP

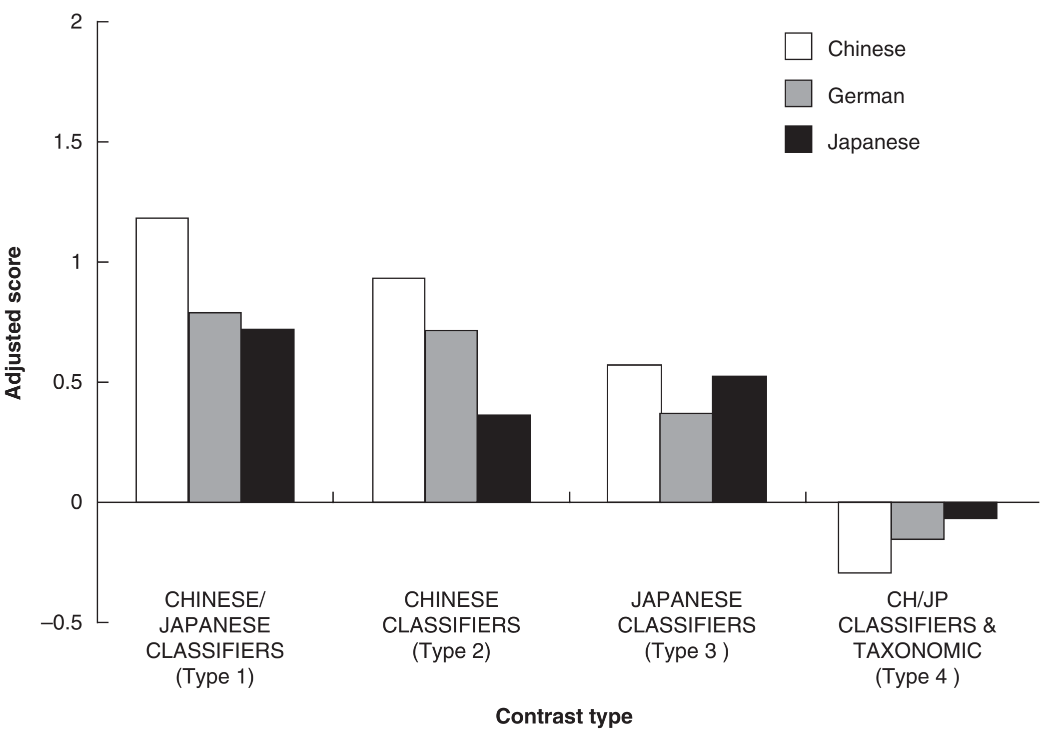


FIGURE 7.6. Adjusted scores for property induction (about blank property) for each contrast type in each language in Study 2.

CLS) greater than German speakers in this task. When Chinese and Japanese speakers were compared on the object pairs belonging to the same classifier category in both languages, again, Chinese speakers gave higher likelihood ratings than Japanese speakers on these pairs.

Induction of Bacteria Consistent with the results of blank property induction, all three language groups gave significantly higher ratings for object pairs belonging to the same classifier category than for their corresponding control pairs (see Fig. 7.7). However, unlike the case with the blank property induction, but consistent with the results of the same task in Study 1, there was no language-specific classifier effect in Chinese speakers here: Chinese speakers did not give any higher ratings than Japanese or German speakers for the object pairs belonging to the same classifier category in Chinese. Again, Japanese speakers did not rate the object pairs belonging to the same classifier category in Japanese any higher than Chinese or German speakers. Also consistent with the results from the similarity judgment task and the blank property induction task, when two objects were related taxonomically, classifier category membership did not increase the likelihood ratings for the same bacteria being found in the two objects.

Word–Picture Priming We first report the results from the version in which the cue noun was presented on its own. In this version, the prime was presented in a word in the bare form, and the participants were asked to judge whether the target picture matched the preceding word prime. Figure 7.8 shows the adjusted response times (difference between the mean latency for the same-classifier items and the control items in the corresponding sets) for the same-classifier items in each of Type 1–4 sets. In the far right of the graph, we also included the adjusted times for the

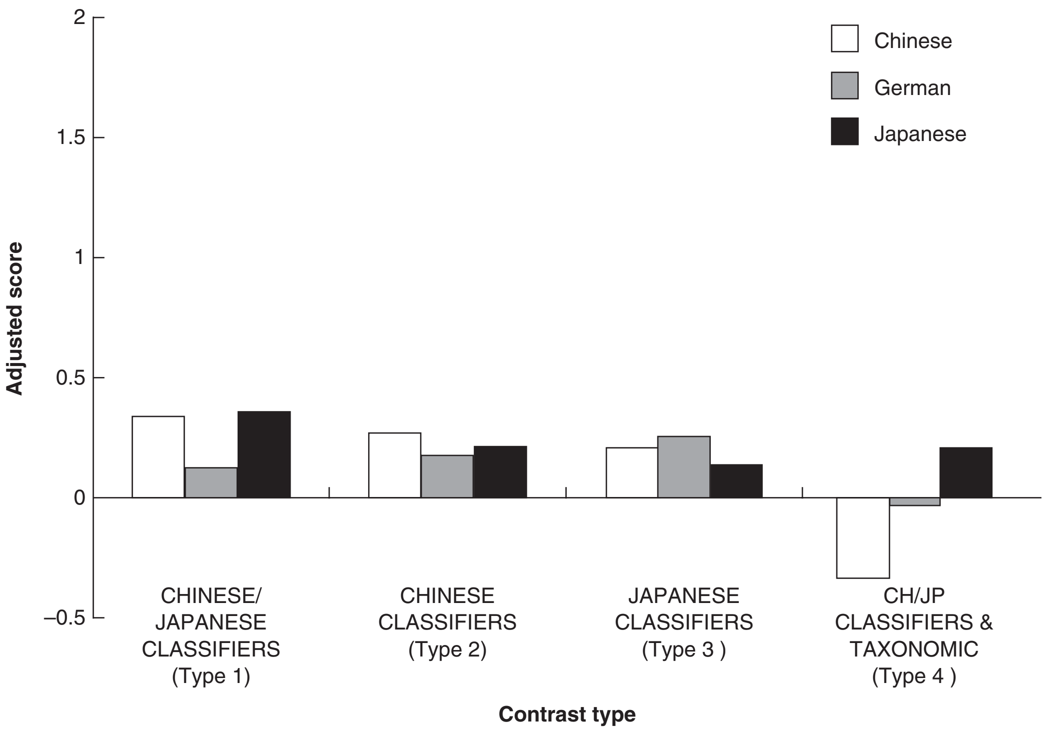


FIGURE 7.7. Adjusted scores for property induction (on the "carrying the same bacteria" question) for each contrast type in each language in Study 2.

taxonomic items, which was obtained by subtracting the latencies for the control items from the taxonomic items (averaged across Type 1–3 sets), so that readers can compare the effect due to classifier relations with the effect due to taxonomic relations. As clearly seen in Figure 7.8, we found that participants from all three language groups took significantly longer to respond to taxonomically related items than to item pairs from the same classifier category (that were not taxonomically related) and control items. The object pairs from the same classifier category (i.e., the classifier items in Type 1–3 sets together) caused longer response latencies relative to pairs from different classifier categories (the control items in Type 1–3 sets together) in all three language groups, again demonstrating that the task was sensitive to important conceptual relations in people's on-line cognitive processing. However, consistent with the results of Study 1, there were no language-specific classifier effects either on the Chinese same-classifer pairs (the classifier items in Type 1 and 2 sets together) or on the Japanese same-classifier pairs (the classifier items in Type 1 and 3 sets together) when we compared the classifier effect with German speakers. Here, when we compared the classifier effect in Chinese and Japanese speakers on the pairs sharing the classifier membership in both languages (Type 1 sets), no difference was found across the two language groups.

Given these results together with the results of the parallel word–picture matching task in Study 1, it is likely that classifier relations are not automatically activated in processing a noun without a classifier. The fact that the delay due to classifier membership was observed in all of the three languages in roughly the same magnitude indicates that the effect of classifier category membership is due to an inherent similarity underlying classifier categories. Perhaps the cognitive influence of classifiers may be seen not so much as a long-term influence on the representation of objects per se but as a temporary shift in the construal of the referent in the way that parallels English speakers' shift of construals when the same noun is marked as a count noun or as a mass noun (e.g., "Jim had a few chocolates" vs. "Jim had some chocolate") (e.g., Middleton, Wisniewski, Trindel, & Imai, 2004). If so, the classifier relations may cause a delay when the classifier is explicitly specified with the noun.

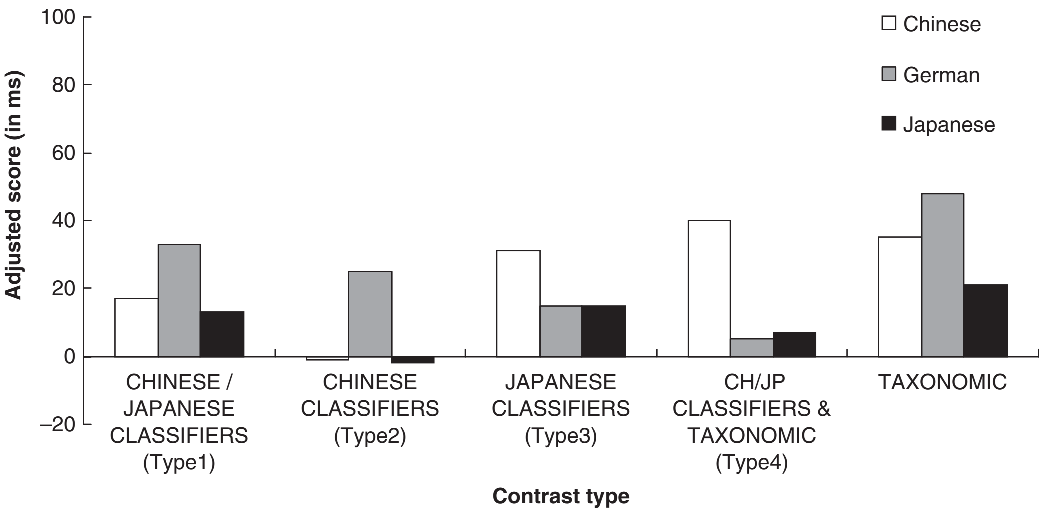


FIGURE 7.8. Adjusted scores for response latencies for each language in the Bare-Noun condition of the Word–Picture Matching task of Study 2.

Note: "Taxonomic" represents the adjusted score for the taxonomic items, which was obtained by subtracting the control (unrelated) items from the mean response time for the taxonomic items in the Type 1-3 sets.

As can be seen in Figure 7.9, the pattern of the results in this *phrase priming task* was the same as in the previous priming task: Although the cue/target pairs from the same classifier categories delayed responses relative to the cue/target pairs from different classifier categories in both Chinese and Japanese speakers, neither the Chinese same-classifier pairs nor the Japanese same-classifier pairs caused language-specific classifier effects on response latencies when compared to the German speakers' latencies.

Summary of Study 2

The results of the similarity judgments and two property induction tasks in Study 2 converged with the results we reported in Study 1 for Chinese speakers. Two objects from the same classifier categories, when they did not share any other conceptual relations, were construed as more similar than two objects from different classifier categories, and this similarity was detected by people whose language did not have a classifier system and could serve as a basis of inductive reasoning, especially when people could not use much background knowledge. However, in none of the tasks tested was the magnified similarity effect due to classifiers observed in Japanese speakers.

Consistent with the results of Study 1, the word–picture matching tasks of Study 2 revealed the sensitivity to the similarity underlying classifier categories in speakers of all three language groups. However, the magnitude did not differ between speakers from classifier languages and a nonclassifier language, regardless of whether the classifier categories were explicitly invoked (using nouns in the classifier phrase) or not (using bare noun). Thus, it seems that the language-specific classifier effect is not evident in fast, automatic cognitive processes.

Overall, although the language-specific classifier effect was not revealed in fast, automatic processing, some influence of the classifier system on the conceptual structure of everyday objects was found in Chinese speakers in non-speeded similarity judgments and blank property induction tasks. However, this effect was not identified in Japanese speakers. Thus, finding a language-specific influence of a given linguistic categorization system in one language cannot be automatically generalized to other languages that have the same linguistic categorization system.

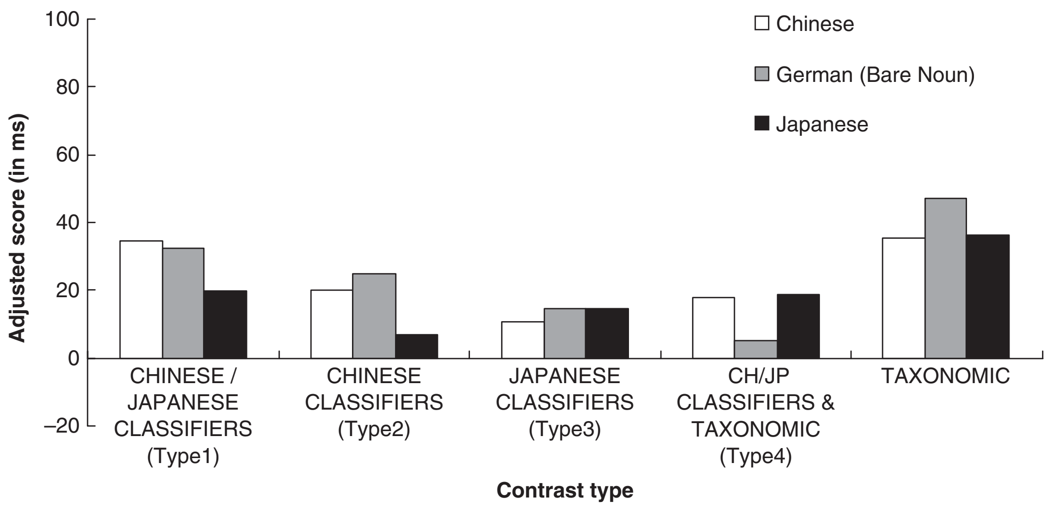


FIGURE 7.9. Adjusted scores for response latencies for each language in the Phrase condition of the Word–Picture Matching task of Study 2.

Note: "Taxonomic" represents the adjusted score for the taxonomic items, which was obtained by subtracting the control (unrelated) items from the mean response time for the taxonomic items in the Type 1-3 sets.

COGNITIVE CONSEQUENCES OF CLASSIFIER CATEGORIES

In summary, what are the cognitive consequences of having such linguistic categories? We set up the following criteria in evaluating the influence of the classifier system: (1) whether the conceptual relation underlying the target linguistic categories serves as a basis not only for similarity judgments but also for inductive reasoning, (2) whether the conceptual relation is evoked automatically, and (3) whether the effect of the conceptual relation is comparable to that of other major conceptual relations such as taxonomic and thematic relations. We then tested three possible scenarios: (1) classifier categories function as the most dominant or one of the major organizers of our concepts and categories (cf. Lakoff, 1987; Zhang & Schmitt, 1998), (2) classifiers heighten the speakers' sense of similarity but the influence of the classifier system is not qualified to be considered as a major organizer of the speakers' conceptual structure, in light of the above criteria, as the classifier influence is not pervasive enough in the context of a broad range of cognitive processes and/or the magnitude of influence is not large enough compared to that of other major conceptual relations, and (3) classifiers are "frozen," linguistic conventions without any cognitive impact.

We found some support for the second but not the first scenario (nor the third) in Chinese speakers. The Chinese participants gave higher ratings to the same-classifier pairs than the German participants both in similarity judgments and inductive inference of a blank property, which suggests that classifier categories have some impact on Chinese speakers' conceptual structure of everyday objects. However, it would be an overstatement to say that classifier categories serve as an additional or alternative basis for *organizing* our concepts, because the magnitude of the classifier effect was limited compared to the other major relations that organize concepts. In Study 1, we demonstrated that taxonomic and thematic relations are important organizers of people's conceptual structures, regardless of whether they speak a classifier language. In contrast, the impact of classifiers was much smaller in magnitude when it was found at all, and the language-specific influence of classifier categories was limited to just two tasks, i.e., similarity judgment and inductive reasoning of blank properties. Study 2 replicated the results of Study 1 for Chinese speakers. However, the language-specific classifier effect was not found in Japanese speakers in any of the tasks, including those in which the classifier effect was observed in Chinese speakers.

Thus, the pattern of the results suggests that the classifier effect found among Chinese speakers is best characterized as a magnified sense of similarity developed through the *habitual* use of classifiers in association with the names of objects. The fact that the language-specific classifier effect was not found in Japanese speakers is in accord with this interpretation. In Japanese, perhaps classifiers are not used frequently enough to result in magnified sensitivity in the semantic features underlying classifier categories. In other words, it does not seem unreasonable to suspect that the language-specific classifier effect in Chinese speakers arises from the speakers' experience of observing conceptually very different objects being marked with the same grammatical morpheme. In this sense, even though the similarity underlying classifier categories is detectable by speakers of a nonclassifier language, the nature of the similarity is very different from the similarity underlying taxonomic or thematic relations.

How are classifier categories different from taxonomic or thematic categories, which apparently play a much more important role in organizing concepts? As discussed earlier, classifier categories are held together only by a single semantic feature or a combination of at most a few semantic features. This characteristic naturally leads to the consequence that category members do not have much in common, which probably makes classifier categories, at best, only weak conceptual categories. Seen this way, it does not seem unreasonable that the classifier effect was observed in similarity judgments and blank property induction but not in other tasks. People are very versatile in perceiving similarity, and even the commonality of a single feature can significantly affect the construal of similarity among objects. On the other hand, it is reasonable that people (including both speakers of a classifier language and a nonclassifier language) did not make inductive generalization of a property based on the kind of similarity that underlies classifier categories, when they could access background knowledge.

IMPLICATIONS FOR LANGUAGE AND THOUGHT ISSUES

The studies reported in this chapter emphasize the importance of putting the effect of linguistic categories in context by comparing other kinds of conceptual relations and considering to what extent the linguistic relativity effect, if there is any, is meaningful for the global organization of human concepts. As we reviewed earlier, Zhang and Schmitt (1998) found that Chinese speakers rated objects from the same classifier categories more similar than English speakers and interpreted this result as evidence for linguistic relativity. Our own research confirmed this effect, but also found that this effect is much smaller compared to other kinds of conceptual relations, such as taxonomic and thematic relations.

This conclusion is also important for interpreting the results of studies examining the Whorfian effect of grammatical categories other than the classifier system. For example, previous studies examining the influence of grammatical gender reported some evidence for linguistic relativity, as reviewed earlier (e.g., Boroditsky et al. 2003; Konishi, 1993; Gomez-Imbert, 1996; Sera et al., 2002). In general, if a cross-linguistic difference is found between a language having the grammatical gender system and a language without the gender system in any task, be it in similarity judgments, categorization, or attribution of male-like/female-like properties, it is taken as evidence for linguistic relativity. However, it is worthwhile to rethink how meaningful the effect is for the global structure of the speakers' object concepts. An interesting question in this light is whether the grammatical gender of an animal name influences inductive or deductive inference of a biological gender-specific property. For example, if a given animal's grammatical gender is feminine but the animal's biological sex is unknown, are the speakers more likely to infer that the animal has a biological female property (e.g., having a female hormone) than when the animal's grammatical gender is masculine? If so, the linguistic influence of grammatical gender categories would seem to have substantial meaning for the speakers' conceptual structures.

The research reported in this chapter suggests that the influence of linguistic categories deeply interacts with task-specific cognitive constraints and availability of background knowledge. Second, it also highlighted the importance of examining the influence of linguistic categories not in light of *whether* there is one, but in light of *how large* the influence is in a broad range of cognitive processes, and *how it is related to* other major conceptual relations underlying our conceptual structure. Relevant to this point, the fact that German participants judged objects belonging to the same classifier category to be more similar than unrelated objects supports the notion that grammatical categories are *motivated* (but of course not determined) *by* universally shared cognitive and perceptual experience (e.g., Zubin & Köpcke, 1986). In this sense, our results are not incongruent with the view that grammatical categories are a reflection of cognitive categories (e.g., Lakoff, 1987). At the same time, our results suggest that grammatical categories such as classifier categories do not function as a major organizer of our concepts to the same extent that taxonomic and thematic categories do. More importantly, the two aspects of our findings—that similarity underlying classifier categories can be detected by German speakers on the one hand and that similarity due to classifier relations is magnified by Chinese speakers on the other hand—cogently suggest that the relation between language and thought is not unidirectional: Linguistic categories reflect universally perceived commonalities in the world, but at the same time they modify universally perceived similarities (see Imai & Mazuka, 2007, for a relevant discussion).

One issue that also warrants some discussion is whether the results we obtained from the two series of studies permit us to draw the conclusion that the classifier system influences Chinese speakers' conceptual structure and cognitive processes, and hence can be interpreted as evidence for the Whorfian hypothesis, given the lack of the classifier effect in Chinese speakers in the on-line tasks. Some researchers might be concerned that Chinese speakers' similarity ratings for same-classifier pairs were higher than those of German speakers because they were aware that the two objects were associated with the same classifier category, which is simple reflection of Chinese speakers' conscious strategy adopted for carrying out the task and hence cannot be considered as the "true" reflection of their cognitive processes.

We set up three points to be considered in evaluating the influence of a given linguistic categorization system, one of which is whether the influence due to the linguistic categorization system is found in unconscious, automatic processes. In this respect, the effect of the classifier system is indeed not as large or important as taxonomic relations or thematic relations. However, in our view, this in itself should not mean that the amplified classifier effect found in Chinese speakers is an experimental artifact and hence is not "real." First, methodologically, we do not think that this conscious "give-higher-ratings-for-same-classifier-object" strategy is the cause for the amplified similarity effect identified in Chinese speakers. If participants had been forced to choose one of two objects, one from the same classifier category and the other with no such relation, and if they could not find reasonable similarity in either object, they might have resorted to this strategy. In the similarity judgment task in our research, however, no direct comparison of the similarity of the same-classifier object and the control object was involved. Furthermore, in the similarity judgment task as well as in the inductive reasoning task, the classifier was not explicitly referred to when presenting the stimuli, and the classifier relations were not the only kind of relation among the objects in the stimulus set, as our stimuli also included taxonomically related objects. Thus, it is not likely that Chinese speakers consciously thought of classifiers while undertaking the task. Even if the Chinese-speaking participants had used this strategy consciously, they did so spontaneously in situations in which this was not required. Similarity and inductive reasoning are two core processes for human cognitive activities. In our view, if participants had thought about classifiers even when no invocation of the classifier was necessary and spontaneously utilized this knowledge when engaging these activities, this would in itself suggest that classifiers indeed affect speakers' cognitive process.

Another aspect that should be considered in examining the effect of language is whether the effect of a target grammatical categorization holds across all languages having that grammatical function. In the research presented, we found the classifier effect among Chinese speakers but not in Japanese speakers. This parallels the finding that grammatical gender influences similarity and categorization in Italian or Spanish but not in German (Vigliocco et al., 2005; Sera et al., 2002).

In closing our chapter, we suggest that the simple Whorfian-vs.-non-Whorfian dichotomy does not deepen our understanding of the nature of our concepts and cognitive processes very much, given the complexity of the interactions among many factors that affect the structure of our concepts and cognitive processes. In future research, we clearly must go beyond simply seeking evidence for the Whorfian hypothesis. What is important, then, is to clarify *how*, rather than *whether*, language-specific categories, be they grammatical or lexical, affect our concepts, categories, and cognitive processes. It is particularly important to specify how the effect interacts with our universal cognitive biases, the structure of the world, the constraints placed by the task or cognitive activity at hand (e.g., what type of information or knowledge is most relevant for the inference), and, finally, the language-specific characteristics of the given linguistic categorization system.

ACKNOWLEDGMENTS

We thank Barbara Malt and Phil Wolff for insightful comments on an earlier version of this chapter. We are also indebted to Zhou Xiaolin, Elsbeth Stern, Sachiko Saitoh, Lennart Schalk, Shu Hua, Zhu Liqi, and Huangfu Gang for help with data collection and discussion. This work was in part supported by a Ministry of Education grant-in-aid for Scientific Research (#15300088 and #18300089) and research grants from Keio University (Keio Gijuku Academic Development Funds and Keio Gijuku Mori Memorial Research Fund) to M. Imai and by fellowships from the Japan Society of the Promotion of Science (JSPS) and Friedrich-Ebert-Stiftung awarded to H. Saalbach.

Notes

1. Here,we consider only so-called "numeral classifiers," and do not include what is sometimes called "noun classes," in which nouns are obligatorily classified into a small number of noun classes as in the case with Dyrbal (Dixson, 1986; see also Lakoff, 1987).
2. We did not conduct ANOVA analyses directly using the difference scores, because it is not statistically desirable. Instead, we conducted a repeated measure ANOVA including the classifier conditions and corresponding control conditions from the four sets (Type 1–4 sets) and then tested specific effects by planned contrasts.

References

Adams, K., & Conklin, N. F. (1973). *Toward a theory of natural classification.* Paper presented at the 9th Regional meeting of the Chicago linguistic society.

Aikhenvald, A. (2000). *Classifiers: A typology of noun categorization devices.* New York: Oxford University Press.

Allan, K. (1977). Classifiers. *Language, 53*(2), 285–311.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Bassok, M., & Medin, D. L. (1997). Birds of a feather flock together: Similarity judgments with semantically rich stimuli. *Journal of Memory and Language, 36,* 311–336.

Beijing Center for Japanese Studies (北 京日本学研究) (2003). *Chinese-Japanese parallel corpus.*

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax and semantics. In *Language in mind: Advances in the study of language and thought.* (pp. 61–79). Cambridge, MA: MIT Press.

Bowerman, M., & Levinson, S. C. (2001). *Language acquisition and conceptual development.* Cambridge, UK: Cambridge University Press.

Craig, C. G. (1986). Jacaltec noun classifiers: A study in language and culture. In C. Craig (Ed.), *Noun classes and categorization* (pp. 263–293). Philadelphia: John Benjamins.

Croft, W. (1994). Semantic universals in classifier systems. *Word, 45,* 145–171.

Denny, J. P. (1979). Semantic analysis of selected Japanese numeral classifiers for units. *Linguistics, 17,* 317–335.

Denny, J. P. (1986). The semantic role of noun classifiers. In C. Craig (Ed.), *Noun classes and categorization*. Amsterdam and Philadelphia: John Benjamins.

Dixon, R. M. W. (1986). Noun classes and noun classification in typological perspective. In C. Craig (Ed.), *Noun classes and categorization*. Amsterdam: John Benjamins.

Downing, P. (1996). *Numeral classifier systems: the case of Japanese*. Philadelphia: John Benjamins.

Gao, Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes.* In press.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*(3), 183–209.

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

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

Gomez-Imbert, E. (1996). When animals become "rounded" and "feminine": Conceptual categories and linguistic classification in a multilingual setting. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity*. Cambridge, UK: Cambridge University Press.

Grinevald, C. (2000). A morphosyntactic typology of classifiers. In G. Senft (Ed.), *Systems of nominal classification* (pp. 50–93). Cambridge, UK: Cambridge University Press.

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

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review, 98*(3), 377–389.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meanings: The role of shape similarity in early acquisition. *Cognitive Development, 9*, pp.45–75.

Imai, M., & Mazuka, R. (2003). Reevaluating linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind* (pp. 429–464). Cambridge, MA: MIT Press.

Imai, M., & Mazuka, R. (2007). Revisiting language universals and linguistic relativity: language-relative construal of individuation constrained by universal ontology. *Cognitive Science, 31,* 385–414.

Ji, L.-J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality & Social Psychology, 87*(1), 57–65.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*(5), 1083–1105.

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research, 22*(5), 519–534.

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

Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General, 130*(1), 3–28.

Lucy, J. A. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. New York: Cambridge University Press.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist, 44,* 1469–1481.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227–234.

Middleton, E. L., Wisniewski, E. J., Trindel, K. A., & Imai, M. (2004). Separating the chaff from the oats: Evidence for a conceptual distinction between count noun and mass noun aggregates. *Journal of Memory and Language, 50,* 371–394.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Natsume, S. (1964). *Bocchan* 'Master Daring.' Tokyo: Kaisei-sha Publisher.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General,* 106, 226–254.

Nisbett, R. E. (2003). *The geography of thought: How Chinese and Westerners think differently ... and why*. New York: Free Press.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291–310.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*(2), 185–200.

Rowling, J. K. (1999). *Harry Potter and the Chamber of Secrets,* London: Bloomsbury Publishing.

Saalbach, H., & Imai, M. (2007). The scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General, 136,* 485–501.

Senft, G. (1996). *Classificatory particles in Kilivila.* New York: Oxford University Press.

Senft, G. (2000). What do we really know about nominal classification systems? In G. Senft (Ed.), *Systems of nominal classification* (pp. 11–50). Cambridge, UK: Cambridge University Press.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodriguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131*(3), 377–397.

Slobin, D. I. (1987). *Thinking for speaking.* Paper presented at the Thirteenth Annual Meeting of the Berkeley Linguistic Society.

Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations. A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology, 28,* 249–257.

Smith, E.E., Shafir, E., & Osherson, D.N. (1993). Similarity, plausibility, and judgments of probability. *Cognition, 49,* 67–96.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301–306.

Vigliocco, G., Vinson, D., Indefrey, P., Levelt, W., & Hellwig, F. (2004). The interplay between meaning and syntax in language production. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 483–497.

Vigliocco, G., Vinson, D., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General, 134* (4), 501–520.

Waxman, S. R., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*(2), 139–156.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. New York: MIT Press.

Wisniewski, E. J., & Bassok, M. (1999). What makes a man similar to a tie? Stimulus compatibility with comparison and integration. *Cognitive Psychology, 39*(3), 208–238.

Yokosawa, K., & Imai, M. (1997). *Conceptual interference on object verification: Effect of superordinate vs. thematic associative relations.* Paper presented at the Meeting of the Asssociation for Research in Vision and Ophthalmology, Fort Lauderdale, FL.

Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied, 4*(4), 375–385.

Zubin, D. A., & Köpcke, K. M. (1986). Gender and folk taxonomy: The indexical relation between grammatical and lexical categorization. In C. Craig (Ed.), *Noun classification and categorization* (pp. 139–180). Philadelphia: John Benjamins.

8 LANGUAGE AND THOUGHT
Which Side Are You on, Anyway?

Terry Regier, Paul Kay, Aubrey L. Gilbert & Richard B. Ivry

The debate over language and thought has traditionally been framed by two opposing stances: "universalist" and "relativist." The universalist view holds that language is shaped by universals of human cognition; on this view, languages make semantic distinctions drawn from a limited palette of universally available options—and when languages do differ semantically, those linguistic differences do not affect cognition. The relativist view, in contrast, often attributed to Whorf (1956), holds that semantic distinctions are determined primarily by largely arbitrary linguistic convention, so that languages are free to vary widely—and that such linguistic differences *do* affect cognition. Thus, on the relativist view, rather than universals of thought shaping language, it is language that shapes thought, in a manner that varies with little constraint across languages.

These opposed stances reflect broader issues of the universality or malleability of human nature, and perhaps for this reason the opposition seems a natural way to conceptually organize the debate. Over the years, consensus has oscillated between these two poles. Most recently, there are some signs that the field has begun to swing toward relativism (Roberson, Davies, & Davidoff, 2000; Lucy, 1992; Gordon, 2004; Roberson, Davidoff, Davies, & Shapiro, 2005).

Here, we respond to this development—but not by a counterswing back to universalism. Instead, we argue that the oppositional framing itself should be jettisonedaltogether, since it has outlived its usefulness and is an obstacle to understanding—despite its apparent simplicity and naturalness. Fundamentally, the problem is that this framing is too coarse-grained. One instance of this coarseness is that the framing bundles together two separate questions:

1. Are semantic distinctions in languages determined by largely arbitrary linguistic convention?
2. Do semantic differences cause corresponding cognitive or perceptual differences in speakers of different languages?

The traditional framing implicitly assumes that the two questions will receive the same answer: either both "yes" (relativist) or both "no" (universalist). A relativist holds that there is no universal vocabulary of thought and perception, so languages are free to vary largely arbitrarily in their semantic partitioning of the world (yes to question 1), and these linguistic differences can leave their imprint on thought and perception (yes to question 2). A universalist, in contrast, holds that there *is* a universal vocabulary of thought and perception, so languages are constrained to reflect it (no to question 1) and cannot alter it (no to question 2).

However, as we will see, available data on color naming and cognition support a picture that is more interestingly differentiated than either of these traditionally opposed positions. We will show that there are clear universal tendencies of color naming, but that linguistic convention may nonetheless play some role in determining category extension—a hedged universalist answer to question 1. In contrast, we argue for a hedged relativist answer to question 2. Our recent research shows that language does affect perception—but primarily in the right half of the visual field, and much less if at all in the left half. This pattern is suggested by the functional organization of the brain, but is unanticipated by the framing of the debate. Thus, the oppositional framing oversimplifies matters by suggesting simple yes-or-no answers to questions that demand more detailed responses. It also oversimplifies matters by collapsing the distinction between the two questions, obscuring the fact that the answers do not match.

Empirical research on the language-and-thought question has concentrated heavily on color naming and color cognition and perception, and we will do the same in this chapter. We first review the debate over color naming and cognition, highlighting the apparent conflation of questions 1 and 2 in that debate. We suggest how some recent findings help to distinguish these questions, and lead to the conclusion of universal tendencies in naming, coupled with Whorfian effects of language on thought. We next show that Whorfian effects of language on perception may be dominant in the right visual field—the experiments we report here again concern color, but only incidentally. The "Whorf on the right" suggestion is a general one, and we expect it to hold for other semantic domains as well. We conclude with a discussion of what these findings mean for the language and thought debate generally, and what useful role, if any, the traditional framing of the debate may play in the future.

A BRIEF HISTORY OF THE LANGUAGE AND THOUGHT WAR, AS FOUGHT ON THE BATTLEFIELD OF COLOR

In the mid-nineteenth century, various scholars, notably William Gladstone (1858) and Lazarus Geiger (1880), noted that the speakers of ancient written languages did not name colors as precisely and consistently—as they saw it—as the speakers of modern European languages. They proposed a universal evolutionary sequence in which color vocabulary evolves in tandem with an assumed biological evolution of the color sense. Although some nineteenth-century scholars, notably Hugo Magnus (1877, 1880), rejected the idea that lexical evolution in the color domain necessarily mirrors a corresponding perceptual evolution, the notion of a universal evolutionary sequence in color nomenclature dominated nineteenth-century scientific thought (Rivers, 1901). As the twentieth century progressed and anthropologists and anthropological linguists increasingly encountered languages and cultures that appeared to be as systematic as the familiar European ones, the notion of cultural and linguistic relativity began to take hold, in opposition to the traditional geist of universal evolutionary progress. By mid-twentieth century, relativism held full sway in both linguistics and anthropology—at least in North America, where the two subjects were fully intertwined. In the color domain, this sentiment took the form of declarations such as that "there is no such thing as a natural division of the spectrum. Each culture has taken the spectral continuum and has divided it upon a basis which is quite arbitrary" (Ray, 1952: 252, quoted in Berlin & Kay, 1969: 159). Similar ideas may also be found elsewhere (Gleason, 1961; Ray, 1953; Conklin, 1955; Nida, 1959; Bohanon, 1963; Krauss, 1968). Psychologists also were inclined to accept the relativist view in regard to color, most notably in Brown and Lenneberg's (1954) finding that colors more readily coded in language were easier to remember (see also Lantz & Stefflre, 1964; Stefflre, Castillo Vales, & Morley, 1966).

Against this background, the comparative color-naming survey of Berlin and Kay (1969) and the field experiments on color cognition of the Dugum Dani of Eleanor Rosch (Heider, 1972; Heider & Olivier, 1972) started to swing the pendulum back to universalism. In a survey of 98 languages (only 20 of which were directly assessed, the others being taken from the literature), Berlin and Kay (1969) found something rather similar to the universal evolutionary sequence originally posited by Geiger. They posited universal focal colors, corresponding to the best examples of English *black, white, red, yellow, green,* and *blue* (or corresponding terms in other languages), and explained the different category boundaries in different languages as resulting from different groupings of these universal foci. Further, they proposed that in the course of its history a language breaks up the categories that group several universal foci in a partially predictable manner corresponding roughly to the sequence Geiger had postulated. Rosch found that Dani speakers, with only two basic color terms in their language, reacted much like English speakers regarding English—and by inference universal—focal colors in several tests of memory and learning. The findings of Berlin and Kay (1969) challenged the view typified by the citation from Ray given above and Rosch's challenged the tradition stemming from the Brown and Lenneberg (1954) experiments.

Although there have always been critiques of the Berlin/Kay/Rosch results (e.g., Hickerson, 1971; Lucy & Shweder, 1988), the recent swing back toward the relativist pole was given a major thrust by the work of Debi Roberson and her associates on the Berinmo of Papua New Guinea (Roberson, Davies, & Davidoff, 2002; Roberson et al., 2000). Roberson and associates found that Berinmo color categories have boundaries that differ from those of English—and that these cross-linguistically varying boundaries seem to affect color cognition. They focused on two Berinmo color categories that are roughly comparable to, but have different boundaries from, English yellow and green. They showed that Berinmo speakers exhibit "categorical perception"[1] of color at the boundary between these two Berinmo color categories—but not at the English yellow/green boundary. English speakers showed the opposite pattern. This confirms earlier findings by Kay and Kempton (1984), who found that English speakers exhibit a categorical perception effect at the English green/blue boundary, whereas speakers of a language that does not make a lexical green/blue distinction (Tarahumara, Uto-Aztecan family) do not. Interestingly, Kay and Kempton (1984) also found that the effect was eliminated in English speakers who were given instructions designed to inhibit the spontaneous activation of color names—suggesting that categorical perception of color stems from the activation of color names. These findings have by now been reinforced with larger and more carefully controlled studies (Özgen & Davies, 2002; Roberson & Davidoff, 2000; Winawer et al., 2007). It now appears to be established that learning the particular categories named by simple words in one's native language produces so-called categorical perception effects at the boundaries; these effects are suppressed by concurrent tasks that interfere with the activation of color names, which fact strongly implies that the categorical effect is verbally mediated.

The picture is clouded by the results of Franklin and Davies (2004) that showed categorical perception of color in prelinguistic infants. This work echoed the earlier studies of Mark Bornstein and colleagues, which showed analogous effects in prelinguistic infants (Bornstein, Kessen, & Weiskopf, 1976) and in macaques (Sandell, Gross, & Bornstein, 1979). The results of Franklin and Davies, and of Bornstein et al., strongly suggest innate category boundaries of some sort. Nevertheless, categorical perception in adult humans has been found to vary across languages in a manner predicted by the differing category boundaries of the languages concerned, as previously described. Furthermore, the fact that verbal interference has been shown in several independent studies to eliminate such categorical perception effects at the boundaries of lexically encoded categories suggests strongly that linguistic categorization plays a role in the low-level processing of color stimuli. All this amounts to a yes answer to question (2): Differences in language structure do seem to influence cognition or perception. In the traditional framing of the debate, this would be considered a "relativist" finding.

We now turn to question (1): Are semantic distinctions in languages determined by largely arbitrary linguistic convention? Roberson and colleagues promote a yes answer to this question as well, thus arguing for a thoroughly relativist position:

> [W]e will propose that *color categories are formed from boundary demarcation based predominantly on language*. Thus, in a substantial way, we will present evidence in favor of *linguistic relativity.* (Roberson et al., 2000: 394. Italics added)

They adduce several pieces of evidence to support this view that it is primarily local linguistic convention that determines linguistic category boundaries. First, English color categories do differ from those of Berinmo. Second, and more significantly, Roberson and colleagues failed to replicate several of Rosch's Dani results concerning the cognitively privileged status of the proposed universal focal colors. This is significant as universal color foci have been taken to be the source of universals in color naming. If this cognitive foundation of color naming universals is either nonexistent or ineffective—as they argue—perhaps color naming in general is less constrained than has been supposed. These authors mention only one constraint on color naming across languages, "grouping by similarity," implying it is the only constraint—and it is a rather loose one:

> The most important [nonlinguistic] constraint [on color terminologies] would be that similar items (as defined by perceptual discrimination) are universally grouped together. Thus, no language would exhibit categories that include two areas of color space but [exclude] an area between them. (Roberson et al., 2000: 395)

> No language has ever been reported to have a category that includes two areas of color space (e.g. yellow and blue) but excludes an area between them (green). There is no associative chain of similarity that could connect yellow to blue without passing through green. Grouping always follows principles of similarity (as defined by perceptual discrimination), and *the only free parameter appears to be the placement of boundaries between categories.* (Roberson, 2005: 65. Italics added)

This view leaves the actual *location* of these categories in color space apparently unconstrained—in direct contrast with the universalist notion that categories are formed around universal color foci. The relativist view moreover receives support from the suggestion that there is in fact no objective, reliable evidence for universals of color naming. John Lucy has argued that color naming universals reside only in the minds of universalist investigators—and not in the languages of the world:

> [Work in the Berlin and Kay tradition] not only seeks universals, but sets up a procedure which guarantees both their discovery and their form . . . . when a category is identified . . . it is really the investigator who decides which 'color' it will count as . . . What appears to be objective—in this case, a statement of statistical odds—is [not]. (Lucy, 1997: 331–334)

You can almost feel the pendulum swing.

Resolving the Question of Color Naming Universals

Both Lucy and Roberson et al. have proposed reasons to doubt existing evidence on color naming universals. Critically, however, neither has actually *demonstrated* that color naming is largely unconstrained.

So are there universals of color naming, or not? This question is now particularly important as we have established a "relativist" answer to question 2: Color language does affect color cognition. If we confirm Roberson's and Lucy's suggestion of a "relativist" answer to question 1 as well, and find that color naming varies across languages without universal constraints stronger than "grouping by similarity," that would support the recent relativist trend. However, if there are more substantial universal tendencies of color naming, that will suggest that, as we have proposed, there is no simple "universalist" or "relativist" answer to the language and thought debate in the color domain—and that the universalist/relativist framing is best dismantled, as the answers to its two framing questions do not match.

We thought that computational methods might be helpful in answering this question, as they have been with related questions (Regier & Carlson, 2001; Regier, 1996; Croft & Poole, 2008; Kirby & Christiansen, 2003; Steels & Belpaeme, 2005). In particular, we believed that applying computational methods to a large body of color naming data would effectively address any concerns, such as Lucy's, about an interpretive middleman possibly skewing the findings.

The body of data we relied on was that of the World Color Survey (WCS). The WCS was undertaken to evaluate criticisms of Berlin and Kay based on the small sample of languages directly assessed, the paucity of unwritten languages of low-technology societies in that small sample, the fact that the participants all spoke English and lived in the San Francisco Bay area, and other perceived methodological flaws. The color terminologies of 110 unwritten languages were assessed in the WCS, with a mean of 24 participants per language; participants were as monolingual as could be found by the field linguists, although many spoke other unwritten languages and some spoke European languages. Each participant named each of the color chips in the naming grid shown in Figure 8.1, and also indicated which chip was the best example of each color term in his or her language. Several papers have appeared, based on qualitative inspection of the data, claiming to have found universal tendencies in naming as well as an evolutionary sequence of color term systems similar to that originally postulated by Berlin and Kay, though not identical to it (Kay, Berlin, & Merrifield, 1991; Kay & Maffi, 1999; Kay, Berlin, Maffi, & Merrifield, 1997). However, these reports have failed to convince the skeptics. In particular, the use of qualitative, informal inspection is precisely what has concerned some critics. What is needed is an objective test, of the sort that can be supplied by a computational or other quantitative analysis.

In a series of studies, we applied computational methods to the WCS and related data, and demonstrated that universal constraints beyond "grouping by similarity" operate on color naming across the world's languages.

Kay and Regier (2003) asked whether color categories across languages tend to cluster in color space at rates greater than chance. To answer this, they first represented each color category in each of the 110 languages of the WCS by the centroid of that category: Thus, each category was represented as a single point in color space, corresponding to the center of mass of those chips that were named by that category.[2] Then, for each category in each language, they found the distance to the nearest category in each other language, and added up these distances. This yielded a measure of *cross-language dispersion* of categories: The larger this quantity, the more dispersed the categories across languages; the smaller this quantity, the more clustered the categories. The critical question was whether the empirically observed dispersion was significantly less than would be expected by chance. This was tested through a Monte Carlo simulation, in

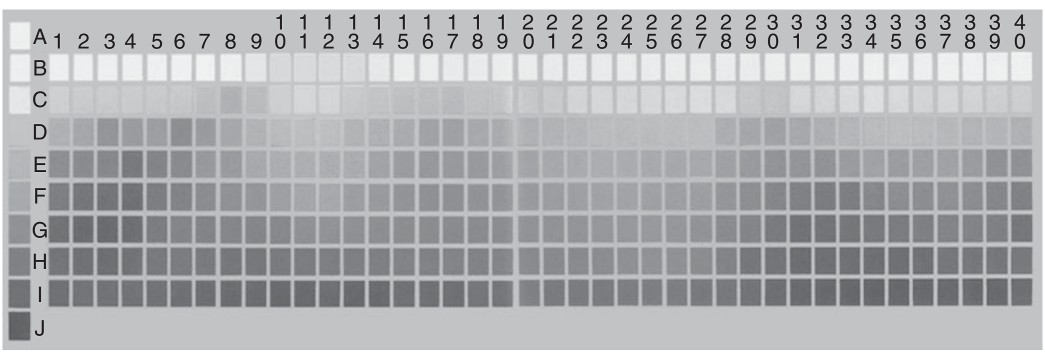


FIGURE 8.1 Grayscale rendition of Color Plate 3. See Color Plate 3 for interpretation. Color naming grid.

which the observed dispersion was compared to a computationally generated distribution of dispersion values that would be expected by chance. But how much dispersion would be expected by chance? This is a slightly tricky question, as a certain amount of dispersion of categories will be found *within* a given language—and the method of generating a random theoretical distribution must respect that fact. Accordingly, Kay and Regier (2003) derived the theoretical distribution through a manipulation of the WCS data itself, as follows. All categories in each WCS language were rotated by a random amount in the hue plane—the same random amount for each language (preserving natural within-language dispersion, as urged above), and different random amounts across languages (randomizing cross-language structure, appropriately, as this is the focus of the test). This process, which is loosely analogous to scrambling a combination lock, resulted in a single randomized theoretical version of the WCS dataset, as shown schematically in Figure 8.2. The dispersion in the randomized dataset can be viewed as being generated by chance.

This process was repeated 1000 times, resulting in a distribution of dispersion values that could be expected by chance. The actual empirically observed WCS dispersion was then compared to this distribution. The actual dispersion of WCS centroids was well below the lowest of the dispersions in the 1000 randomized datasets, meaning that actual WCS categories are clustered across languages to a degree greater than chance, $p < 0.001$. An analogous Monte Carlo simulation showed moreover that color categories in the WCS cluster near those in the data of Berlin and Kay (1969) to a degree greater than chance. These results objectively demonstrate universals of color naming.

These findings leave open an important issue, however: the status of the proposed universal focal colors or universal best examples. These focal colors are often taken as the source of color-naming universals—but as we have seen, Roberson et al. (2000) have shown that foci may not be universally cognitively privileged as was earlier claimed, and they use this finding to cast doubt on the entire notion of universal foci. They suggest instead that color categories are defined at their boundaries by language, and that best examples are then extracted secondarily as the *centers* of these language-defined categories (Roberson et al., 2000: 395).

Are the boundaries of color categories organized around universal best examples (foci)—or are best examples determined from language-demarcated boundaries? Regier, Kay, and Cook (2005) sought to answer this question. They did this by examining the best example choices given by all speakers, for all color terms, in all languages of the WCS, taken in aggregate. They first asked if these WCS best example choices were similar to those of English; Figure 8.3 shows that they are. The contour plot shows the number of WCS best example choices that fell on each chromatic

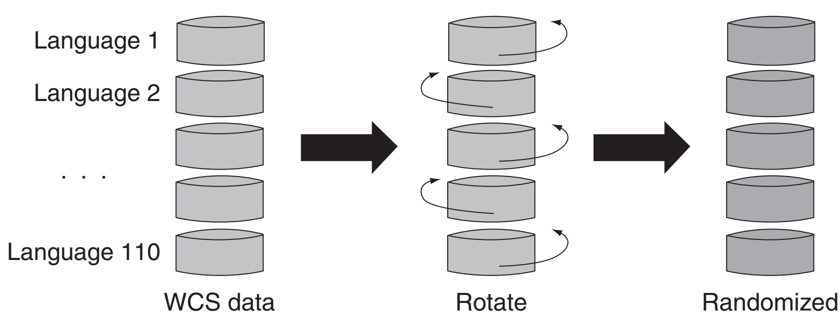


FIGURE 8.2 Creating a randomized dataset.
(Reprinted from Kay & Regier, 2003.)

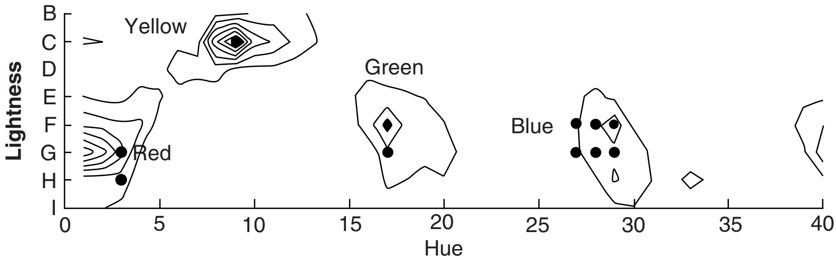


FIGURE 8.3 Best examples of WCS color terms, compared with those of English.
(Reprinted from Regier et al., 2005.)

chip of the WCS stimulus array. The outermost contour represents 100 hits, and each subsequent inner contour represents an increment of 100 hits. The black dots represent the best examples of English red, yellow, green, and blue, provided by one U.S. speaker, as reported by Berlin and Kay (1969). The WCS distribution is evidently quite close to English. It is worth emphasizing that the WCS data are from languages of nonindustrialized societies—so a similarity between WCS best examples and those of English cannot be attributed to the global spread of industrialization; rather, it suggests that best examples are determined by genuinely universal forces. This pattern would not be predicted if best examples are derived secondarily as the centers of categories whose boundaries are determined primarily by local linguistic convention—instead, it suggests a central role for universal foci in color naming, after all.

What then are we to make of Berinmo, the language that has been most saliently held up recently as a counterexample to universal tendencies in color naming? A recent reanalysis of the Berinmo data suggests that the language may not be so atypical. Kay and Regier (2007) found that the boundaries of Berinmo color categories, which Roberson and colleagues suggested were determined by linguistic convention, are in fact relatively close to category boundaries in WCS languages—closer than would be predicted if "grouping by similarity" were the main constraint on color naming across languages.

In sum, whether we consider category centroids (Kay & Regier, 2003), best examples (Regier et al., 2005), or boundaries (Kay & Regier, 2007), universal tendencies in color naming are objectively supported in the world's languages, and the relativist view of color naming is empirically challenged. Color naming is universally constrained, and far from arbitrary.

Nonetheless, there is also evidence that appears to challenge, or at least soften, the universalist view of color naming, and to suggest that linguistic convention may play some limited role in determining the boundaries of named color categories. Specifically, even languages with similar color-naming systems do differ in their category extensions. For instance, the Berinmo color-naming scheme is broadly similar to that of Himba, a language of Namibia that Roberson et al. (2005) have studied—but the category boundaries clearly differ across the two languages. Because the color-naming systems of these two languages appear to be organized around the same grouping of the universal foci, the difference in boundaries apparently stems from something other than the foci.

Color Naming Is Near Optimal

What sort of account accommodates both universal tendencies and such cross-language differences? Regier, Kay, and Khetarpal (2007) suggested a possible answer, building on an earlier proposal by Jameson and D'Andrade (1997): Color naming across languages reflects *near-optimal partitions* of an irregularly shaped perceptual color space.

One possible explanation [for universals in color naming] is . . . the irregular shape of the color space . . . Hue interacts with saturation and lightness to produce several large ''bumps''; one large bump is at focal yellow, and another at focal red. . . . We assume that the names that get assigned to the color space . . . are likely to be those names which are most informative about color. (Jameson & D'Andrade, 1997: 312)

Regier et al. (2007) formalized this proposal as follows. We represented the colors of the grid shown in Figure 8.1 as points in the CIEL*a*b* color space; we chose CIEL*a*b* because the distance between two colors in this space is a reasonable approximation to their psychological dissimilarity. We then considered a categorical partition of these colors to be *well-formed* to the extent that the partition maximized perceptual similarity within categories and minimized it across categories (Garner, 1974). We hypothesized that the color-naming systems of the world's languages correspond to maxima or near-maxima in well-formedness—and in that sense, to theoretically near-optimal color-naming systems.

We used simulations to create theoretically optimal color-naming systems with $n = 3, 4, 5, 6$ categories. We initialized each simulation by randomly assigning each color in the grid to one of the n categories, and then adjusted category labels through steepest ascent in well-formedness until a maximum was reached. The results are displayed in Figure 8.4, compared with mode maps[3] for selected languages from the WCS database.

Well-formedness optimization places boundaries in roughly the right places for these languages, and correctly predicts some details. For instance, in three-term systems, the composite red/yellow term excludes the lighter shades of yellow—both in the simulation and in naturally occurring three-term systems, exemplified here by Ejagam. In contrast, when there is a separate yellow term, that term includes the lighter shades—both in the simulation and in reality.

There are also many languages in the WCS with color-naming systems that are not very similar to these theoretical optima. Nonetheless, we predicted that all languages would be at least near optimal. To test this proposal, we compared the well-formedness of a given language's color-naming system to that of a set of hypothetical systems derived from the original by rotation, as illustrated

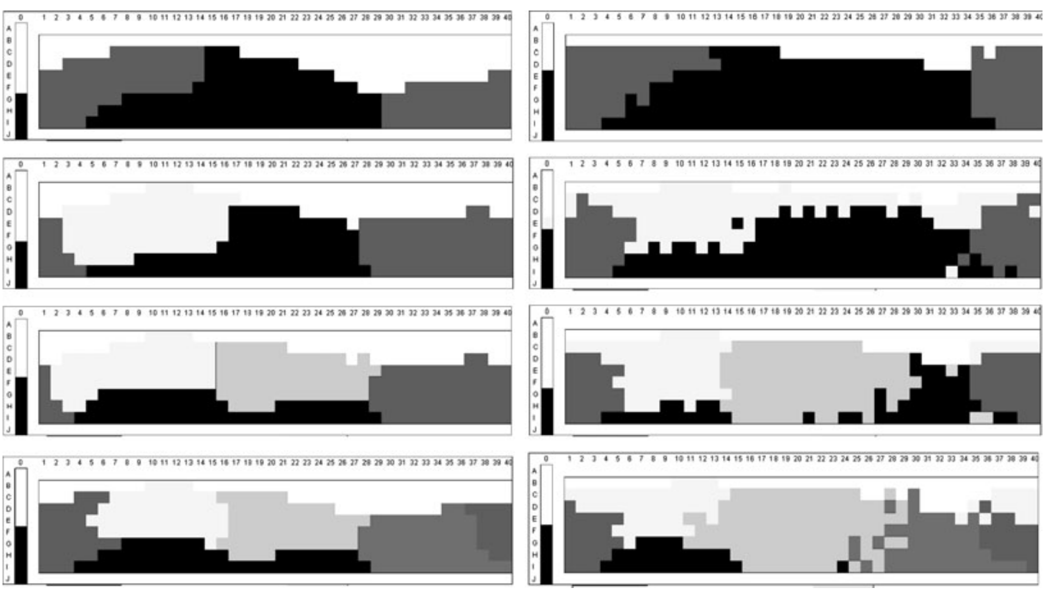


FIGURE 8.4 Grayscale rendition of Color Plate 4. See Color Plate 4 for interpretation. Theoretical optima (left) compared with selected WCS languages (right), for $n = 3, 4, 5, 6$ categories. The WCS languages are, from top to bottom: Ejagam (Nigeria, Cameroon), Culina (Peru, Brazil), Iduna (Papua New Guinea), and Buglere (Panama).

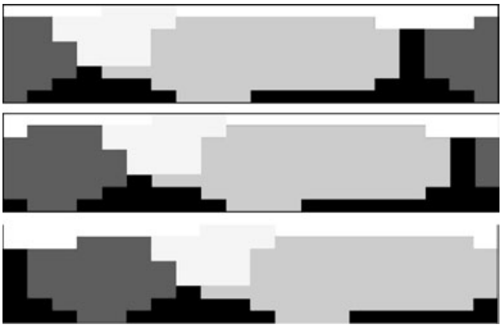


FIGURE 8.5. Grayscale rendition of Color Plate 5. See Color Plate 5 for interpretation. Creating hypothetical color-naming systems by rotation. The top panel shows the color-naming system of Berinmo; the lower panels show the same system rotated by four (middle panel) and eight (bottom panel) hue columns.

in Figure 8.5. Specifically, we rotated the original language's system by 2, 4, 6, etc. (and −2, −4, −6, etc.) Munsell hue columns, yielding a hypothetical variant for every two columns around the entire hue circle. We predicted that the naturally occurring (unrotated) system would have higher well-formedness than any of the rotated variants. Why? Because by hypothesis the boundaries of the naturally occurring system are near optimal, whereas in the rotated systems the boundaries have been deliberately shifted away from these hypothetically near-optimal positions.

We first examined Berinmo, and found that the Berinmo color-naming system indeed has higher well-formedness than any of its rotated variants. This casts further doubt on Berinmo's proposed status as a counterexample to universals of color naming. Instead, it appears that the Berinmo system is located where it is along the hue dimension because the structure of perceptual color space makes its actual location the optimal location. At the same time, however, we also found evidence suggesting where linguistic convention may get some wiggle room. The rotated Berinmo variant at +2 columns had well-formedness nearly as high as the naturally occurring system, demonstrating that small differences in boundary placement can sometimes yield only very modest differences in well-formedness. This fact may explain why similar languages differ somewhat in their boundary placements. It may be that the universal structure of perceptual color space makes some systems preferable to others—and linguistic convention is then free to select from among the set of highly well-formed systems, some of which will resemble each other. This view accounts both for universal tendencies of color naming (e.g., Kay & Regier, 2003) and for the observation that similar languages sometimes differ in the placement of their boundaries (e.g., Roberson et al., 2005).

We repeated the rotation-based analysis for all languages of the WCS, and obtained comparable results: For most languages (82 of 110), the unrotated (attested) system has higher well-formedness than any rotated variant. For most of the remaining languages, the well-formedness of the unrotated system was almost as high as the maximum rotated variant—supporting the proposal that color-naming systems of the world's languages are near optimal, while also allowing for a limited amount of language-specific determination of category boundaries.

To summarize to this point, empirical support has been established for the "universalist" tenet that there are constraints on color naming across the world's languages that go well beyond "grouping by similarity"—but at the same time, our findings leave open the possibility that linguistic convention may play some role in selecting from among the class of well-formed color-naming systems. This leaves us with an interestingly complex view of color naming: ultimately universalist, but with a relativist tinge. In the following section, we consider whether linguistic category differences affect perception—and argue for a view that is ultimately relativist, but with a universalist tinge. Universalism and relativism are both wrong, and also both right.

WHORF HYPOTHESIS IN THE RIGHT VISUAL FIELD

As we have seen, the universals-versus-relativity distinction paints with a brush too broad to capture an interestingly differentiated reality. It misleadingly collapses together the two central questions of the language-and-thought debate, and also inappropriately demands simple yes-or-no answers to these questions. We have seen that the evidence on color naming supports a hedged universalist view. Here we show that recent research on the question "Does language affect perception?" supports a hedged relativist view. This conclusion flows from a finding unanticipated by the framing of the debate: Language may affect perception in the right half of the visual field and much less, if at all, in the left half. Based on this view, language simultaneously affects perception, and affects it much less, if at all, in the same individual, depending on which part of the visual world is considered. Here, we first review literature that suggests why Whorfian effects might be lateralized to the right visual field, and then discuss recent work that directly supports the idea.

Motivation

What motivation is there for the idea that the Whorf hypothesis is lateralized to the right visual field (RVF)? There is a chain of findings that makes the idea seem fairly reasonable, a priori.

As mentioned earlier, Kay and Kempton (1984) found an effect of language on color cognition.[4] They compared the color similarity judgments of speakers of English, which has distinct words for "green" and "blue," and Tarahumara, which does not, instead having a single, broader, named color category encompassing most of green *and* blue. They found enhanced dissimilarity of colors at the green/blue boundary in English speakers, but not in Tarahumara speakers—suggesting that linguistic distinctions may heighten perceptual differences.

How might this happen? One possibility is a "naming strategy": The sight of a color activates the name that would ordinarily be used to label that color (e.g., "green")—and then when comparing two colors, we also implicitly compare their names. Two colors with different names would appear to be more distinct than colors with the same name, because of the involvement of names in the comparison process. Note that this is an *on-line* effect: The linguistic influence occurs during the act of perception. Another possibility, and one arguably more consonant with Whorf's original proposals, is that the habitual use of a particular language permanently changes our perceptual apparatus: The language's distinctions get "burned in" to our perceptual machinery, so to speak, and it is by these permanent changes, rather than by an on-line process, that language affects perception.

As mentioned earlier, Kay and Kempton (1984) found evidence for the on-line option. When participants were shown two colors that would ordinarily be given different names, but were *told* that these two colors had the same name, there was no enhancement of dissimilarity between the two colors: The Whorfian effect was eliminated. This argues against a permanent change, and in favor of a process in which on-line representations of names shape perception—as long as these representations are not interfered with, and thereby incapacitated. This conclusion is also supported by a number of more recent studies. In particular, Roberson and Davidoff (2000) found that apparent effects of language on perception were eliminated by a verbal interference task (see also Winawer et al., 2007). This strongly supports the interpretation that the effect is fundamentally linguistic in origin—rather than due to cultural or environmental differences between populations: When an experimental manipulation effectively incapacitates language processing, the effect vanishes. But at the same time, it also strongly supports an on-line over a permanent-change interpretation.[5]

What does that mean for us? If Whorfian effects are mediated on-line by names, that implicitly raises the question: Where are these names represented? A likely answer is in the left hemisphere (LH) of the brain, as preferential involvement of the LH is found for most language tasks[6] (Corballis, 1991; Hellige, 1993), including those involving access to names (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996).

It is also well established that for most senses, perceptual input from one side of the body projects to the *contralateral* hemisphere of the brain. For instance, the right visual field (RVF) projects directly to the LH, whereas the left visual field (LVF) projects directly to the right hemisphere (RH). (Similarly, information received in the right ear projects largely to the LH, and the left ear to the RH.) This pattern suggests that visual stimuli in the RVF might have more immediate access to, and be more affected by, the linguistic representations of the LH, as compared with visual stimuli in the LVF. This proposal is essentially a Whorfian analogue of some already well-established findings concerning the perception of linguistic material: The right ear appears to be dominant over the left ear in the recognition of spoken words, and the RVF appears to be dominant over the LVF in the recognition of written words (Kimura, 1961, 1973). Here, we take this known pattern of lateralized linguistic influence one step further—one Whorfian step further—and propose that it extends to linguistic effects on the perception of *non*linguistic material as well: "lateralized Whorf," if you will.

But Is It True?

What sort of data would directly support this proposal? The lateralized Whorf hypothesis makes three predictions:

1. The discrimination of stimuli with different names should be faster in the RVF than in the LVF, as the difference in names will heighten perceptual differences in the RVF.
2. The discrimination of stimuli with the *same* name should be *slower* in the RVF than in the LVF, as the sameness of the name will impede perceptual discrimination in the RVF.
3. This overall pattern should be disrupted by concurrent tasks that interfere with verbal processing, but not by concurrent tasks of comparable difficulty that interfere only with nonverbal processing.

Gilbert, Regier, Kay, and Ivry (2006) tested these predictions in a color discrimination task designed to probe the lateralized Whorf hypothesis. They defined a continuum of four hues spanning the "green"/"blue" boundary in English: Two of these colors were instances of "green" (as determined in a color naming task) and two were instances of "blue." This continuum is illustrated in Figure 8.6, although the specific colors shown here may not be fully accurate renditions of those used. Here, colors A and B are different hues of "green" and C and D are different hues of "blue."

On each experimental trial, participants were asked to fixate on a centrally displayed fixation cross, and then a circle of colored squares appeared around it, as in Figure 8.7.

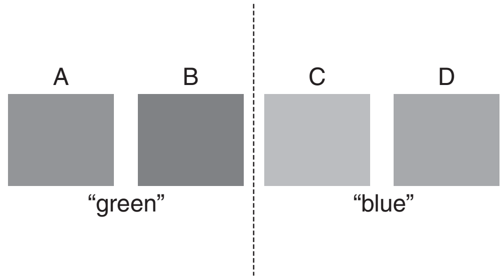


FIGURE 8.6. Grayscale rendition of Color Plate 6. See Color Plate 6 for interpretation. Four colors (A–D) spanning the green/blue boundary—two greens and two blues.

(Reprinted from Gilbert et al., 2006.)

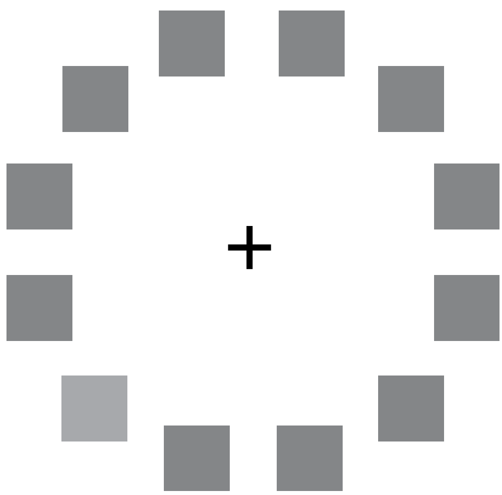


FIGURE 8.7. Grayscale rendition of Color Plate 7. See Color Plate 7 for interpretation. Visual search task: is the odd-man-out on the left or the right?

(Reprinted from Gilbert et al., 2006.)

All of the squares were of the same color except for one that was of a different color; we refer to this odd-man-out as the "target" and the other squares as "distractors." Critically, the color of the target had either the same name as that of the distractors (e.g., two different hues of "green") or a different name (e.g., the target was a hue of "green" and the distractors a hue of "blue"). The target square appeared in either the RVF or the LVF, and participants were asked to indicate the side of the circle (left or right) on which the target appeared, by making a keyboard response using the corresponding hand. The task was performed under three conditions: (1) without any interference task, (2) with a concurrent verbal interference task (silently remembering a word for a color other than green or blue, e.g., "red," which presumably requires verbal rehearsal), and (3) with a concurrent nonverbal interference task (remembering a spatial grid of squares, which is assumed to not require verbal rehearsal). These three conditions allowed a direct test of the three cardinal predictions. The dependent variable was reaction time.

Eleven right-handed Berkeley undergraduates, all native English speakers, performed the visual search task under the three conditions previously listed. Figure 8.8 shows reaction times from the no-interference condition. Here, the effect of language on perceptual discrimination appears to be restricted to the RVF: When the target appeared in this visual field, RTs for stimulus pairs with different names ("between categories") were faster than for pairs with the same name ("within category"); in contrast, when the target appeared in the LVF no such difference was observed.[7]

Moreover, in support of Prediction 1, RTs for stimulus pairs with different names ("between categories") were faster for RVF targets than for LVF targets. In support of Prediction 2, RTs for stimulus pairs with the same name ("within category") showed the opposite pattern: slower for RVF targets than for LVF targets.

The results of the verbal interference condition are shown in Figure 8.9. Under this condition, the original pattern of lateralization is disrupted, supporting Prediction 3. In fact, the results were reversed. RTs to stimulus pairs with different names ("between categories") are now *slower* for RVF targets than for LVF targets. Similarly, RTs to stimulus pairs with the same name ("within category") are now *faster* for RVF targets than for LVF targets. This reversal was not predicted; only the more general idea of disruption was. Thus, we take these results to support Prediction 3, but to also raise as-yet-unanswered questions as to why the verbal interference task actually leads to a significant reversal of the lateralized Whorf effect.

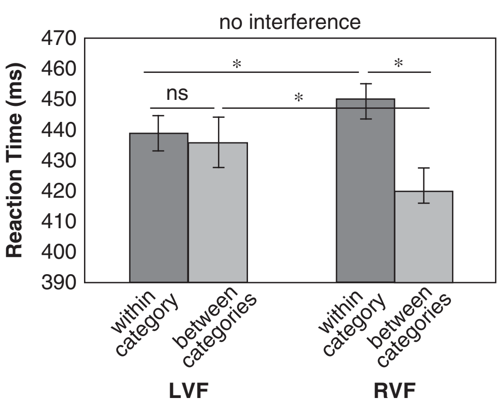


FIGURE 8.8. No-interference condition: an effect of language in the RVF, but not the LVF.[8]

(Reprinted from Gilbert et al., 2006.)

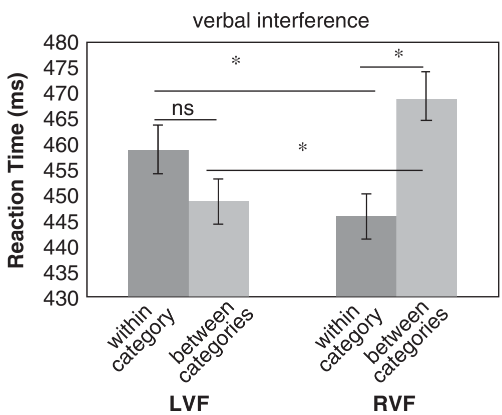


FIGURE 8.9. Verbal interference disrupts the pattern of lateralization.

(Reprinted from Gilbert et al., 2006.)

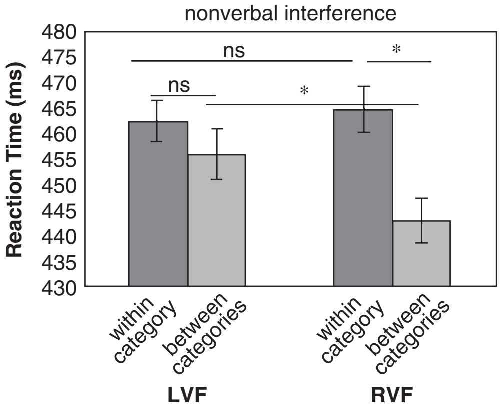


FIGURE 8.10 Nonverbal interference largely preserves the pattern of lateralization.
(Reprinted from Gilbert et al., 2006.)

The results of the nonverbal interference condition are shown in Figure 8.10. These are similar to the results obtained without interference—again in support of Prediction 3. The one qualitative difference is that here, RTs to stimulus pairs of the same name ("within category") are the same for left and right visual field targets. But for stimuli with different names ("between categories"), we obtain the same RVF superiority as we did without any interference. This pattern, taken together with the disruption caused by verbal interference, suggests that the lateralized Whorf pattern that was obtained without interference was due to language: This pattern is selectively disrupted by a linguistic, but not a nonlinguistic, interference task.

The idea behind this study was originally sparked by curiosity as to whether Whorfian effects would be lateralized to the RVF in a split-brain patient—that is, a patient whose corpus collosum had been surgically severed. A further experiment confirmed that this is indeed the case. However, the results just presented are perhaps more striking, as they demonstrate the same pattern in normals—despite the possibility of information transfer between hemispheres.

These findings open up a number of interesting questions. The study examined only English speakers, on only one semantic distinction ("green" vs. "blue"), in only one semantic domain (color), in only one perceptual modality (vision). Yet the logic behind the study is very general, and could in principle apply to speakers of any language, on any perceptual distinction that is marked in that language, and in any sensory modality for which lateralized inputs are primarily projected to one hemisphere. Recall, for instance, the language mentioned earlier, with a single term covering both green and blue—and therefore without a green/blue category boundary. We would expect speakers of such a language, unlike English speakers, to show a "within-category" pattern across all of the stimulus pairs in the experiments described. We would also expect linguistically driven "lateralized Whorf" results to be observed at other color category boundaries, in other parts of color space—and in other visually based semantic spaces (e.g., spatial relations), or in semantic spaces based on other perceptual modalities that project contralaterally to the brain.

Is There Any Other Evidence?

One possible objection is that this is just one study, conducted in one laboratory. The lateralized Whorf argument as a whole would be more compelling if there were converging evidence coming from elsewhere. Is there any such evidence?

There are a number of encouraging signs in the literature. Malone and Hannay (1978) found RVF superiority overall in the discrimination of color hues from memory—whereas Davidoff (1976) and Hannay (1979) found the opposite pattern, an LVF superiority, in hue discrimination. The apparent inconsistency is possibly resolved by noting that the RVF superiority was found in a study that examined pairs of colors that were quite dissimilar—and therefore likely to have different names—whereas the LVF superiority was found in studies that examined pairs of colors that were more similar—and therefore more likely to have the same name. Moreover, Hannay (1979) found that the higher the number of color pairs for which both members were given the same name in a color-naming task, the larger the LVF superiority in discrimination—suggesting a linguistic basis for the lateralization. The study of Gilbert et al. (2006) builds on these earlier findings by demonstrating RVF superiority in discrimination precisely at language-demarcated category boundaries (and LVF superiority elsewhere), and by confirming the linguistic nature of the effect through the use of interference conditions.

Drivonikou, Kay, Regier, Ivry, Gilbert, Franklin, and Davies (2007) have recently replicated the general findings of Gilbert et al. (2006) in a different laboratory, and at the blue–purple boundary as well as the green–blue boundary. However, there is one respect in which their findings diverge from those of Gilbert et al., and suggest a slight modification of the "lateralized Whorf" proposal. Although they did find a stronger category effect in the RVF than the LVF, they also found a weak LVF category effect. Thus, in some circumstances language may affect perception in both halves of the visual field—but still more in the right half than the left. Interestingly, the reaction times reported by Drivonikou et al. were also slightly slower overall than those reported in the original Gilbert et al. study. It may be that the weak LVF category effect reported by Drivonikou et al. resulted from interhemispheric transfer of information across the corpus callosum—whereas the participants in the study by Gilbert et al. did not exhibit such an LVF category effect because their responses were too fast to be affected by this transfer.

As noted, the lateralized Whorf hypothesis is very general; thus we might also expect supporting evidence from domains other than color. One finding that may be relevant concerns the perception of spatial relations. Kosslyn, Koenig, Barrett, Cave, Tang, and Gabrieli (1989) showed participants a dot located somewhere either above or below a horizontal bar, and asked either "is the dot above or below the bar?"—this was considered a categorical judgment—or "is the dot within 2 cm of the bar?"—this was considered a metric, or coordinate, judgment. They found that categorical judgments were faster in the RVF, whereas coordinate judgments were faster in the LVF. They concluded that the LH was specialized for categorical spatial perception and the RH for coordinate spatial perception—and in subsequent work suggested that this is due to the nature of neural connectivity within each of the two hemispheres (Kosslyn, Chabris, Marsolek, & Koenig, 1992). However, another possibility is that the categorical nature of the LH is tied to language—and that the study by Kosslyn et al. (1989) is in essence a spatial forerunner of the study by Gilbert et al. (2006). This possibility can be tested by introducing verbal and nonverbal interference conditions to the paradigm of Kosslyn et al.

There is also direct evidence of a lateralized Whorf effect in a domain other than color. Gilbert, Regier, Kay, and Ivry (2008) recently replicated the findings of Gilbert et al. (2006), but using pictures of dogs and cats rather than blue and green colors as stimuli.

In sum, it appears that Whorfian effects of language on perception may be lateralized to the RVF. When we ask whether language affects perception, then, the answer appears to be neither a simple yes nor a simple no—the two answers implicitly offered by the traditional framing of the language-and-thought debate—but instead yes in the RVF and perhaps less so in the LVF, a possibility as yet unexamined in the debate.

CONCLUSIONS

We begin our concluding remarks by returning to the question posed in the title of this chapter: "Which side are you on, anyway?" We hope by now to have convinced the reader that she or he would be ill-advised to wholly back either the universalist or the relativist view of language and thought—and would be better off instead thinking outside the standard "universals versus relativity" framing. The traditional framing is simplistic, and hides interesting realities. One such reality is that at least in the color domain, there are clear universals governing the semantic distinctions that languages make—but there may also be some limited element of arbitrariness in exactly where category boundaries are drawn. This is an ultimately universalist finding, but with a relativist twist.

The second reality obscured—or at least left entirely unanticipated—by the traditional framing is that language may affect perception primarily in the right half of the visual field and much less if at all in the left half. These "lateralized Whorf" results—ultimately relativist this time, but again with a twist—reinforce the impression left by the review of color naming: The world is a more interestingly complicated place than is suggested by the options presented in the traditional framing of the debate.

What *useful* role might the universals-versus-relativity opposition play? After all, it seems unlikely to simply vanish, given its naturalness and its connection to inescapably engaging questions about the extent to which we are creatures of our environment, or of an innately given human nature. Certainly it is a convenient way of quickly sketching the major issues and linking them to larger questions that will seem interesting to a broad audience. But perhaps this is the extent of its usefulness—as a means of starting, rather than pursuing, a conversation about language and thought.

ACKNOWLEDGMENTS

We thank Barbara Malt and Phillip Wolff for helpful comments on an earlier draft of this chapter. This work was supported by NSF Grants 0418283 (to T.R.) and 0418404 (to P.K.).

Notes

1. Strictly speaking, categorical perception refers to enhanced perceptual discrimination of stimuli that straddle a category boundary, as compared to equivalently spaced stimuli that fall within the same category. However, the term is also sometimes applied to enhanced cross-category discrimination from memory, rather than just from perception. A number of studies on the "categorical perception" of color actually fall under this broader usage of the term.
2. Kay and Regier (2003) used CIEL*a*b* color space, a three-dimensional color space in which the L* dimension represents lightness, and the two remaining dimensions, a* and b*, define a plane orthogonal to L*, such that angle in that plane represents hue and radius represents saturation. Distance in this space is roughly comparable to psychological dissimilarity.
3. The mode map for a language is a depiction of the color grid showing for each color chip the most frequent name it received from participating speakers of that language. Each colored region corresponds to a named color category.
4. Throughout this section, we will be exploring the "lateralized Whorf" question by examining the naming and cognition of *color*. However, this is only a matter of convenience, not one of principle: the issues at play reach beyond color, to the effect of language on the perception of any visual stimulus.
5. These findings act as a sort of Rorschach test. Those who "want" the Whorf hypothesis to be true can point to the fact that the manipulation clearly implicates language. At the same time, those who "want" the hypothesis to be false can point to how easy it is to eliminate effects of language on perception, and argue on that basis that Whorfian effects are superficial and transient.
6. At least in right-handers. Language function is less clearly lateralized in left-handers.
7. Despite our efforts, the four colors A, B, C, D were not perfectly evenly spaced in color space. This fact complicates comparison of "within category" responses to "between categories" responses—but such a comparison is still possible. In CIEL*a*b* color space, the (within category: green) A–B distance is less than the (between category) B–C distance, which is less than the (within category: blue) C–D distance. Follow-up analyses that treated the data for within-green and within-blue conditions separately, rather than pooled together, produced results qualitatively the same as those reported. This suggests that the influence of language on perception is stronger than any bias that may have been introduced by the uneven spacing.
8. In all figures, "*" means "$p < .05$"; "ns" means "not significant", both in protected t-tests.

References

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

Bohanon, P. (1963). *Social anthropology*. New York: Holt, Rinehart & Winston.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2*(1), 115–129.

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49,* 454–462.

Conklin, H. C. (1955). Hanunóo color categories. *Southwestern Journal of Anthropology, 11,* 339–344.

Corballis, M. C. (1991). *The lopsided ape: Evolution of the generative mind*. New York: Oxford University Press.

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

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature, 380,* 499–505.

Davidoff, J. (1976). Hemispheric sensitivity differences in the perception of colour. *Quarterly Journal of Experimental Psychology, 28*(3), 387–394.

Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L. (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences, 104,* 1097–1102.

Franklin, A., & Davies, I. R. L. (2004). New evidence for infant color categories. *British Journal of Developmental Psychology, 22,* 349–377.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum.

Geiger, L. (1880). *Contributions to the history of the development of the human race*. London: Tubner and Company.

Gilbert, A., Regier, T., Kay, P., & Ivry, R. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103*(2), 489–494.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the Whorfian effect beyond the realm of color discrimination. *Brain and Language, 105*(2), 91–98.

Gladstone, W. E. (1858). *Studies on Homer and the Homeric age*. London: Oxford University Press.

Gleason, H. A. (1961). *An introduction to descriptive linguistics*. New York: Holt, Rinehart & Winston.

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

Hannay, H. J. (1979). Asymmetry in reception and retention of colors. *Brain and Language, 8*(2), 191–201.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93,* 10–20.

Heider, E. R., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337–354.

Hellige, J. B. (1993). *Hemispheric asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.

Hickerson, N. P. (1971). Review of basic color terms. *International Journal of American Linguistics, 37,* 257–270.

Jameson, K. A., & D'Andrade, R. G. (1997). It's not really red, green, yellow, blue. In C. L. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 295–319). Cambridge, UK: Cambridge University Press.

Kay, P., Berlin, B., Maffi, L., & Merrifield, W. (1997). Color naming across languages. In L. Maffi & C. L. Hardin (Eds.), *Color categories in thought and language* (pp. 21–56). Cambridge, UK: Cambridge University Press.

Kay, P., Berlin, B., & Merrifield, W. (1991). Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology, 1*(1), 12–25.

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

Kay, P., & Maffi, L. (1999). Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist, 101*(4), 743–760.

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences, 100*(15), 9085–9089.

Kay, P., & Regier, T. (2007). Color naming universals: The case of Berinmo. *Cognition, 102,* 289–298.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15,* 166–171.

Kimura, D. (1973). The asymmetry of the human brain. *Scientific American, 228,* 70–78.

Kirby, S., & Christiansen, M. H. (2003). From language learning to language evolution.

In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 272–294). Oxford: Oxford University Press.

Kosslyn, S., Chabris, C., Marsolek, C., & Koenig, O. (1992). Categorical versus coordinate spatial relations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 562–577.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 723–735.

Krauss, R. M. (1968). Language as a symbolic process in communication. *American Scientist, 56*, 265–278.

Lantz, D., & Stefflre, V. (1964). Language and cognition revisited. *Journal of Abnormal and Social Psychology, 69*, 472–481.

Lucy, J. (1992). Grammatical categories and cognition: a case study of the linguistic relativity hypothesis. Cambridge, UK: Cambridge University Press.

Lucy, J. (1997). The Linguistics of "Color." In C. L. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 320–346). Cambridge, UK: Cambridge University Press.

Lucy, J. A., & Shweder, R. A. (1988). Whorf and his critics: Linguistic and non-linguistic influences on color memory. *American Anthropologist, 81*, 581– 607.

Magnus, H. (1877). *Die geschichtliche Entwickelung des Farbensinnes*. Leipzig: Veit.

Magnus, H. (1880). *Untersuchungen über den Farbensinn der Naturvölker*. Jena: Fraher.

Malone, D. R., & Hannay, H. J. (1978). Hemispheric dominance and normal color memory. *Neuropsychologia, 16*(1), 51–59.

Nida, E. A. (1959). Principles of translation as exemplified by Bible translating. In R. A. Brower (Ed.), *On translation*. Cambridge, MA: Harvard University Press.

Özgen, E., & Davies, I. R. L. (2002). Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General, 131*, 477–493.

Ray, V. F. (1952). Techniques and problems in the study of human color perception. *Southwestern Journal of Anthropology, 8*, 251–259.

Ray, V. F. (1953). Human color perception and behavioral response. *Transactions of the New York Academy of Sciences (ser. 2), 16*, 98–104.

Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press.

Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General, 130*(2), 273–298.

Regier, T., Kay, P., & Cook, R. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences, 102*(23), 8386–8391.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences, 104*, 1436–1441.

Rivers, W. H. R. (1901). Introduction and vision. In A. C. Haddon (Ed.), *Reports on the Cambridge Anthropological Expedition to the Torres Straits, Vol. II, Physiology and Psychology, part I*. Cambridge, UK: Cambridge University Press.

Roberson, D. (2005). Color categories are culturally diverse in cognition as well as in language. *Cross-Cultural Research, 39*(1), 56–71.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*(6), 977–986.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50*(4), 378–411.

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*(3), 369–398.

Roberson, D., Davies, I. R. L., & Davidoff, J. (2002). Color categories are not universal: Replications and new evidence. In B. Saunders & J. van Brakel (Eds.), *Theories, technologies, instrumentalities of color: Anthropological and historiographic perspectives* (pp. 25–35). Lanham, MD: University Press of America.

Sandell, J. H., Gross, C. G., & Bornstein, M. H. (1979). Color categories in macaques. *Journal of Comparative and Physiological Psychology, 93*, 626–635.

Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28*(4), 469–529.

Stefflre, V., Castillo Vales, V., & Morley, L. (1966). Language and cognition in Yucatan. *Journal of Personality and Social Psychology, 4*, 112–115.

Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*, 7780–7785.

9 RELATIVELY SPEAKING

An Account of the Relationship between Language and Thought in the Color Domain

Debi Roberson & J. Richard Hanley

Is the "striped" appearance of the rainbow a construct of our higher mental processes or is it determined at an early perceptual level by the organization of human color vision? If your language does not have separate terms for "blue" and "green" (and many languages, including Welsh, do not) would you perceive these shades as more similar than a speaker of English? Although the nature of the relationship between natural language and our mental representation of the experienced world has been probed by philosophers, psychologists, linguists, and anthropologists in many areas (e.g., number: Gumperz & Levinson, 1997; Gordon, 2004; spatial relations: Majid, Bowerman, Kita, Haun, & Levinson, 2004; Choi, McDonough, Bowerman, and Mandler, 1999; time: Boroditsky, 2001; shape: Lucy, 1992; Roberson, Davidoff, & Shapiro, 2002), the color domain has been a principal focus of investigation.

One reason why so much research has focused on color categorization is that the range of colors visible by humans is large (approximately 2 million just noticeable differences), but the range of color terms available to describe them is generally small (languages contain between 2 and 22 basic[1] terms). Not only do some languages have gross differences in terminology (e.g., the use of a single term to refer to everything that an English speaker would call either blue, green, or purple), but even those languages with similar color vocabularies have slight variations in the range of stimuli covered by a particular term (e.g., the different ranges covered by the English term "blue" and the Italian term "*blu*").

Another reason why the color domain may be a particularly fruitful ground for investigating the relationship between language and thought is that the acquisition of color vocabulary by children is typically rather slow and error prone compared to the acquisition of terms in other domains (Bornstein, 1985; Au & Laframboise, 1990; O'Hanlon & Roberson, 2006). By 2 years of age, children can learn a novel category label from a single exposure if it is a highly imageable concrete noun (Heibeck & Markman, 1987), leaving only the briefest of intervals to study category acquisition. In contrast, children may still make color-naming errors at 6 years of age (Roberson, Davidoff, Davies, & Shapiro, 2004), so there is an extended period in which the process of color term acquisition can be studied.

This chapter is divided into six sections. The first sets out the background of the debate about the relationship between language and cognition in the color domain. The second explains how recent studies of color recognition employing visual search tasks have clarified this relationship. This section also argues that these studies point to the existence of two separate systems that influence perception and categorization of color, one of which is linguistically based and one of which is not affected by language. The third section critically evaluates recent claims by Kay, Regier, and Cook (2005) that there are similarities between color terms in the world's languages that point to the existence of color universals. In the fourth section, we examine children's color term acquisition in an attempt to trace the mechanisms by which color categories are acquired. It also discusses whether infants have an innate prepartitioned organization of color categories that is overridden during the learning process. In the two final sections, we outline some outstanding questions, note some methodological constraints on the conclusions that can be drawn from the accumulated evidence, and argue that much more empirical investigation is still needed in this field.

BACKGROUND TO THE DEBATE

Historically, the debate concerning color categorization has been sharply divided between what have come to be known as the universalist and the relativist positions. At one extreme, the view is that color categories are based on perceptual primitives, given either by the visual environment (Shepard, 1992) or by the properties of the human visual system (Kay & McDaniel, 1978), and are therefore universal. By contrast, the relativity hypothesis emphasizes the role of cultural needs in shaping both language and cognition: "the essential idea of linguistic relativity (is) the idea that culture, *through* language, affects the way we think, especially perhaps our classification of the experienced world" (Gumperz & Levinson, 1997, p. 612). According to relativists, therefore, color categories are learned and are likely to vary as a function of cultural and linguistic differences (Ratner, 1989).

The wide variability of color naming that exists in different cultures is clearly consistent with the relativist proposal that language can influence which categorical distinctions a child comes to develop (Lantz & Stefflre, 1964; Ratner, 1989). However, as we explain later in this section, advocates of the universalist view have contested this conclusion by claiming that linguistic color categories and the mental organization of the perceived color space can be completely independent of one another (Rosch Heider, 1972a, 1973; Bornstein, 1985).

The relativist[2] view of a close link between linguistic categorization and the cognitive organization of color was based on the findings of Brown and Lenneberg (1954). They showed that colors that were easier to name in English were easier for (English-speaking) participants to remember across short retention intervals and easier to communicate to others. Brown and Lenneberg, in common with many subsequent investigators, believed that memory performance provided insights into the way in which color categories are structured in the human cognitive system (we return later to the issue of whether this claim is justified). Cross-cultural support for the relativist view came from studies that demonstrated a close relationship between memory and naming in other languages (Lantz & Stefflre, 1964; Stefflre, Castillo Vales, & Morley, 1966). It was concluded that ease of naming would be a generally good predictor of memory accuracy in all languages across a wide range of color stimuli.

Berlin and Kay's (1969) survey of color terms in different languages led them to a radically different view. They suggested instead that some ways of organizing colors into categories were better than others. They proposed a common evolutionary trajectory for color vocabularies with the optimal arrangement as the endpoint. In that case, the number of color terms available in different languages reflected the point along that trajectory that a particular language/culture had reached. Western languages, such as English, had reached the endpoint of that evolutionary trajectory and used an optimal set of color terms and categories. Pressure for color vocabularies to evolve toward an optimal set would arise because all humans shared underlying cognitive representations of the optimal organization (black, white, grey, red, yellow, green, blue, pink, purple, orange, and brown), even if they did not express those categories in their language. This organization at a language-independent cognitive level was proposed to be innate and hardwired into the organization of color vision pathways (Kay & McDaniel, 1978; Bornstein, 1985). Saunders and van Brakel (1997) review these proposals in detail.

Berlin and Kay's (1969) hypothesis was supported by Rosch Heider's studies of a small remote branch of a hunter–gatherer tribe, the Dugum Dani (Heider & Olivier, 1972; Rosch Heider, 1972a, 1973, 1975). The Dani used only two basic terms for the whole range of visible colors [although Rosch Heider (1972b) did in fact suggest that many of her Dani participants used additional color terms]. Her results indicated that Dani memory patterns were not well predicted by their color-naming performance. Instead, they showed patterns of memory for colors similar to speakers of American English. For example, the Dani remembered colors that were good examples of English color categories (e.g., red, blue, green) better than colors that were hard to name for English speakers. These findings supported the proposal that a particular set of color categories might be panhuman cognitive universals that could transcend terminological differences. Under this view there could be such large differences between the "structure of the color space in memory" and the structure of the lexical categories used to describe them (Heider & Olivier, 1972, p. 351) that the two sets (one in language and the other in thought) could be effectively orthogonal.

A complete disconnection of this nature between thought and language would have widespread implications for theories of cognition generally. Under such a universalist view, no learning or transmission of cognitive color categories would be required (as everyone would have the same set from birth). However, speakers of languages that express a different (nonoptimal) number of categories would still at some point have to learn the appropriate reference set of exemplars for the linguistic terms that were used in their culture (Bornstein, 1985). Nevertheless, it remains unclear how a detailed category structure of this nature could be innately specified or how these differences between linguistic and cognitive categorization come to exist in certain cultures. It is also unclear why even English-speaking children appear to find color terms very difficult to learn (Bornstein, 1985) given that their language codes all members of the proposed universal categories.

Moreover, there are empirical as well as conceptual difficulties for this version of the universalist position. It is now clear that there are methodological problems with Rosch's experiments with the Dani that make her results difficult to interpret. Lucy and Schweder (1979) noted that the two sets of colors (best examples versus poor examples of English categories) used by Rosch were not equally discriminable because the arrays were ordered by hue and brightness. It turned out that the best examples had fewer close competitors surrounding them. Garro (1986) repeated Rosch's experiment with English speakers using a randomized array and found that best examples of English color categories were still better remembered than poor examples. However, these findings left unanswered the question of how the Dani would have performed with a randomized array.

To investigate this issue further, a new series of investigations was conducted that involved adult speakers of two other languages with a small number of color terms (Davidoff, Davies, & Roberson, 1999; Roberson et al., 2000; Roberson, Davidoff, Davies, & Shapiro, 2005). Both the Berinmo language, which is spoken in Papua New Guinea, and the Himba language, which is spoken in Northern Namibia, contain only five basic color terms compared to the 11 terms present in English (see Fig. 9.1). When recognition memory for color was examined in both these cultures, Rosch's results were replicated as long as the arrays were ordered by hue and brightness. However, when the array was randomized, and the number of close competitors equated for the best and poor examples, the Himba and Berinmo no longer showed any memory advantage for English best examples. Instead, speakers of each language recognized good examples of *their own* linguistic color categories better than poor examples, regardless of these items' status in English color categories (Roberson et al., 2005). Paired-associate learning of colors to pictures of

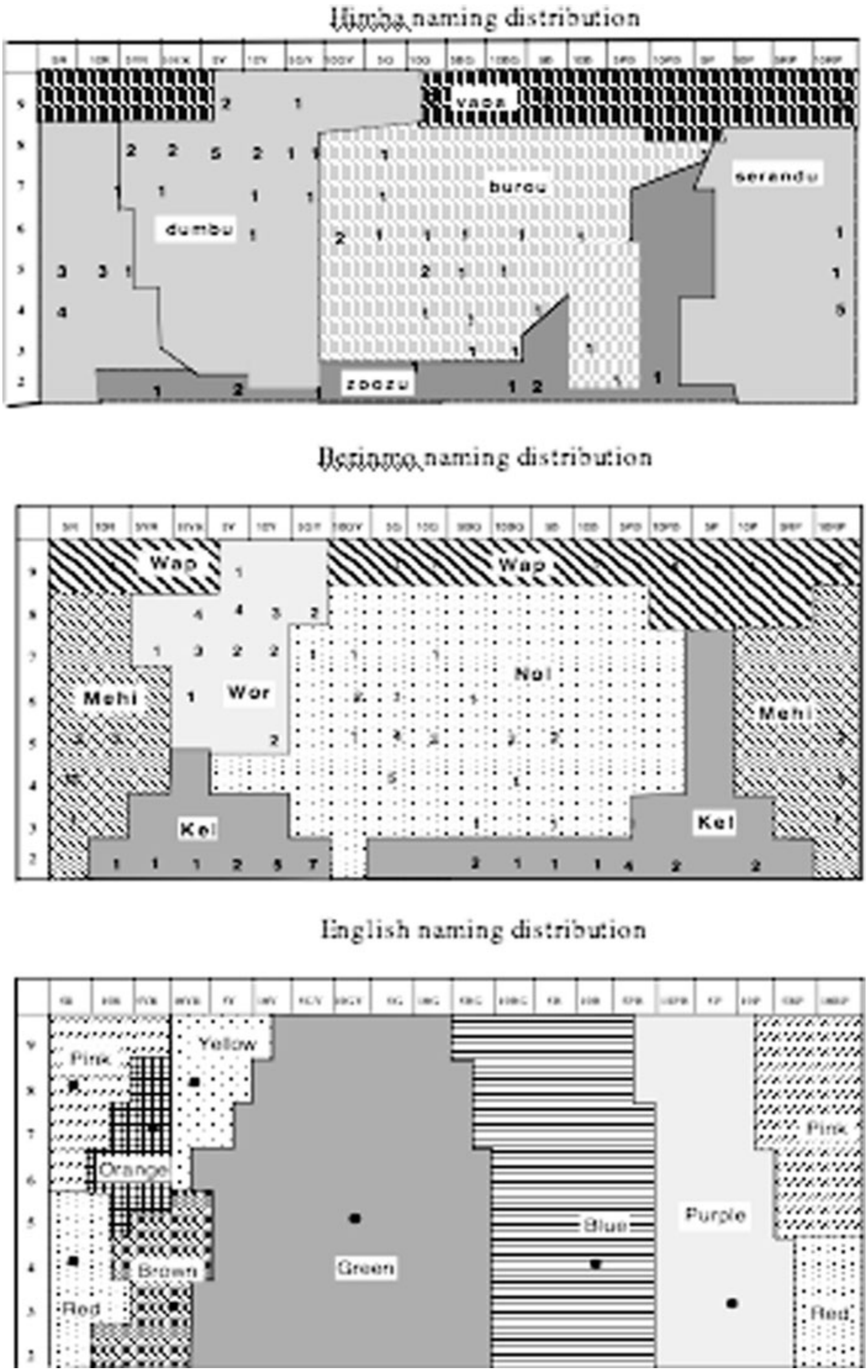


FIGURE 9.1. Distribution of Himba-named categories and choices of best exemplar for the 160 chip-saturated array (for 31 observers) compared to those of English and Berinmo speakers for the same array. Numbers represent the number of individuals choosing an exemplar as the best example of the category. Dots on the English graph represent the locations of best examples for English speakers.

familiar objects also failed to show any advantage for the proposed universal prototypical examples in either Berinmo or Himba speakers. A similar lack of preeminence for this particular set, either in naming or categorization, was reported by Jameson and Alvarado (2003) in Vietnamese speakers. These findings show that there is no single set of prototypical colors that is universally cognitively privileged. Rather, those stimuli that are best examples of an individual's own named categories are remembered more easily than those that are not.

In a further series of experiments, Roberson et al. (2000, 2005) investigated categorical perception (CP) of color in speakers of English, Berinmo, and Himba. CP refers to the sharp peak in the relative discriminability of colors that cross a category boundary compared to discriminability within a color category (Harnad, 1987) so that continuous quantitative differences along a continuum come to be perceived as discrete qualitative changes at category boundaries. For English speakers, it is claimed that pairs of colors that cross the boundary (e.g., between blue and green) are discriminated faster and more accurately than pairs of colors with equal physical separation that are both good examples of green or blue.

The issue investigated by Roberson et al. was whether speakers of Berinmo and Himba would also show CP at the boundaries of the English categories green and blue. Moreover, would the Berinmo and Himba show CP at category boundaries within their own language that do not exist for English speakers. Participants were shown a colored target and had to decide which of two stimuli, presented 5 seconds later, was identical to the target. For each language tested, performance was facilitated when the target and distractor stimuli had different color names (e.g., in English, a blue target with a purple distractor) relative to the same name (e.g., in English, two different shades of blue). The results indicated that all three groups of participants showed CP, but only at color boundaries that were explicitly marked in their own language. Crucially, there was no effect of the proposed universal boundary between green and blue for speakers of Himba and Berinmo whose languages do not make this distinction.

One criticism that has been made of the findings of Roberson et al. (2000) with the Berinmo has been that the speakers of this language live close to the equator and that as a consequence their eyes may have been prematurely damaged by strong ultraviolet sunlight (UVB) (Lindsey & Brown, 2002). As the lens ages, it becomes denser and less clear, a process known as "brunescence," which may particularly affect the ability to discriminate colors in the blue–green range. Lindsey and Brown (2000) gave young adults in the United States observation conditions (colored lenses) that simulated lens brunescence in the elderly and found that their naming classification of colors in the critical region changed. They suggested that individuals born with an innate boundary between green and blue, such as the Berinmo, could fail to distinguish them linguistically because by adulthood they would have lost their discriminative ability in that region. However, such an argument cannot explain why Roberson et al. (2000, 2005) found no evidence of any deficits when they tested the color vision of their Himba and Berinmo participants. Hardy, Frederick, Kay, and Werner (2005) provided the most direct test of Lindsey and Brown's hypothesis by examining older adults in the United States who were known to suffer from lens brunescence. They found that color naming for stimuli that were nominally green, blue–green, or blue was almost identical for observers with and without lens brunescence when viewing the same (unfiltered) stimuli. Thus, it seems that the effects of lens brunescence cannot explain the differences that Roberson et al. (2000) observed between the performance of Berinmo and English speakers.

RECENT STUDIES OF COLOR CATEGORIZATION IN ADULTS

Much of the research that we have discussed so far has employed memory tasks to investigate the relationship between language and color categorization. In these experiments, a strong link between naming and recognition might have emerged because individuals chose to rely on verbal coding to retain information about color during the retention interval. Even in the categorical perception experiments of Roberson et al. a color had to be held in memory for 5 seconds before the target and distractor items were presented. Consequently, these studies leave open the question of whether linguistic coding affects perception or the ability to retain a color in memory (Munnich & Landau, 2003).

Gilbert et al. (2006) devised a matching-to-sample visual-search task that appeared to make little or no demands on memory to investigate CP for colored targets. Participants were told to fixate on a cross in the center of the computer screen. They were then asked to report the location of an "oddball" colored target appearing among an identically colored array of distractors. Participants showed clear evidence of CP on this task. They were faster to detect a difference between the target and background when the target and background colors came from different categories (e.g., blue target and green background) than when both target and background came from the same category (e.g., different shades of blue) even when the amount of physical separation between target and background was held constant.

A critical question is whether CP on this visual search task occurs only at boundaries between colors in the putative universal set or whether it also occurs at boundaries that are not marked in English. This issue has been recently investigated with speakers of Russian by Winawer et al. (2007) and with speakers of Korean by Roberson, Pak, and Hanley (2008). The Russian participants of Winawer et al. showed CP at the boundary between *siniy* (dark blue) and *goluboy* (light blue), which are distinct "basic" color terms for speakers of Russian. English speakers, who would call all these stimuli "blue," did not show the same cross-category advantage. The Korean participants of Roberson et al. showed CP at the boundary between *yeondu* (yellow–green) and *chorok* (green), which are distinct "basic" color terms for speakers of Korean but not for speakers of English. No evidence of CP was shown by native English speakers at this boundary.

Because the experimental tasks were not tests of memory, these two studies provide a clear demonstration that categorical perception of colors is constrained by culture and language. Consequently, these two studies provide overwhelming evidence that superior discrimination of stimuli that cross a category boundary (such as that found for English speakers at the boundary between blue and green) does not provide evidence for a set of universal color categories that are hard-wired in the human visual system.

CP effects were once thought to reflect low-level visual processing (Bornstein & Korda, 1984). Is it therefore the case that these results mean that linguistic processing affects early stages of color processing? Two sets of additional findings have provided information about the precise point at which verbal codes influence color categorization in this experimental paradigm. First, Gilbert et al. (2006) and Winawer et al. (2007) showed that CP was not observed in perceptual tasks when participants carried out a concurrent verbal task. Under verbal suppression, all equally spaced separations of color were equally easy to discriminate. Second, Gilbert et al. (2006) and Roberson et al. (2008) reported that the CP effect was found only for colors that were presented in the right visual field, presumed to preferentially access language-processing areas in the left hemisphere. No difference between within- and across-category pairings of targets and distractors was observed for colors presented in the left visual field, which gains preferential access to the right hemisphere. Gilbert et al. (2006) also showed that CP was found only in the left hemisphere of a patient in whom the corpus callosum, which connects the two hemispheres of the brain, had been surgically severed.

Further evidence that left hemisphere brain regions associated with language processing are actively associated with perceptual processing of color has been provided by a recent functional magnetic resonance imaging (fMRI) study (Tan et al., 2008). Easy-to-name colors evoked stronger activation in areas associated with language than hard-to-name colors. Tan et al. suggest that these results support the rapid automatic activation of verbal color codes during perceptual decisions about color.

It is not hard to produce an explanation of the way in which left-hemisphere language processes might produce categorical perception. Let us assume that decisions about whether a target and background are the same can be made on the basis of either a right-hemisphere perceptual code or a left-hemisphere verbal code, and that when the two codes conflict, accuracy and speed will be reduced. Automatic activation of color names should therefore impair judgments about whether, for example, two different shades of blue are from the same category because the linguistic information that they are the same will conflict with the perceptual information that they are different. Decisions for items from different categories (e. g., a blue and a green) will be faster and more accurate because both linguistic and perceptual codes indicate that the target and background are different. When the left-hemisphere language system is suppressed by verbal interference, or is not accessed because information is presented directly to the right hemisphere, the verbal code is not generated and there is never any source of conflict with the perceptual code. Hence there is no advantage for comparisons that fall across linguistic category boundaries.

If this account of categorical perception is true, then it follows that the ability to decide whether two colors are different probably depends on right-hemisphere processing systems and does not require any form of verbal mediation. Color categories and color categorization, however, are entirely the product of the left-hemisphere language system. Because verbal and perceptual codes for color are automatically activated relatively rapidly, there can be errors if the codes conflict. But where these two sources of information yield congruent information; memory for colored stimuli is good and decisions about color can be made accurately and rapidly. It therefore appears that categorical perception of color, contrary to what has often been claimed, does not in fact reflect superior discrimination of colors when they cross a category boundary. Instead it appears to reflect the fact that decisions about color are hampered when perceptual codes and verbal codes are in conflict. This conflict occurs when a task requires that two different shades of a primary color must be treated as different even though they share the same label.

ARE THERE UNIVERSAL "TENDENCIES" IN COLOR NAMING?

Recent formulations of the universality hypothesis have acknowledged that differences in color category boundaries between languages influence memory for color and that linguistic boundaries determine the points at which categorical perception for color is observed (Kay et al., 2005). Thus, there is now some common ground between the universalist and relativist positions.

At the same time, however, Kay et al. (2005) still maintain that in different languages, there are strong universal tendencies both in color naming and in selection of the best examples of categories, which are held to cluster near the prototypes for English white, black, red, blue, green, and yellow (Kay et al., 2005). However, in place of the 11 originally proposed universal categories (Berlin & Kay, 1969) recent formulations suggest that there are instead universal tendencies in color naming. For instance, they propose that the naming systems of different languages that all use five color terms are more similar to each other than would be expected by chance. Regier, Kay, and Khetarpal (2007) suggested that instead of a single optimal system with 11 color terms, there might be optimal ways of dividing color space into 3, 4, 5, 6, etc. categories (see also Kay & Regier, 2007). They allow that these proposed "optimal" partitions could be based on some properties of perceptual color space (Jameson, 2005), on some properties of the visual environment (Yendrikhovskij, 2001; Shepard, 1992), or on sociolinguistic negotiation among speakers (Steels & Belpaeme, 2005). They also acknowledge that even though the number of languages surveyed to date is only a tiny fraction of the 6000+ languages still extant, many of them fail to fit the proposed optimal pattern.

There are, however, a number of methodological factors that may have inflated the apparent similarity of different naming systems in some investigations (Hickerson, 1971; Wierzbicka, 1999; Saunders & van Brakel, 2001). First, the range of stimuli used to collect cross-cultural naming data for the World Color Survey (Kay, Berlin, & Merrifield, 1991) contains 320 highly colorful (saturated) stimuli (from the outer "skin" of the Munsell sphere) that constrain the possible pattern of color labeling to a fixed set. Most traditional communities, lacking printing and dying facilities, may see such colorful stimuli for the first time when naming them for an experimenter. In some cultures speakers willingly extend their color terms to all these colors (e.g., Berinmo speakers). In other cultures, informants are more reticent, and many stimuli are left unnamed (e.g., Himba speakers; many South American informants interviewed by MacLaury, 1997). In addition, the possible extension of their color terms to less saturated stimuli (from the inner portion of the Munsell sphere) is rarely explored. Restrictions of the stimulus set may have led to an underestimation of the degree to which linguistic categories in one language might differ from those in another. In our own studies, the similarity between Himba and Berinmo naming patterns for the most colorful stimuli (0.61 interlanguage agreement) does not extend to less saturated stimuli (0.27 interlanguage naming agreement). Himba and English speakers use a large number of secondary terms to label less colorful stimuli (e.g., maroon, dun, olive, khaki) whereas Berinmo speakers readily extend basic terms to such stimuli.

In conclusion, reliance on the naming of only the most colorful stimuli may have led, in the past, to overestimation of the similarity of divergent languages' color term systems (Lucy, 1992; Lucy & Schweder, 1979; Saunders & van Brakel, 1997; Levinson, 2001). A similar observation of the limitations imposed by conducting cross-cultural naming studies with a restricted stimulus set is made in regard to cross-cultural studies of object naming by Malt, Sloman, and Gennari (2003).

Second, for simplicity, the naming maps commonly produced contain only the modal names given, and only those modal terms deemed by the experimenter to be "basic." Consequently, much individual variation in naming is lost both within and across languages. Reporting only modal names reduces the size of cross-linguistic discrepancies between areas in which name agreement is low or many participants fail to provide any name for stimuli (Jameson & Alvarado, 2003). It may also result in routine "regularization" of large and complex data sets in order to decide which terms should be counted as "basic" and included and which should be counted as "nonbasic" and excluded (see Saunders & van Brakel, 2001; Lucy, 1997). Indeed, some languages have been reported to have no "basic" color terms (Kuschel & Monberg, 1974).

Even if there are genuine similarities between certain color systems, do we really need to invoke color universals to explain them? There are obvious cultural factors that could explain at least some of these similarities. First, similar cultural needs, such as evolutionary pressure for successful frugivory (Sumner & Mollon, 2000; Komarova, Jameson, & Narens, 2007), could cause some category divisions to be more likely than others. The existence of dyes means that certain shades of color can be artificially generated much more easily than others, and may be labeled in many languages as a consequence. Second, it is clear that cultural contact between speakers of different languages has increased the similarity of the color categorization systems that these languages employ. For example, the introduction of the term "burou" for colors in the blue–green range into Herero and subsequently into Himba came directly from the German word *blau* during the time that Namibia was a German colony.

In conclusion, therefore, we believe that methodological problems with the way in which the data have been collected render unsafe any claim that there are universal tendencies that lead different cultures and different linguistic systems to divide up the color space in similar ways. Moreover, if there are genuine similarities between color categories in different cultures, it seems quite possible that they can be explained as easily by shared culture as by universal properties of the human visual system.

CHILDREN'S COLOR TERM ACQUISITION

A number of recent computational models of color category learning have highlighted the importance of communication between agents in establishing an optimal set of categories for a perceptual continuum of color (Steels & Belpaeme, 2005; Belpaeme & Bleys, 2005; Jameson, 2005). These studies have suggested that language plays an important role in initially establishing shared categories within a community, since simulations of category acquisition without communication among agents fail to speedily establish an optimal set.

Sometime between infancy and adulthood, children acquire a set of color terms and this linguistic categorization appears subsequently to influence their judgments of color. Given that color terms vary so widely across languages and that this variation has such profound behavioral consequences, the question arises as to when and how the differences come about. One possibility is that human infants are born with a prepartitioned set of cognitive color categories that are universal and innately specified (e.g., blue, green, yellow, red, and possibly also pink, purple, brown, and orange), but that these are overwritten during development by those categories in current use within the infant's culture and language (Bornstein, 1985). If so, given the evidence we have reviewed from color categorization in adults, language learning would appear to completely eliminate that original set.

Bornstein (1985) predicted that acquiring color terms would be more difficult for children learning a language in which an innate, hue-based, universal set must be overwritten by a new set, even if the new set contained fewer terms in total. Bowerman and Choi (2001) suggested that language acquisition would have to overcome great resistance in order to restructure mental life, where *any* robust and prepotent organization of the perceived world exists prelinguistically. Thus, the acquisition of a new set of named categories whose divisions cut across a proposed universal set should show a developmental pattern radically different from that of English-speaking children who would have to learn to map appropriate labels only to a set of already present cognitive categories.

Roberson et al. (2004) tested these claims by comparing the color naming and memory of young children in the UK who were learning English and children from Namibia who were learning Himba. Naming and comprehension were studied systematically over a 3-year period in order to establish a reliable measure of children's color term acquisition in speakers of different languages. They tested 28 English 3-year-old children before they entered preschool and, subsequently, through 3 years of formal education. They also tested 63 Himba 3-year-old children, few of whom received any formal education during the period of the study. Children's color term knowledge and memory for colors were tested at 6-month intervals over 3 years. The children completed a color term listing task ("tell me all the colors that you know"), color naming ("what color is this?"), color term comprehension ("can you find a red one?"), and a recognition memory task in each of the six testing sessions.

Despite the considerable environmental, linguistic, and educational differences between the two groups, the process of color-term learning appeared to be remarkably similar in the two groups. There was no predictable order of category acquisition across either group, consistent with other recent studies (Macario, 1991; Mervis, Bertrand, & Pani, 1995; Pitchford & Mullen, 2001; Shatz et al., 1996). Within each group individual children displayed almost every possible order of acquisition and, at the end of the study, there were still some children from both language groups who could not correctly use all their color terms (even though the English children had had 3 years of specific instruction).

In recognition memory, from an initial reliance on perceptual similarity, an influence of language categories became evident as soon as children acquired color terms. Himba and English children who knew no color terms showed similar patterns of memory errors and, critically, both patterns appeared to be based on perceptual distance rather than a particular set of predetermined categories. Of those children knowing one or more color terms at the first time of testing, Himba children showed better memory for the items that are good examples of Himba, but not of English categories, whereas English children showed the reverse pattern. Such rapid divergence in recognition patterns for the two groups, from the time that the first terms are learned, suggests that color categories in both languages are learned using similar mechanisms. These data, like those for adult Himba and Berinmo speakers, argue for a pivotal role of language in shaping color categorization. Considering the trajectory of color term acquisition in the two cultures, the longitudinal results suggested that children continue to refine the range of referents for each of their color terms for some years after they first show evidence of term knowledge for the best exemplars of categories.

For both populations, once color terms were acquired, memory performance was determined by the number of terms known. Children identified more examples of terms that they knew than of terms that they did not know, regardless of the absolute number of terms known. Knowledge of even one color term changed patterns of color recognition, and from this point on there were language-dependent differences between the two groups. In addition, the type of recognition errors made by each group of children diverged over time, so that more errors were made to best examples of the appropriate language categories than to other alternatives, even though some items that were best examples of a category in one language were poor exemplars in the other.

Overall, children from the two cultures seemed to acquire their color terms in the same gradual fashion, and knowledge of the appropriate terms influenced memory accuracy and memory errors so that the patterns of performance, from a common beginning, diverged increasingly over time. There was no evidence that English-speaking children learned their color terms more easily than Himba children, even though the English terms map directly onto a proposed innate set. Nor was there evidence that either group of children appeared to have a prepartitioned representation of color at 3 years of age, before they learn color terms (but see also Franklin et al., 2005).

Nonhuman primates have wavelength discrimination similar to humans (Sandell, Gross, & Bornstein, 1979; Matsuzawa, 1985). So, if there is a set of universal color categories, we might also expect to find evidence for their existence in these animals. However, a recent study of color discrimination in baboons (Fagot, Goldstein, Davidoff, & Pickering, 2006) failed to support the hypothesis that color categories are explicitly instantiated in the primate color vision system. In a match-to-sample paradigm, in which human participants showed a sharp category boundary between blue and green, none of the baboons showed any inclination to partition the range of blue and green stimuli on which they were trained into two categories despite good color discrimination.

Current evidence thus suggests that if there is an innate set of cognitive categories that is present in young infants, then (1) they are species specific and thus do not result from some property of the visual system that is shared with other primates, and (2) they are not retained once adult linguistic categorization is in place. An alternative possibility might be a scenario for color vision similar to that observed for auditory stimuli. In the auditory case, infants up to the age of 6 months appear to be sensitive to a wide range of categorical differences, including some that are not marked in their native language, such as the phonemic distinction between "*l*" and "*r*" for Japanese speakers (Werker & Tees, 1984). Their auditory system becomes selectively tuned to the appropriate categories for their native language sometime in the second 6 months of life. After 12 months of age, infants lose the ability to make some distinctions that younger infants make successfully. Such a possibility has not yet been investigated in infants with regard to color categorization.

Finally, a number of studies show that children achieve competent use of color terms relatively late compared to their acquisition of terms for other dimensions (Andrick & Tager-Flusberg, 1986; Mervis et al., 1995; Braisby & Dockrell, 1999; Sandhofer & Smith, 2001). Nevertheless, children appear to understand that color terms form an independent lexical semantic category by 2 years of age (Backsheider & Shatz, 1993; Sandhofer & Smith, 1999). At this stage, knowledge of color terms seems unrelated to the ability to use them. For example, when asked "tell me all the colors that you know," 3-year-old English and Himba children were as likely to list terms that they were unable to use correctly as terms that they could use correctly (Roberson et al., 2004). Learning color terms may promote selective attention to color, so children achieve a comprehensive conceptual representation of the color domain only after acquiring a sizable color vocabulary (Sandhofer & Smith, 1999). A difficulty in learning the referents for novel color terms might arise because the ability to abstract any object properties (color, size, form, and motion) develops slowly (Pitchford & Mullen, 2001) or because children are predisposed to attend to object shape when interpreting novel object labels (Smith, Jones, & Landau, 1992; O'Hanlon & Roberson, 2006). Whatever the explanation, it is clear that learning the appropriate set of referents for color terms is a more difficult task for young children than might have been expected than if they were simply learning appropriate labels for innately specified universal color categories.

OUTSTANDING QUESTIONS

Although it appears that categorical effects of color in adults depend on access to the language system, some studies have reported evidence of categorization of color in 4-months-old infants. Bornstein, Kessen, and Weiskopf (1976), Catherwood, Crassini, and Freiberg (1990), and Franklin and Davies (2004) have claimed that young infants do make apparently categorical distinctions of color continua. However, a study by Davidoff, Roberson, de Haan, and Davies (2007) found no differences in novelty preference for changes in color that were either within or across adult category boundaries. Much more research is needed to establish why minor changes in experimental paradigm produced such different results. If this evidence of prepartitioning proves robust, more research is also needed to examine how and why it is lost when language is learned.

What drives so many different cultures to arrive at even coarsely similar solutions to the problem of categorizing the continuum of visible colors? We should be wary of assuming that those similarities that do exist provide evidence for color universals before alternative explanations have been fully explored. If an eventual set of 11 basic categories was in some way optimal, why would some cultures maintain a small set of linguistic categories in their own language when surrounded by other languages that have larger sets? Why would Russians and Koreans develop additional basic color terms beyond those used in English? Whatever the origin of the observed differences between the color terminologies of different societies, any comprehensive model of color categorization needs to explain both the observed similarities and the differences between color-naming systems. The origins of linguistic color categories in different societies might be constrained by either cultural or environmental needs, or both, and both may change, over time, in different ways in different communities. For a discussion of these issues in domains other than color, see Nisbett, Peng, Choi, and Norenzayan (2001), Sera et al. (2002), and Wierzbicka (1999, 2005).

There remain several other outstanding questions that are both fundamental to the debate and beyond the scope of empirical investigations to date. Is the development of adult color categorization a unique case? If not, to what extent does it follow a similar pattern to other modalities that come to be perceived categorically? In studies of object classification, evidence from cross-linguistic studies best fits a hybrid model in which some broad, shared, nonlinguistic understanding of a domain combines with varying cultural pressures to differentiate particular aspects of a dimension at particular times in their history (Malt et al., 2003). Recent computational models of color category instantiation support the view that a combination of shared domain structure (in terms of both available learning mechanisms and the range of visible colors in the environment) and language is needed to explain shared color category structure. Such a combination of factors might lead to differences between linguistic categorization systems that also vary depending on the degree (and nature) of interactions between linguistic communities.

CONCLUSIONS

Early research in the field led to the conclusion that there are separate levels of categorical representation of color, one that is cognitive and impervious to language and another that is at a more superficial linguistic level (Rosch Heider, 1972a; Heider & Olivier, 1972). The evidence that we have reviewed in this chapter also points to the existence of distinct linguistic and nonlinguistic color systems. However, there is no evidence that the linguistic system is in any way superficial. Linguistic categorization in different languages and cultures partitions the same range of visible colors in different ways and these differences affect decisions about color even on visual search tasks. Evidence suggests that categorical effects in color perception and memory occur as a result of access to lexical codes for color in adults. Moreover, children appear to acquire adult-like patterns of discrimination and memory for color as soon as they learn the appropriate color terminology for their language and culture. This argues against the view that linguistic categorization of color is superficially overlaid on some more important cognitive structure.

When linguistic categorization is prevented (in adults), or is not yet in place (in the case of young children), participants behave as if they perceive an undifferentiated continuum of just-noticeable differences. There does, therefore, appear to be a separate nonlinguistic system (possibly in the right hemisphere) that can make extremely fine discriminations between colors and decide whether two colors are identical. There is no evidence that language learning has any effect on the way that this system processes color. We do not believe, however, that this system "knows" precise information concerning similarities and differences between two shades of color (e.g., that one is brighter than another, that one is more saturated than another, or that two different shades may share the same name). We do not believe, therefore, that this is the system that makes us see the rainbow as composed of seven distinct colors. Categorical knowledge of this kind is available only to the left hemisphere language-based color system, and people with different linguistic categories may well see a smaller or greater number of colors in the rainbow as a consequence.

Theorists who have supported the universalist position in the past now accept that linguistic differences between speakers of different languages influence color categorization (e.g., Kay et al., 2005). The relativist position has also been modified because theorists who support the relativist position acknowledge the existence of a separate color-processing system that is completely independent of linguistic influence. Nevertheless, some important differences still remain between the relativist position that has been put forward in this chapter and the universalist position put forward by Regier and colleagues in Chapter 8 in this volume. We believe that further investigation of the remaining controversial issues is needed. It is also important that dual process theories of color perception of the kind advocated in this chapter are subjected to rigorous empirical scrutiny. The domain of color has been and remains a fruitful ground for examining the relationship between language and thought. Now is not the time to discontinue the debate or the investigation.

ACKNOWLEDGMENTS

We are grateful to Barbara Malt and Phillip Wolff for helpful comments on an earlier draft of this chapter.

Notes

1. The criteria set by Kay, Berlin, and Merrifield (1991) for terms considered to be "basic" are terms that are monolexemic, present in the ideolect of every observer, and not subsumed within the range of another term.
2. Often reported as a Whorfian view, following the writings of Benjamin Lee Whorf (1956).

References

Andrick, G. R., & Tager-Flusberg, H. (1986). The acquisition of color terms. *Journal of Child Language, 13,* 119–134.

Au, T. K., & Laframboise, D. E. (1990). Acquiring color names via linguistic contrast: The influence of contrasting terms. *Child Development, 61,* 1808–1823.

Backsheider, A. G., & Shatz, M. (1993). *Children's acquisition of lexical domains.* Paper presented at the twenty-ninth regional meeting of the Chicago Linguistic Society, Chicago.

Belpaeme, T., & Bleys, J. (2005). Explaining universal color categories through a constrained acquisition process. *Adaptive Behavior, 13,* 293–310.

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

Bornstein, M. H. (1985). On the development of color naming in young children: Data and theory. *Brain and Language, 26,* 72–93.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 115–129.

Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research, 46,* 207–222.

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

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475–511). Cambridge, UK: Cambridge University Press.

Braisby, N., & Dockrell, J. (1999). Why is color naming difficult? *Journal of Child Language, 26,* 23–47.

Brown, R., & Lenneberg, E. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49,* 454–462.

Catherwood, D., Crassini, B., & Freiberg, K. (1990). Infant response to stimuli of similar hue and dissimilar shape: Tracing the origins of the categorization of objects by hue. *Child Development, 60,* 752–762.

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. M. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development, 14,* 241–268.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories of a stone-age tribe. *Nature, 398,* 203–204 and *402,* 604.

Davidoff, J., Roberson, D., de Haan, M., & Davies, I. R. L. (2007). Colour categorization in infants. Report on ESRC grant R000223894. http://www.esrcsocietytoday.ac.uk/

Fagot, J., Goldstein, J., Davidoff, J., & Pickering, A. (2006). Cross-species differences in color categorization. *Psychonomic Bulletin & Review, 13,* 275–280.

Franklin, A., & Davies, I. R. L. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology, 22,* 349–377.

Franklin, A., Clifford, A., Williamson, E., & Davies, I. R. L. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology, 90,* 114–141.

Garro, L. C. (1986). Language, memory and focality: A reexamination. *American Anthropologist, 88,* 128–136.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103,* 489–494.

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

Gumperz, J. J., & Levinson, S. C. (1997). Rethinking linguistic relativity. *Current Anthropology, 32*, 613–623.

Hardy, J. L., Frederick, C. M., Kay, P., & Werner, J. S. (2005). Color naming, lens aging and grue—What the optics of the aging eye can teach us about color language. *Psychological Science, 16,* 321–327.

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition.* Cambridge, UK: Cambridge University Press.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58,* 1021–1034.

Heider, E. R., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337–354.

Hickerson, N. P. (1971). Review of Berlin & Kay (1969). *International Journal of American Linguistics, 37*, 257–270.

Jameson, K. A. (2005). Why GRUE? An interpoint-distance model analysis of composite color categories. *Cross-Cultural Research, 39,* 159–204.

Jameson, K. A., & Alvarado, N. (2003). Differences in color naming and color salience in Vietnamese and English. *COLOR Research & Application, 29,* 128–134.

Kay, P., Berlin, B., & Merrifield, W. R. (1991). Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology, 1*(1), 12–25.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic colour terms. *Language, 54*(3), 610–646.

Kay, P., & Regier, T. (2007). Color naming universals: The case of Berinmo. *Cognition, 102*(2), 289–298.

Kay, P., Regier, T., & Cook, R. S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Science, 102,* 8386–8391.

Komorova, N. L., Jameson, K. A., & Narens, L. (2007). Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology, 51,* 359–382.

Kuschel, R., & Monberg, T. (1974). "We don't talk much about colour here": A study of colour semantics on Bellona Island. *Man, 9,* 213–242.

Lantz, D., & Stefflre, V. (1964). Language and cognition revisited. *Journal of Abnormal and Social Psychology, 69*, 472–481.

Levinson, S. C. (2001). Yélî Dnye and the theory of basic color terms. *Journal of Linguistic Anthropology, 10,* 3–55.

Lindsey, D. T., & Brown, A. M. (2002). Color naming and the photoxic effects of sunlight on the eye. *Psychological Science, 13,* 506–512.

Lucy, J. A. (1992). *Language and diversity of thought: A reformulation of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. A. (1997). The linguistics of 'color.' In C. L. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 320–346). New York: Cambridge University Press.

Lucy, J. A., & Shweder, R. A. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist, 81,* 581–605.

Macario, J. F. (1991). Young children's use of colour in classification: Foods and canonically coloured objects. *Cognitive Development, 6,* 17–46.

MacLaury, R. E. (1997). *Color and cognition in mesoAmerica: Construing categories as vantages.* Austin: University of Texas Press.

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

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

Matsuzawa, T. (1985). Color naming and classification in a chimpanzee (Pan-Troglodytes). *Journal of Human Evolution, 14,* 283–291.

Mervis, C. B., Bertrand, J., & Pani, J. R. (1995). Transaction of cognitive-linguistic abilities and adult input: A case-study of the acquisition of color terms and color-based subordinate object categories. *British Journal of Developmental Psychology, 13,* 285–302.

Munnich, E., & Landau, B. (2003). The effects of spatial language on spatial representation: Setting some boundaries. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought.* Cambridge: MA: MIT Press.

Nisbett, R. E., Peng, K. P., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108,* 291–310.

O'Hanlon, C. G., & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology, 94*, 275–300.

Pitchford, N. J., & Mullen, K. T. (2001). Conceptualization of perceptual attributes: A special case for color? *Journal of Experimental Child Psychology, 80*, 289–314.

Ratner, C. (1989). A sociohistorical critique of naturalistic theories of color perception. *Journal of Mind and Behavior, 10*, 361–373.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences, 104*, 1436–1441.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2004). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General, 133*, 554–571.

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. (2005). Colour categories in Himba: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50*, 378–411.

Roberson, D., Davidoff, J., & Shapiro, L. (2002). Squaring the circle: The cultural relativity of good shape. *Journal of Cognition and Culture, 2*, 29–52.

Roberson, D., Davies I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a Stone-age culture. *Journal of Experimental Psychology: General, 129*, 369–398.

Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right hemisphere is verbally mediated: Evidence from Korean. *Cognition, 107*, 752–762.

Rosch Heider, E. (1972a). Universals in color naming and memory. *Journal of Experimental Psychology, 93*, 10–20.

Rosch, E. H. (1972b). Probabilities, sampling and ethnographic method: The case of Dani colour names. *Man, 7*, 448–466.

Rosch Heider, E. (1973). Natural categories. *Cognitive Psychology, 4*, 328–350.

Rosch, E. (1975). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance, 1*(4), 303–322.

Sandell, J. H., Gross, C. G., & Bornstein, M. H. (1979). Color categories in macaques. *Journal of Comparative and Physiological Psychology, 93*, 626–635.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology, 35*, 668–679.

Sandhofer, C. M., & Smith, L. B. (2001). Why children learn color and size words so differently: Evidence from adults' learning of artificial terms. *Journal of Experimental Psychology: General, 130*, 600–620.

Saunders, B. A. C., & van Brakel, J. (1997). Are there non-trivial constraints on color categorization? *Behavioral & Brain Sciences, 20*, 167–178.

Saunders, B. A. C., & van Brakel, J. (2001). Rewriting color. *Philosophy of the Social Sciences, 31*, 538–556.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodriguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131*, 377–397.

Shatz, M., Behrend, D., Gelman, S. A., & Ebeling, K. S. (1996). Colour term knowledge in two-year-olds: Evidence for early competence. *Journal of Child Language, 23*, 177–199.

Shepard, R.N. (1992). The perceptual organisation of colors: An adaptation to regularities of the terrestrial world? In J. H. Barkow, L. Cosimedes, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Child Development, 55*, 363–380.

Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28*, 469–489.

Stefflre, V., Castillo Vales, V., & Morley, L. (1966). Language and cognition in Yucatan: A cross-cultural replication. *Journal of Personality and Social Psychology, 4*, 112–115.

Sumner, P., & Mollon, J. D. (2000). Catarrhine photopigments are optimized for detecting targets against a foliage background. *Journal of Experimental Biology, 203*, 1963–1986.

Tan, L. H., Chan, A. H. D., Kay, P., Khong, P-L., Yip, L. K. C., & Luke, K-K. (2008). Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences, 105*, 4004–4009.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech-perception—Evidence for perceptual reorganization during the 1st year of life. *Infant Behavior & Development, 7*, 49–63.

Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought and reality: Essays by B. L. Whorf*. Cambridge, MA: MIT Press.

Wierzbicka, A. (1999). "Universals of colour" from a linguistic point of view. *Behavioral & Brain Sciences, 22*, 724–725.

Wierzbicka, A. (2005). In defense of 'culture' *Theory & Psychology, 15,* 575–597.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104,* 7780–7785.

Yendrikhovskij, S. N. (2001). Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology, 45*, 409–417.

10 WORLDS WITHOUT WORDS

Commensurability and Causality in Language, Culture, and Cognition

Peter Gordon

Relations between words and the world can be profitably explored by correlating cross-linguistic differences with cross-cultural cognitive differences. By looking at cultures whose languages encode relations between objects and events in ways that differ radically from our own, we can determine whether there are biases toward encoding reality that reflect the preferred order within the linguistic system. The extreme case of such linguistic-cognitive correlation occurs when a language actually lacks words for a particular domain that is considered central to higher cognitive functioning in our own culture or vice versa. The present chapter examines the domain of numerical cognition in cultures that lack fully elaborated counting systems. I will focus on one such culture, the Pirahã of Lowland Amazonia, and will outline experiments on numerical cognition that attempt to show that linguistic impoverishment in this domain can lead to conceptual systems that are cognitively incommensurable. I consider the results of these studies in the context of research with other cultures that show numerical variability and consider claims about the causal role of language in determining the conceptual content of numerical cognition. I also consider how we can most accurately reconcile the intertwined roles of culture, language, and cognition.

Working with cultures that lack linguistic terms for everyday concepts that we take for granted is perhaps the strongest case that can be found in the attempt to discover whether there is a correlated effect on cognition. Such cross-cultural comparisons naturally bring to mind issues of linguistic determinism and relativity in the tradition of the Sapir-Whorf hypothesis (Sapir, 1921; Whorf, 1956). If the absence of words within a domain precludes the possession of basic concepts and the ability to reason within that domain, then this would represent one of the strongest cases for linguistic determinism. If true, then there would be evidence for the incommensurability of conceptual systems that possibly derives from differences in linguistic structure and/or function. In the original formulation of the Sapir-Whorf hypothesis, determinism and relativity focused on the question of whether language biases the manner in which people conceptualize the world around them. Neither theorist intended to claim that thought was either totally determined by language or that language was the medium for thought, in any sense. Rather, they characterized the hypothesis as describing how the habitual means of expressing an idea could differ from one language to another, and that this, in turn, would affect the way that piece of reality was differentially experienced or codified by speakers of the respective languages. In this sense, there was never a claim that thought was rigidly defined by the language, but only that the most common and salient means of expressing an idea in the native language would set up the default way in which experience would be biased. In Whorf's words:

> the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds—and this means largely by the linguistic systems in our minds. . . . we are parties to an agreement that holds throughout our speech community and is codified in the patterns of our language. (Whorf, 1956, pp. 213–214)

Notice that this passage does not suggest that we actually think in the same language that we speak—contrary to the common straw-man case that is often made against determinism. Whorf's claim here is simply that language acts as a filter on experience, setting up the appropriate categories that will be employed in codifying that experience. Despite the relatively benign tone of this passage, there are nevertheless points of contention. The ''kaleidoscopic flux'' reminds us of William James' (1890) ''blooming, buzzing confusion''—the idea that the mind of the newborn does not come innately endowed with the categories that bring order to the world. James believed that we must acquire the categories of thought that shape our perception of the world through experience. In the past few decades, many researchers in infant perception and cognition have endeavored to show that the infant's world is not quite so disorganized as originally believed (Baillargeon, 1994a,b; Bornstein, 1985; Spelke, 1998, 2000). Instead, theorists such as Liz Spelke have proposed that core aspects of perceptual and cognitive organization might be innately specified. Within this framework, innate organized mental structure is said to be as envisioned by Kant's idea of *synthetic a priori* knowledge—knowledge that encompasses empirical facts about the world prior to any relevant experience of such (Kant, 1848).

This Kantian idea of what we would today refer to as ''innate knowledge'' has informed much of the nativist perspective in cognitive science in the latter half of the twentieth century, especially following Chomsky's (1965) influential proposals about the nature of language. In proposing a nativist approach to linguistics, Chomsky was primarily concerned about the innateness of syntactic knowledge—how the form of the grammar might emerge from the mental structure afforded by the human brain rather than being acquired through experience. However, the idea that learning plays only a minimal role in development has also captured the imagination of those within the nativist camp who are interested in the acquisition of conceptual structure. Although it is difficult to accept the extreme form of conceptual nativism proposed by Fodor (1981; Fodor, Garrett, Walker, & Parkes, 1980), which suggests that all 50,000 concepts underlying the typical English vocabulary are innate, many are attracted to the idea that the conceptual repertoire is universal in some sense, and that cultures pick and choose from the same stock of concepts to deal with their everyday needs. The idea that concepts might be shaped by language is an idea to be resisted perhaps because it leaves too much to the vagaries of experience. If concepts come from a universal human stock, then perhaps that stock might represent the core knowledge of the human mental system that might be innately specified.

Whorf's second contentious claim in the previous quote is that mental organization is ''largely [organized] by the linguistic systems in our minds.'' Such a statement reads as if it were self-evident that the only avenue to conceptual organization is through language. In fact, the idea of language as a vehicle to conceptual organization is seen in other influential ideas from the mid-twentieth century. For example, Quine (1960) postulated that the ontological distinction between discrete objects and continuous substances emerged from experience with linguistic reference and quantification rather than being based on some language-independent ontological analysis of natural kinds. Like Whorf, Quine produced no independent empirical evidence for the conjecture that ontological distinctions were bootstrapped from linguistic quantification in this manner. Indeed, when experiments were carried out to examine such claims, it was found that children make ontological distinctions between objects and substances before acquiring the linguistic markers of such (Soja, Carey, & Spelke, 1991). The parallels between Whorf and Quine perhaps reflect a common zeitgeist of language as the prime organizer of human thought processes.

It is unfortunate that Whorf had a tendency to make assumptions about language and thought that were not independently validated and rarely justified beyond a kind of circularity that took the existence of linguistic differences as evidence for cognitive differences (see (Brown, 1970). Such criticisms eventually led to the almost wholesale rejection of many of the ideas of relativity and determinism within the psycholinguistic and linguistic community. The now famous lampooning of Whorf's claims about the number of words Eskimos have for snow is taken as a bitter lesson in the sloppy manner in which many of the claims were made (Pullum, 1989, 1991). Although this original critique pointed to a paper by Laura Martin (1986) claiming that Eskimos actually had only a few words for snow, in a more recent account, Pullum (1991) notes that it might be the case that some Eskimo languages do have many words for snow, but it depends on what we mean by ''Eskimo,'' what we mean by ''snow,'' and what we mean by ''word.'' In some languages, there could be an indefinite number of words for snow given the several hundreds of morphological variants for each noun (although these would not necessarily indicate different kinds of snow rather than case-marked forms of the same root and such). In addition, it is not clear that variability in the Eskimos' conceptualization of snow extends much beyond that found in English speakers who have words for snow, slush, snow drifts, sleet, packed snow, granular snow, and so on. Such oversights were not unusual in Whorf's style of linguistic analysis. For example, he claimed that Hopi Indians have no way to express time because such expressions were embedded within concepts of hope and intentionality (Whorf, 1956). However, one could equally argue that the ''will'' of the English future tense constructions (cf. *I will go*) and expressions including the adverb ''hopefully'' could indicate a similar manner of encoding time, but our own experiences suggest a less mystical and more mundane representation of temporal relations. Such problems speak to the leaps of logic that trip up the enterprise of actually investigating the relationship between language and thought.

Much of the earliest research in psycholinguistics, as it was developing an identity as a scientific discipline, was an attempt to evaluate claims about linguistic determinism. One of the first child language experiments by Roger Brown (1957) examined whether children could use the grammatical morpheme frames: *a X*; *some X*; and *X-ing* with nonsense words as variables, to indicate the referent as an object, a substance, or an action, respectively. This paradigmatic experiment for testing the acquisition of grammatical form classes was curiously referred to in its title as ''Linguistic Determinism and the Part of Speech,'' which speaks to the importance of the Whorfian framework at the time that seems odd in the present day.

The first earnest attempts to empirically test the Whorfian hypothesis were in the domain of color, primarily carried out by Eric Lenneberg and his colleagues. For example, Brown and Lenneberg (1954) did show significant differences in the encoding of color that appeared to arise from cross-linguistic differences. Zuni Indians, who do not distinguish linguistically between red and orange, performed less well on memory tests that required differentiation of these two colors. Similarly, both the organization of color space and memory for colors have also been shown to be affected by lexical organization in a series of recent studies by Davidoff, Davies, & Roberson (1999). They found that the language of the Berinmo tribe of Papua New Guinea carves up the color space in a very different manner than English and that such organization, in turn, affected the speakers' ability to remember colors that either were or were not differentiated lexically in the language. Berinmo speakers showed superior memory for color chips that crossed the partition of green into two lexical categories in their language. They showed no such memory advantage for chips that crossed the blue-green hue boundary for which there is no lexical distinction in the Berinmo language.

For many years, the retelling of this chapter in psycholinguistic history has been like a game of Chinese Whispers, where the initial consensus in the 1950s was that the data supported Whorf's claims, to the later consensus in the 1970s and beyond that determinism had been completely refuted (Lucy, 1992). The most commonly cited refutations of Whorfian theory were the studies of Heider (1972) and Berlin and Kay (1969). According to these authors, cultures that have fewer color words than English nevertheless shared the same conceptual organization in terms of the core and the periphery of the color space (Berlin & Kay, 1969; Heider, 1972). For example, even though speakers might lack color vocabulary, they had better memory for a saturated fire-engine red than one that was less prototypical of this color category. Later studies suggested that the results found in the original studies were an artifact of greater contrast in the proposed core colors, and when contrast levels were controlled for, there were no effects of the proposed universal status of the representation of color (core vs. peripheral) (Lucy & Shweder, 1979).

It is a bit ironic that what turned out to be Whorf's undoing from within the empirical sciences was actually based on a domain that he never claimed would be influenced by language. Whorf explicitly rejected the idea of linguistic influence on basic perceptual functions, which he referred to as ''irrational'' and therefore impervious to conceptual reorganization by linguistic or other means: ''Thinking may be said to be language's own ground, whereas feeling deals in feeling values which language indeed possesses but which lie rather on its boundaries. These are Jung's two rational functions, and by contrast his two irrational functions, sensation and intuition, may fairly be termed non-linguistic'' (Whorf, 1956, p. 66).

In the case of color, it has become clear there is an important distinction to be made about what kinds of linguistic effects might be expected to affect cognition. If language does have an effect, it would probably be in the categorical encoding of experience and the storage of such categories. Direct low-level perception is less likely to be affected. Having words for red, blue, and green is unlikely to affect the neural processes that send signals from the retinal color receptors down to the areas of visual cortex that process those signals for color perception. However, the organization of color in memory could well be affected. We can easily imagine effects of color naming on the neural processing of color categorization. For example, language could act as an attentional gate that would sharpen some color boundaries and blur adjacent regions of color space that were not categorically labeled as distinct colors in the language. This, in turn, could have long-term effects on how neurons might be dedicated to processing colors that are adjacent in the named versus the unnamed color space. This is basically a form of the use-it-or-lose-it principle that has been useful in thinking about neural development and cognitive function in general. In the same way that a speaker of one language might have a distinct neural organization depending on the categorical phoneme boundaries employed in the phonological system of their native language, so speakers of different languages could, in principle, have differential neural organization shaped by attentional factors associated with different lexical organization.

INCOMMENSURABILITY AND THE SAPIR-WHORF HYPOTHESIS

In considering the different versions of the Sapir-Whorf hypothesis I have proposed that linguistic differences can affect cognition, and moreover can actually lead to incommensurable conceptual systems between cultures (Gordon, 2004a). Specifically, I argued that the lack of number words in a language would lead speakers of the language to be incapable of conceptualizing the idea of exact numerosity. Although Whorf did not address the question of whether languages could be incommensurable, his mentor, Sapir, specifically denied the possibility that languages could even differ in this manner:

> The outstanding fact about any language is its formal completeness ... To put this ... in somewhat different words, we may say that a language is so constructed that no matter what any speaker of it may desire to communicate ... the language is prepared to do his work ... The world of linguistic forms, held within the framework of a given language, is a complete system of reference. (Sapir, 1924, p. 153)

The idea of the semantic completeness and equipotentiality of language is echoed by many prominent linguists and, for some, the assumption of equipotentiality seems almost axiomatic in cross-linguistic research (e.g., Ladefoged, 1982). The consensus that seems to have weathered all flavors of linguistic theory is that even if your language cannot express a concept in a direct way, at least you can get around this shortfall by circumlocution. So, if a language does not have a word for *molecule*, and its speakers do not have a scientific tradition to support such a concept, the concept could nevertheless be expressed with some phrasal definition such as *a very, very small thing that makes up stuff in the world*. In such cases, although the concept is not fully expressed, neither is it for many speakers of English whose scientific understanding is minimal. Lack of a fully elaborated concept of a molecule is not the result of an inability to speak English, but the lack of relevant scientific education. So, at all times in this endeavor, we must be careful to separate linguistic effects from effects of education and really focus on the concepts that we consider to be a core part of the language. As a test of this distinction, we can ask ourselves whether we expect all speakers of a language to possess a particular grammatical device or lexical domain. Someone who did not understand the basic grammatical morphemes and syntactic constructions of English would be suspected of being either a second language learner or of having some kind of language disorder. Someone who did not know the basic counting scheme up to some quite large number or the basic color words would be similarly suspect. On the other hand, lack of familiarity with the further reaches of the number scale for words such as *googolplex* is not taken as a failure to grasp the language. In a similar vein, someone who did not understand certain complex terminology associated with a specialized field of study would also not be suspected of a basic failure to learn the language. In summary, candidate cross-linguistic differences for thinking about relativity and determinism should be within this core set of items and structures in the comparative languages.

Can we find cases for which cross-linguistic differences account for conceptual differences that do not simply reflect cultural differences? Furthermore, can we find cases for which the differences are so extreme that they would actually violate the assumptions of equipotentiality and constitute a case of incommensurability? In the case of color words, the most recent research indicates that the conceptual color space is organized differently in different cultures for which, presumably, language is the determining factor (Davidoff et al., 1999). On the other hand, it is not the case that a particular color cannot be described by such speakers of languages that fail to lexicalize certain distinctions, nor that they cannot remember colors that fall outside the core color space of their linguistic labels. It is one thing for a color distinction to be more difficult to remember, it is another for such a memory to be impossible. For example, speakers of languages that organize colors differently could refer to noncore colors by compounding existing color words (orangish-yellowish blue) or by likening them to the colors of known objects. Such a strategy leads to the adoption of color words like *orange*, which originally named the fruit. To really show incommensurability, we need to examine a domain in which there is converging evidence that language plays a fundamental role in the conceptual foundations, and that lack of such linguistic experience cannot be ameliorated by circumlocution or such. We have to show that concepts within the conceptual domain simply cannot be entertained by speakers of a language lacking the appropriate labels.

Although I am not certain that numerical cognition in the Pirahã meets such lofty goals, I present what I believe is the best case so far for incommensurability and explore the criticisms since its first presentation.

THE CASE OF NUMBER

It is well known that there is significant variation across the world's cultures in the form and extent of number systems in language. We can characterize the range of counting systems across the world based on whether the count uses body parts for the counting base (fingers and/or other parts) or if it uses a small range of individuation or ''subitizing'' as a base for enumeration (perhaps based on working memory limitations of about two to four number words). Intersecting these two bases for counting, a second parameter involves whether the system uses recursion of some sort to extend counting beyond the limited range allowed by the base. Those cultures that use or used fingers as a counting base are revealed by being based in either 5 or 10, as in our own counting systems. Often such counting systems are fully recursive with potentially infinite count sequences, but sometimes they are not. For example, the Mundurukù (Pica, Lemer, Izard, & Dehaene, 2004) use five number words, some of which refer to fingers, but they are not recursive and stop at ''5'' (see below for further clarification on this point). Cultures such as the Yupno of Papua New Guinea that coopt other body parts such as facial landmarks, nipples, and toes tend to have counting ranges in the low 30s due to ending the count somewhere at the left testicle (at least in males). Sometimes a culture such as the Pirahã, with a small counting base in the subitizing range, will also use fingers for enumeration, but will not name the numerosities encoded by the fingers, and will therefore fail to employ words for numerosities beyond two to four items of the small enumeration range (see the following for a more detailed examination of this phenomenon).

Sometimes a culture with a small number base will actually employ a limited form of recursion over that number base. Menninger (1969) reports that the Gumulgal of Australasia had a recursive, base-2 counting system—*urapun (1), okasa (2), okasa urapun (21), okasa okasa* (22), *okasa okasa urapun* (221)—which led to a potentially infinite count sequence, even if such infinity was never actually realized in practice. In fact, it seems that such systems rarely count beyond a handful because the memory demands become overwhelming (Bender & Beller, 2007a,b). One of the more interesting side notes on the development of numerical systems concerns the origins of recursion. Bender and Beller (2007b) note that base systems might have emerged from the classifier systems that we find in most natural languages. Normally, classifiers are used to enumerate nouns according to the kind of thing they are, or the kinds of portions within which they are individuated (***pieces*** *of furniture,* ***bales*** *of hay,* ***plates*** *of food,* ***glasses*** *of water*, etc.) Bender suggests that in the cultures in the Pacific Islands, for example, counting of things such as coconuts was accomplished by arranging them in groupings of perhaps 8 or 20 items, and counting might have occurred all the way to the millions in the bustling trade communities of this region. Having established conventional groupings of these items, the groupings, as quantities, could then be referred to with numerical classifiers. So, 120 would be *six* ***twenty-units*** *of coconuts*. Through linguistic evolution such compounding could clearly evolve into a recursive and potentially infinite count sequence. As in the English count sequence, the derivation is relatively transparent (e.g., sixty = six tens). With the advent of quantification over chunked group-units, a fully elaborated recursive count system could be bootstrapped from the common linguistic classifier system that is not itself inherently recursive.

ONE-TWO-MANY

With respect to the range of numerical systems, let us consider what may be the simplest possible numerical system. These are systems based on small numbers that are not recursive. They are often referred to as the *one-two-many* systems and we will consider one such system found in the Pirahã culture. In the Pirahã language, the ''number'' words consist of *hói* (falling tone), *hoí* (rising tone), and *baagi*, which are roughly translated as ''one,'' ''two,'' and ''many,'' respectively. Such systems are not really systems of counting because people with such limited numbers are unlikely to look at an array of two nuts and tag them ''hói, hoí . . . Aha there are hoí nuts!'' In other words, such systems are not enumerative in the sense of providing an on-line tagging system. The quantities are so small that memory is perfectly able to keep track of such quantities without tagging them. What are such systems used for? As I noted previously (Gordon, 2004a,b), these terms tend to be vague in the numerical reference, and we often came across *hói* and *hoí* being used to refer to much larger sets of 4, 5, or more. Everett (2005) argues categorically that these are not number words, and the Pirahã do not have number words in any form.

The Pirahã word *hói*, which is sometimes translated as ''one,'' has a dual function and the same word is used to mean something like ''small.'' In particular, it is used in opposition to the word *xogi* meaning something like ''big'' or ''great in extent.'' The use of *hói* and *xogi* in opposition is quite general and can be used for comparisons between lesser and greater extents in many dimensions of comparison (Everett, 2005). Everett argues that *hói* never means ''one,'' it only ever means ''small,'' and *hoí* (~ 2) means something like: ''a bit bigger than small.'' My own use of *1-2-many* to describe the quantification system derives from the tradition within the linguistic anthropological literature, and from Everett's own initial description of the system (D. Everett, personal communication, 1990). Having pointed already to the imprecision of these terms as numerical tags, I am obviously not claiming that the meanings of these terms exactly map onto the English small number words. Instead, I suggest that the terms make a fuzzy distinction between discrete and continuous quantification. To the extent that the Pirahã terms quantify numerically at all, then such quantification is also fuzzy, perhaps in the same way that a carpenter asking for ''a couple of nails'' would not be terribly upset if he was handed three or four nails instead of two.

On the other hand, I think there is a question about what the Pirahã ''number'' words denote and whether they have any numerical function. Everett's basic claim is that *hói* and *hoí* have only the meanings related to general size, which is generalized to discrete quantities, and have no numerical function. My own impression (albeit based on about 25 years less experience than Everett) is that *hói* is polysemous in something like the vague way that words in a language can have several distinct but related meanings. In the way that the several uses of *go*, as in *going crazy*, *going to Vermont*, and *going commando* (not wearing underwear), are roughly related in the sense of marking transitions, they are also quite distinct meanings with quirky restrictions. In the case of *hói (~1)*, there are reasons to believe that its use in counting contexts is distinct from its use meaning ''small'' and in opposition to ''big.'' It is true that we, as speakers of English, seem quite happy with discrete quantificational terms such as ''*a few Xs*,'' which is a bit more than ''*a couple of Xs*,'' and ''*some Xs*,'' which is a bit more than ''*a few Xs*'' and a bit smaller than ''*several Xs*'' and so on. On the other hand, we do not have a term in English that means ''a little bit bigger than small.'' I think the difference here is that discrete quantifiers operate vaguely over exact numerical amounts, whereas size is always relational—what is tiny from one perspective might be massive from another, and so there is no standard ''smallness'' from which we can marginally increment size relations, except in the case where one has previously established a smallness by the previous use of the adjective. The occasions for such comparisons seem so rare as to make lexicalization seem extravagantly wasteful in a language that possesses only a few thousand words. Number, on the other hand, is not relational in this sense. Two atoms have the same numerical value as two solar systems and so there is no need to set up an initial item to contrast numerical quantity. In all likelihood, the phonological similarity between *hói* and *hoí* indicates that the relative meanings are established through some form of sound symbolism in which rising tone is somehow more highly valued than a falling tone and translates iconically into a greater value.

Hói and *hoí* are often used sequentially in discrete numerical contexts and such contexts are often followed by the word *baagi* (or *aibaagi*, or *baagiso*—as derivational variants) to denote numerical quantities that are larger than previously denoted. In addition, these terms are used spontaneously by Pirahã speakers in counting-like contexts (e.g., laying out objects one at a time in a line with no other function than to be enumerated) where it would seem unfelicitous to use size terms or general quantifiers such as *small, few*, and *some*. That is, if someone were to lay out objects, one at a time, accompanied by finger counting, it would seem odd to hold out one finger and spontaneously comment ''It's small,'' whereas the Pirahã do comment that a singular object is ''*hói*'' while indexing numerical quantity with fingers. There seems to be some kind of quasi-counting function to this activity, even though the count words are imprecise. Of course, I should caution that my argument rests on pragmatic assumptions that might not necessarily hold between the cultures.

A typical example of the use of these terms in counting-like contexts is found in a video, in which Keren Everett is doing a task in which she lays out lemons incrementing the array at each stage, and asking the informant to use fingers and words to enumerate that array after adding a single lemon (see Gordon, 2004b). In this video, the Pirahã participant uses *hói* for enumerating one lemon, and as more lemons are added up to four, each result of the addition is labeled *hoí*, at which point the participant uses the term *baagi* to indicate larger amounts (see Table 10.1). These terms are accompanied by finger gestures that show increasing error as the set size increases. There are several things to notice about this video. First, the informant never uses the term *xogi* to refer to larger numbers of lemons. If it were the case that the only meaning of *hói* (= ''small'') is in contrast to *xogi* (= ''big'') then surely the latter could be used to refer to big set sizes. But this never happens. Supporting this idea, in a recent study, Frank, Gibson, Fedorenko, and Everett (2008) report on an experiment in which they ran Pirahã participants in a numerical naming task and never reported any participants using *xogi* to name large numbers of objects. But taking Everett's account to its logical conclusion, the words *baagi* (= ''many'') and *xogi* (= ''big'') should be pretty much synonymous, as there is no dimension of number to be quantified that is distinct from overall size. Yet I only know of *baagi* being used in numerical contexts and *xogi* being used to denote generalized extent. Again, such claims are subject to further empirical testing.

TABLE 10.1. Use of Number Words and Finger Gestures in Lemon-Counting Task

Lemons	Number Word	Fingers
1	hói	
2	hoí aibaagi	2
3	hoí	3
4	hoí aibai	5–3
5	aibaagi	5
6	aibaagi	6–7
7	hói	1–8
8		5–8–9
9	aibaagi	5 –10
10		5

There is another striking impression about this video, which is that the subject appears to be indicating numerosity in coordinating the words *hói, hoí*, and *baagi* with the display of fingers to mark the relevant quantities. Note that if we were presented with the same task, and were constrained not to use number words, it seems unlikely that we would say, ''there's a few,'' ''there's some,'' ''there's several,'' and so on. There is something distinctly reminiscent of enumerative-like behavior in this scene, but the exact meanings of the terms remain a bit of a mystery.

For the reasons noted, I suggest that *hói* is polysemous, but have never insisted that the term is directly comparable to the English translation of ''one.'' Hói and hoí are definitely vague when it comes to their numerical denotation. Therefore, strictly speaking, the Pirahã language does not have number words if we take that to mean words that have exact numerical values. On the other hand, I do suggest that *hói* is polysemous and has a distinct quasi-numeric function when used within the context of *hoí* and *baagi*, and a more continuous function when used in the context of the *hói-xogi* (small-big) opposition.

I suspect that such a state of affairs is the rule rather than the exception when it comes to counting systems referred to as ''1-2-many'' in different cultures. In fact, we know from the work of Pica et al. (2004) that the reference to numerical quantities in the five number words for the Mundurukù is equally vague. Researchers who have worked with cultures using small number systems suggest that numerical denotation is generally imprecise in these systems (Hammarström, 2008). Perhaps this makes sense. With such a limited range, it is unlikely that the numbers are actually used for counting in any useful way in any of these cultures. Instead, they might be used to indicate relative amounts within a small range of items. Because mathematical precision is not at a premium, it seems likely that meanings can be quite imprecise. Fortunately, there are now clearer data on how the *hói* and *hoí* are interpreted by Pirahã speakers. In the study mentioned, Frank et al. (2008) had Pirahã informants name quantities of objects from 1 through 10, which were incremented by one for each trial. They were then asked to do the same in reverse order for quantities 10 through 1, this time decrementing the total by one on each trial. The results are shown in Figure 10.1. The data show that, in the context of the array that is incremented by one each trial, *hói* (~1) is the only term used to denote singular objects and is never used for any other numerical value. *Hoí* (~2) is the only term used to denote quantities of exactly 2, and is used about one-third of the time to denote values 3 to 10. Finally *baagi* is used about two-thirds of the time for values 4 through 10 and just under one-half of the time when the quantity was 3. When magnitudes were decremented, starting with 10 and ending with 1, *baagi* was used about two-thirds of the time for 10 and 9, was used less than half the time for 8 and 7, and was not used for any smaller values. *Hoí* (~2) complemented the use of *baagi* for 10 through 7, then shared 6 through 4 with *hói* (~1), which, in turn, gradually took over the majority of uses

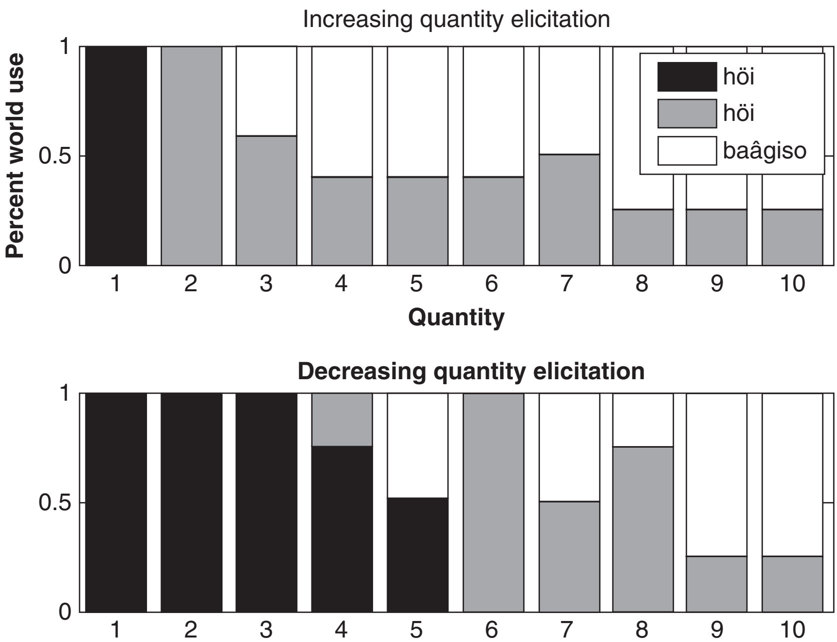


FIGURE 10.1. Pirahã use of hói, hoí, and baagi in describing numerical quantities of increasing and decreasing set size.

(From Frank et al., 2008.)

until it was used exclusively for values 3 through 1.

Why is the numerical designation fixed at ''one'' on the ascending task but fuzzy on the descending task? It is possible that these terms actually contain a mixture of ordinal and cardinal functions that are not well defined. In other words, when a single item is presented in an ascending count, it is both ''first'' and small in number. The next trial with two items is then designated as *hoí* or the ''next'' in numerical quantity, when compared to *hói*. If *hoí* is some combination of ''next biggest'' and ''a few,'' and *baagi* means something like ''some'' or ''many,'' then this would cover the uses found in these data. Because the designations are relative to an ascending count, we see that, in the descending count, the designations are undefined in the absence of an ascending sequence and what we see is something akin to guessing. Of course, this is not the only interpretation of these data, but we must account for the asymmetry between ascending and descending sequences in the meanings of these terms and something of this sort seems to fit the bill for now.

PIRAHÃ NUMERICAL COGNITION

In my own studies of Pirahã numerical cognition, independent of language, I examined whether members of the tribe were able to perform simple numerical quantification tasks that did not require counting. Instead, I asked them to indicate quantities by matching tokens to objects or making judgments about quantities of hidden objects inside containers. The data from these tasks are given in Figure 10.2. The tasks were designed so that the task demands in terms of visual-spatial transformation and memory representation should make them increasingly difficult. In the simplest tasks, I was interested in whether the participants could do a simple one-to-one mapping task of the sort found in some of Piaget's (1952) earliest studies on numerical abilities in children. In this task, I sat across the table from the participant, with a stick dividing my area from theirs. I laid out a certain number of AA batteries orthogonally to the stick (see the diagrams in Fig. 10.2) and asked the participant to match the array and make it the same. Frank et al. (2008) recently ran a similar task with somewhat different stimuli. Instead of matching arrays of batteries in my procedure, they had participants match deflated balloons on one side with cotton reels on the matching side.[1]

In contrast to the errors for larger numerical quantities committed by participants in my procedure (see top left graph in Fig. 10.2), Frank et al. (2008) found that performance was almost flawless on their version of the task. In some ways, the outcome of Frank et al. is less surprising than mine. The 1-1 matching task is not a task that requires the participant to develop a representation of the cardinality of the whole set. Rather, the task can be done incrementally, with no reference to the set as a whole, requiring only that one start matching at one end and stop at the other. Why then would the participants have made errors on my version of the task? A clue lies in the different items used in the two versions of the task. In my version, the AA batteries were prone to be disturbed by the wobbly table, and they would roll around. Exact placement in a 1-1matching task with the corresponding battery was therefore less stable, and any misalignment of the batteries could lead to error. Presumably, misalignment was more likely with a larger number of items. On the other hand, the deflated balloons and cotton reels used in the procedure by Gibson et al. had flat surface that does not move. In some ways, the difference in performance cuts at exactly where we should suspect there to be problems. When an array of items is unstable, the only way to ensure that the task has been done correctly is to have a sense of the cardinality of the target set and the matching set. If that concept of equivalence of cardinality is missing from the conceptual repertoire, then such differences will not be noticed. In some ways, this is like a number conservation problem in which the spatial configuration of two arrays is disturbed and the numerical equivalence must be inferred rather than aligned (Piaget, 1952). At this stage, I also speculate that keeping track of 1-1 equivalence might also

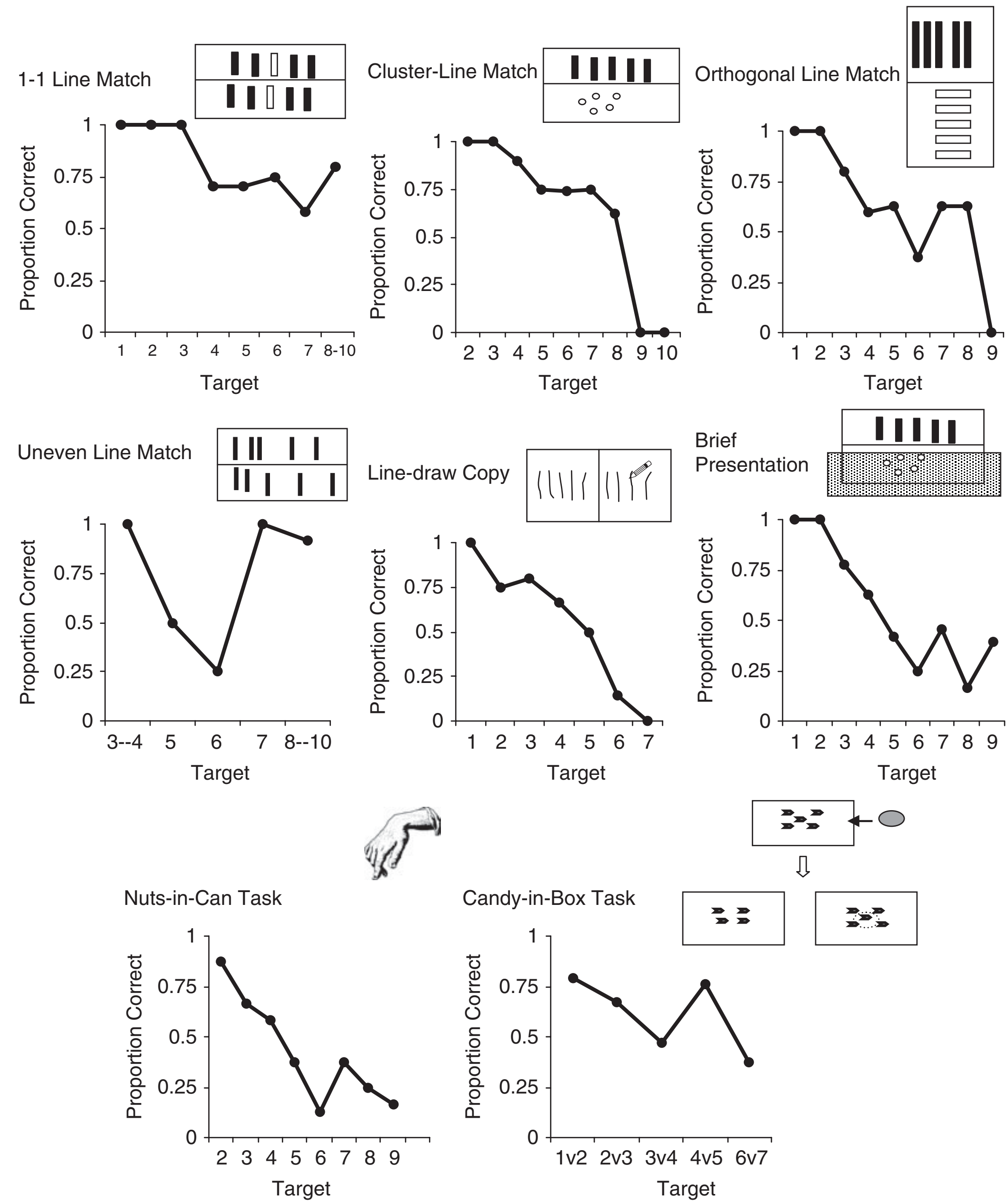


FIGURE 10.2. Performance of Pirahã participants on number-matching tasks.
(From Gordon, 2004a.)

be easier when the two arrays are distinct, containing different kinds of items. Somehow, when dealing with up to 20 AA batteries (10 in the target set and 10 in the matching set) such a task might be more confusing than dealing with visually distinct arrays of items such as cotton reels and balloons that clearly mark the target and matching sets as distinct and could possibly set up distinct representations.

Other tasks in my study also required a sense of cardinality either to compensate for visual–spatial realignments such as the cluster-match task (matching the battery line to a cluster of nuts) or the orthogonal matching task (matching a horizontal to a vertical line of batteries). In others, cardinality was required to form a memory representation of the target set. These include, for example, the brief presentation task, in which a target group of nuts was initially obscured by a screen, then revealed to the participant for about 1 second and then obscured again. The participant was then asked to reconstruct the number of items in the array with the line of batteries. Because the participant cannot inspect the target set, the only representation is one that would be encoded in memory. If such a representation includes no encoding of exact numerosity, then reproducing the numerosity would be increasingly more difficult with increasing set size, according to the standard psychophysical laws of magnitude estimation. Other memory tasks in my study included the nuts-in-a-can task and the candy-in-the-box task. In the former, a collection of nuts was first shown to the participant for extended inspection. Then, each nut was placed inside a can one at a time. This was followed by the removal of one nut at a time from the can, during which the participant was asked to state whether there were still more nuts in the can or if it was empty. Again, without a representation of the exact numerosity, guessing when the can was empty would become more difficult with increasing set size.

In the candy-in-the-box task, the participant was shown a small box with a picture of some number of fish on it. A candy was placed inside the box; I shuffled the box behind my back along with a second box, which had a picture with either one more or one fewer fish compared to the target box. If participants chose the box with the candy, they were rewarded by being able to keep (and eat) the contents. Despite the attraction of the candy, participants performed poorly on this task at all target values, performing at about 75% with a chance value of 50% (see Fig. 10.2).

Taken together, the results show that the Pirahã perform poorly on tasks in which a representation of cardinality is required for quantities greater than about 2. The exact threshold cut-off value and the slope of error rate plotted against set size depended on the complexity of the task. To the extent that there was a regular relationship between set size and accuracy, this suggested that this function was a typical magnitude estimation function found in the psychophysical literature based on Weber's law. Analysis of performance in these tasks with respect to measures of magnitude estimation revealed a rather interesting pattern. Even though the Pirahã performed quite poorly on these tasks when numerical values exceeded about two items, their mean response values almost exactly matched the target values. In other words, even though they were likely to be incorrect on larger valued arrays, their errors were evenly spread above and below the target value in equal measure. In accordance with Weber's law, the variability of the responses increased in proportion to the increase in target set size. In other words, the larger the target value, the larger the mean error. The ratio of target size to error is captured in the coefficient of variation (CV) (the mean divided by the standard deviation). According the Weber's law for magnitude estimation, this ratio should be constant. For the Pirahã, the CV was constant at about 0.15 across tasks and subjects when target values were greater than 2 (see Fig. 10.3). Interestingly, this same CV value is the same as that found in numerical estimation tasks performed by American undergraduates during concurrent vocalization (Whalen, Gelman, & Gallistel, 1999). Although the CV can vary depending on the particular task, the fact that the value for the Pirahã was in the same range does suggest that there might be similar processes of numerical estimation at play in both cultures. In general, the results clearly show that magnitude estimation within the domain of discrete individuals is preserved despite the lack of number words in the language.

In addition to the 1-1 matching tasks described, Frank et al. (2008) replicated the orthogonal matching task, the brief presentation task, and the nuts-in-the-can task with their Pirahã participants from a different village. Unlike the conflicting data in the 1-1 matching task, the results of these more complex tasks were highly similar to those

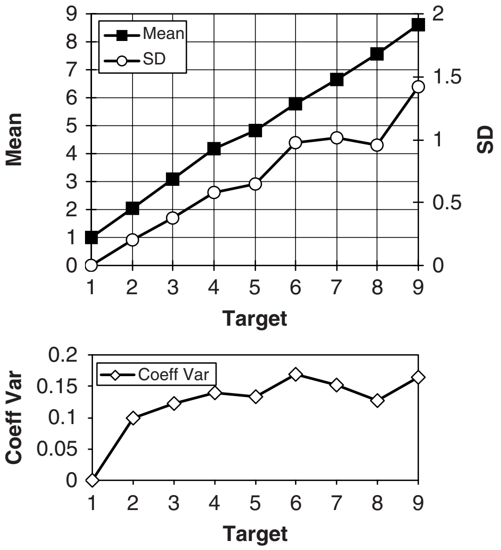


FIGURE 10.3. Mean, standard deviation, and coefficient of variation for number-matching tasks with Pirahã participants.
(From Gordon, 2004a.)

obtained in my own studies with a similar CV value.

INCOMMENSURABILITY AND CAUSALITY

We return then to the original question about whether the data from Pirahã numerical cognition provide a case for incommensurability of cognitive systems, and whether we can attribute such incommensurability to linguistic factors viz. the lack of exact number words in the Pirahã language. If this is a case of conceptual incommensurability then we should have obtained convincing evidence that the Pirahã do not have a means to conceptualize exact numerical cardinalities that would be encoded by number words in a numerate language such as English. Beyond evidence of the ability to encode values in the small number range of three or less, we do not see evidence that the Pirahã can encode exact cardinalities cognitively. That is, when the task requires a representation of larger numerical values, the Pirahã are restricted to a form of approximation seen in experiments in which verbalization of number words is prevented by means of a rapid presentation of items and/or concurrent vocalization (cf. Whalen et al., 1999). To the extent that there is an inability to conceptualize exact number on the part of the Pirahã, it can be said that there is, indeed, incommensurability between conceptual systems. This does not mean that there is no hope that the Pirahã can learn to count, only that such abilities are outside the range of abilities within the everyday existence of the culture. Whether the Pirahã can learn to count is provocatively alluded to in reports of the Everetts when they described attempting to teach Portuguese counting to Pirahã adults and children. The children had no problems in learning to count, but the adults were apparently befuddled and gave up quite quickly (apparently insisting that the children stop attending the classes, too). Although this was not a controlled experiment, the outcome does suggest that the inability to conceptualize numerical values goes quite deep in the adults, perhaps bearing indications of a critical period of some sort.

The second question relates to the causal structure of the cognitive pattern here. Recently, Pinker (2007) argued that limitations of Pirahã numerical cognition cannot be attributed to linguistic origins. He instead attributes the lack of number words and the lack of exact number sense to the same common cause, which is the simplicity of the culture. He suggests that if language were the causal factor in limiting Pirahã numerical abilities, then the Mundurukù (Pica et al., 2004), who have five number words, should be able to conceptualize quantities from 1 to 5 accurately:

> Could there be a control group for the Pirahã—a people whose culture was similar to theirs but whose language differed in its inventory of number words? Such a culture would afford a true test of Linguistic Determinism unconfounded by culture... The Mundurukù are also an illiterate hunter-gatherer people in Brazilian Amazonia, but their language has number words up to five. This is not, however, enough to grant them exact number concepts up to five... The Mundurukù, like the Pirahã, used number words ... approximately ... [their] ability

to visualize the result of a subtraction ... was imperfect for numbers greater than three, and got worse for larger and larger numbers. So the presence of additional number words in their language did little or nothing for their exact number sense. (Pinker, 2007, p. 140)

It is interesting that Pinker puts such store in the Mundurukù study as being the final arbiter of the case. Even though I might disagree with the claim that the cultures are exactly comparable and that they represent a control group for the Pirahã,[2] it is instructive to look closely at the data from Pica et al. (2004), which is reproduced for one of the tested groups in Figure 10.4. These data are from the subtraction task that Pinker refers to in Pica et al., which is the one test that requires exact enumeration. Even though Pica et al. draw their own interpretation of the trend line in these data, when we look at the actual data points, we see that the Mundurukù do remarkably well for values 1 through 5 at about 80–90% correct, with a precipitous drop in performance for values 6 through 8.

Not only do these data invalidate Pinker's argument, but the argument itself was specious to begin with. As Pinker points out, the meanings of the so-called number words 1 through 5 do not denote exact numerosities, just as the Pirahã number words are inexact. In

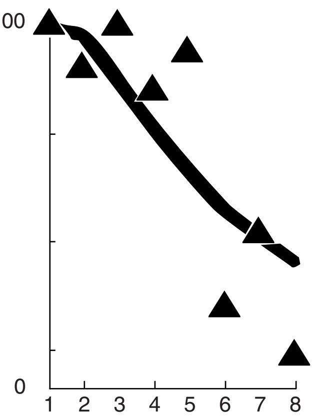


FIGURE 10.4. Performance of Mundurukù participants on a subtraction task.
(From Pica et al., 2004.)

fact, Pinker takes this inexactness as evidence that having number words in the Mundurukù language for 1 through 5 does not guarantee that they will learn these exact numerical concepts, and therefore that the lack of numerical abilities in both tribes cannot be causally attributed to language, hence there is no Whorfian effect here. But if Mundurukù speakers allow the words that are interpreted by Pica et al. as meaning ''three,'' ''four,'' and ''five'' to each span values between 3 and10, then how can we say that these words have exact numerical meanings but those meanings do not match their reference? It is as if we were suggesting that God had given the tribe these number words and they just could not learn them. Surely these words could not ''mean'' what they are claimed to mean if they do not denote exact quantities when they are used. If it doesn't quack, then it probably isn't a duck. If we make a tentative assumption that such terms mean something similar to our vague quantifiers such as *few, some, several,* and *many,* then it is not reasonable to expect such terms to denote exact quantities in the speakers.

In rejecting the linguistic explanation for lack of numerical abilities in either the Pirahã or the Mundurukù, Pinker echoes Casasanto (2005) and Everett (2005) in stating that the cause of innumeracy is not the lack of number words, but the lack of a cultural need for counting. He thereby explains both the lack of number words and the lack of number concepts in terms of a common causal cultural root:

It can't be a coincidence that the Pirahã language just happens to lack big number words (unlike the English language) and the Pirahã speakers just happen to hunt and gather in remote stone-age villages ... A more plausible interpretation is that lifestyle, history, and culture of a technologically underdeveloped hunter-gatherer people will cause it to lack both number words and numerical reasoning. (Indeed Daniel Everett, the linguist who studied the Pirahã for twenty-three years, rejected Gordon's conclusions and attributed their limitations ... to general patterns in their culture.) (Pinker, 2007, p. 139)

When Pinker cites Everett in supporting his case for cultural causation, it is unlikely that this represents an endorsement of the actual claims that Everett makes about the role of culture in cognition. Everett's proposals represent an extreme form of functionalism. He claims that facts about the Pirahã language and culture invalidate many Chomskian claims about formal language universals. Given Pinker's general support for the Chomskian modular and nativist perspective on language, his reference to Everett's position is hardly likely to be an endorsement of the actual claims about cultural causation. Everett rejects the idea that grammar can be explained from an autonomous formal perspective. Instead, he claims that the Pirahã case demonstrates that language structure can be understood only within the context of the whole cultural milieu. To whit, he claims that a variety of phenomena related to Pirahã language and culture can all be explained by a cultural prohibition against speaking of anything that goes beyond immediate experience (these phenomena include—according to Everett—the absence of creation myths and fiction, the simplest kinship system yet documented, absence of numbers and of a concept of counting, absence of color terms, absence of embedding in grammar, absence of relative tense systems, lack of a native pronoun system, being monolingual despite 200 years of contact, absence of collective memory greater than two generations, absence of drawing or art, the simplest material culture documented, and absence of quantificational terms). According to Everett (2005): "Pirahã culture constrains communication to non-abstract subjects which fall within the immediate experience of interlocutors" (p. 621). Although there is much to be debated about each of these claims about Pirahã culture and cognition (see Nevins, Pesetsky, & Rodrigues, 2009; Everett, 2009), the more general claim that the Pirahã cannot speak of anything outside of their immediate experience strikes me as implausible based on my own experiences with them. In any case, the link between this claim and the limitations on numerical ability seems obscure.

The more general position taken by Casasanto and Pinker is the claim that technological cultures cannot survive without numerical language and cognition. In stating the proposition that change would occur if technology developed (or perhaps that technology could not develop without a number system) it does not follow that the lack of such technology is what causes the lack of numerical concepts. In fact, it seems clear that technology is not the singular necessary condition for the development of numerical systems of reference, as many nontechnologically advanced cultures have advanced counting systems. In addition, Beller and Bender (2008) have pointed out that the evolution of numerical complexity is not unidirectional. That is, there is no universal cultural progression whereby underdeveloped cultures start out without exact counting systems, and exact counting invariably develops in response to technological innovations in the culture. In the history of Polynesian and Melanesian cultures, it was not so much technology that led to the development of complex counting systems, but the extensive trade and travel between islands that required exact quantification of goods. However, when groups moved to locations where trade was no longer as significant, the people in these new habitats, who had once used a complex number system that enumerated in the thousands, fell back onto a simple two-base counting system.

It could be argued that this is a case of cultural need driving numerical language (albeit backwards), but such relationships are not being denied here. What I am questioning is whether referring to vague relations between culture and technology constitutes a causal explanation for the lack of exact numerical cognition. Furthermore, if trade rather than technology is the cultural variable that pushes the development of numerical cognition, then the Pirahã should have extensive numerical cognition in spades. They trade regularly with anyone they come into contact with including other tribes, researchers, missionaries, and riverboat traders. Such trade just does not occur with exact enumeration, but by simple bartering, which often means that the Pirahã are exploited by the unscrupulous riverboat traders. If one accepts that Polynesian cultures developed elaborate counting systems as a result of trade rather than developing technology, then cultural causation should dictate that extensive trading should be a sufficient condition for exact numerical cognition to develop in any culture. Yet, the extensive trading behaviors of the Pirahã are not sufficient to prompt them in this direction.

On the other side of the cultural causality hypothesis, Casasanto (2005) and Pinker (2007) argue that the Pirahã have no concept of exact number because they are hunter-gatherers. This implies that being a hunter–gatherer is a sufficient condition for not having exact number concepts. In fact, this is demonstrably not true. In a survey of hunter–gatherer cultures and their number systems, Hammarström (2008) documents 85 that have small-number based systems (e.g., 1-2-many), 76 that have full counting systems, and 35 that are somewhere in between (small number systems with additional use of body parts). It is true that all cultures that have small-number systems are hunter-gatherers. In other words, being hunter-gatherer appears to be a necessary condition for lack of exact number systems, but it is not sufficient. Therefore, it cannot be the cause of such.

At a more general level, the problem with the debate about causal explanation, which I attempted to address in my response to Casasanto (see Casasanto, 2005; Gordon, 2005), is that there is a certain level at which it makes no sense to argue whether it is lack of words or lack of cultural need that causally explains the lack of numerical concepts in the Pirahã. The question of what causes a lack of number concepts is ill formed because it seeks to find a cause for nonexistence. To illustrate the problem here, consider the question of what is the cause of the Pirahã not being able to play chess? One can imagine many counterfactual cultural scenarios in which the Pirahã might be introduced to chess: If missionaries had taught them to play chess, if their culture were more advanced, if they watched chess on TV, and so on. But such flights of fancy are no more than ''Just-So'' stories and do not constitute a causal mechanism that explains why chess playing or numerical cognition occurs in one culture and not another. They speak only of vague notions like ''need'' and ''desire.'' On the other hand, chess playing cannot occur without a chess set of some kind, and counting cannot occur without words or other symbols for tagging numerosities.

One of the problems with comparing the linguistic and the cultural accounts of Pirahã innumeracy is that we are dealing with causal factors that operate at different grain levels. The cultural explanation is vague and amorphous. To the extent that we can define this position in terms of concepts such as ''technological advancement'' or ''trade,'' it fails to capture the causal structure of numerical competence either way. The linguistic explanation has the potential to provide a cognitive–mechanistic explanation for the relation between knowing number words and knowing about exact numerical quantities. What we need to ask ourselves is whether possession of number words is the key to understanding number: Is it the equivalent of having a chess set for playing chess? There are several intriguing experiments that suggest that language is fundamental to learning about exact number. For example, David Barner and colleagues have found that acquiring plural markers in language appears to be the key to learning about relative cardinalities. English-speaking children at around 22 months who have no plural morphology in their language fail to show evidence of numerically based preferences in a numerical choice task, whereas plural-knowing children pass numerical choice tasks (Barner, Thalwitz, Wood, & Carey, 2007). Chinese children who have no plural markers in their language appear to be delayed in their acquisition of a word for ''one,'' although they are not delayed in numerical choice tasks compared to English-speaking children (Barner, Chow, & Yang, 2009).

It seems untendentious that any language having a fully elaborated counting system that is capable of encoding large cardinalities provides sufficient conditions for its speakers to develop concepts of exact number, provided that these number words actually have exact numerical values as their conventionalized understood meaning. Languages with small number counting systems, or no numbers at all, such as those found in the Pirahã and Munduruků, characteristically do not encode exact quantities and so do not provide sufficient conditions for exact numerical knowledge in their speakers. So, the question is whether having number words is *necessary* for learning exact number concepts. Can people learn to count without words? The question here revolves around what we mean by a word. Certainly, we do not mean a spoken word, because we are well aware that people who do not speak develop number systems in other modalities such as signing in deaf populations. However, it is not the case that all deaf signers learn conventional sign systems. Some are not exposed to a standardized sign language and instead develop home sign systems that lack many of the structural properties of conventional sign systems. Such home sign systems also tend not to include conventionalized number systems.

Maria Coppola, Molly Flaherty, Liesje Spaepen, and Annie Senghas have studied number systems in Nicaraguan deaf populations. The situation in Nicaragua is very interesting because the National Sign Language of Nicaragua (NSL) was developed only within the past 30 years. Because there were no facilities for deaf people to converge in the pre-Sandinista period, no conventional sign language emerged and deaf people were left to their own devices to develop invented home sign systems. Such systems tend to lack the formal properties of fully developed languages of the kind found in standardized sign languages such as American Sign Language (ASL) and, now, NSL. Among the Nicaraguan population there are some who acquired the full system of NSL by learning to sign at an early age from other signers (so-called signers of the "second generation" and beyond). Others, who did not attend the deaf school in Managua, communicate only with their invented home sign systems.

When it comes to encoding number in a conventionalized sign language such as ASL and NSL, we see that the small numbers of 1–5 tend to look like finger counting. As numbers increase, instead of continuing with the iconic representation of finger counting, the language develops conventionalized abbreviations such as a flick of the wrist to indicate a landmark or number base. So, 12 might be indicated with two fingers and a wrist swivel. Nicaraguan home signers, on the other hand, develop no such conventionalized representations. Instead, their number words remain iconic forms of finger counting. Users of the iconic finger-counting system tend to have problems in forming exact representations of number in a way that NSL signers do not. Home signers tend to make errors and cannot hold representations of exact number in memory with accuracy. In one case, a deaf man who used a home sign led a full life with a family and thoroughly enjoyed playing craps. However, his memory for numbers was quite limited when he was required to reconstruct the number of taps given on his leg. The comparative results from NSL and home signers suggest that finger counting is no substitute for real linguistic counting. But what is the major difference between finger counting and linguistic counting? Basically, a language-based counting system has the following features:

1. It has symbols for a limited set of numerical values.
2. It has a method of encoding larger numerical values through some representations of a base and a combinatorial system for incrementation over the base with the unit symbols.
3. The method for encoding the base is arbitrary in some sense.

One property that seems crucial here is the arbitrariness of the symbolic system that is not in the iconic finger-counting system. Because finger counting is iconic, it relies on the counter to keep track of a representation that is no more compact than the things being counted. As a consequence, there is the possibility that errors will emerge as quantities become larger. In addition, the values encoded with fingers do not have their own distinct names, and so they are not really representational in the kind of arbitrary symbolic way that numbers encoded by words are. Iconic systems such as finger counting do not develop lexical representations and do not accrue associations that would form through encounters with words in a semantically connected lexicon. In addition, a property that seems to be crucial to the task of acquiring an arithmetic relationship between numerical values is the representation of a number line on which the symbolic names of numerical values (i.e., number words) are lined up (see Dehaene, 1997). If the representation is purely an iconic finger line, then the representation is not lexical in this sense and probably never develops a higher-order cognitive representational format of this kind.

On a final note, I would like to speculate as to why some hunter–gatherers are counters and some are not. It is possible that counting can emerge in a culture, not because of technology or trading or some other such ''need,'' but from fortuitous accidents of linguistic exploration. Someone in the culture develops a game of chunking of quantities together and naming those chunks with a new word. The names of the chunks then enter into syntactic relationships that enable recursion, as in the case of the classifier systems noted earlier (Bender & Beller, 2007b). At this point, the language has the potential to develop a fully elaborated recursive counting system. There is an interesting case that was informally observed by David Wilkins (personal communication, 2004), who worked with the Arrernte tribe in Australia. The Arrernte language also had a small number 1-2-many type of counting system. However, the Arrernte had also developed a way of encoding large numbers of objects up to 100 or more by chunking groups of objects into sets of four, and counting by fours. This is very much like the grouping of coconuts found in the Pacific Islanders (Bender & Beller, 2007b). We can imagine, in this case, the possible emergence of a recursive count system. If the four-chunks were given a name, a chunk could then itself be counted and a fully recursive counting system might emerge. Because most Arrernte speakers are now bilingual in English, the development of such a system will probably never happen. However, if such a scenario is something that does happen in cultural evolution, we can see that numerical systems would emerge in this case through cultural invention. Such an invention might be encouraged by technological advancements, trading, and other circumstances, but the focal role of language cannot be ignored.

SUMMARY AND CONCLUSIONS

I have argued in this chapter that studies of numerical cognition in the Pirahã really do constitute a case of incommensurability of cognitive systems across cultures, and that the differences in numerical cognition can be rather directly related to linguistic differences. Claims that cultural factors are more parsimonious in explaining the original cause of innumeracy in the Pirahã lack both rational and empirical support. I argue that language is the tool that enables numerical cognition, in the same way that a chess set enables chess playing. It provides, in Gibsonian terms (Gibson, 1979), the *affordance* for counting. I speculate that words are special in this sense because they are symbolic and arbitrary in nature. As such, they enable the cognitive system to set up lexical representations that can enter into systematic relationships with other members of the number vocabulary such as the time line and arithmetic relations.

Notes

1. It might be thought that AA batteries could be unfamiliar to the Pirahã, and therefore culturally inappropriate. Although batteries are not indigenous to the culture, the Pirahã were very familiar with these as objects that the Everetts used extensively for flashlights and recording equipment. In addition, the Pirahã often borrowed flashlights for alligator hunting and were familiar with changing batteries when the power was depleted.
2. The Mundurukù appear to be significantly more integrated with mainstream Brazilian culture than are the Pirahã. For example, a large

proportion of the Mundurukù are bilingual in Portuguese and attend schools, whereas the Pirahã are exclusively monolingual and have no schooling. Ladefoged (2001) commented with some astonishment that they were the most linguistically isolated of any of the hundreds of cultures that he had studied. In addition, Pica et al. (2004) note that the Mundurukù are sometimes employed by the FUNAI, Brazilian Indian Agency. The Pirahã have never held such positions due to their isolated culture and monolingual status that makes communication with FUNAI agents impossible.

References

Baillargeon, R. (1994a). How do infants learn about the physical world? *Current Directions in Psychological Science, 3,* 133–140.

Baillargeon, R. (1994b). Physical reasoning in young infants: Seeking explanations for unexpected events. *British Journal of Developmental Psychology, 12,* 9–33.

Barner, D., Chow, K., & Yang, S. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology, 58,* 195–219.

Barner, D., Thalwitz, D., Wood, J., & Carey, S. (2007). On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science, 10,* 365–373.

Beller, S., & Bender, A. (2008). The limits of counting: Numerical cognition between evolution and culture. *Science, 319,* 213–215.

Bender, A., & Beller, S. (2007a). Counting in Tongan: The traditional number systems and their cognitive implications. *Journal of Cognition and Culture, 7,* 213–239.

Bender, A., & Beller, S. (2007b). Numeral classifiers in specific counting systems: Cultural context, linguistic principles, and cognitive implications. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual conference of the Cognitive Science Society* (pp. 821–826). New York: Lawrence Erlbaum.

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

Bornstein, M. H. (1985). Human infant color vision and color perception. *Infant Behavior & Development, 8*(1), 109–113.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal Psychology, 55,* 1–5.

Brown, R. W. (1970). *Psycholinguistics*. New York: Free Press.

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49,* 454–462.

Casasanto, D. (2005). Crying "Whorf." *Science, 307,* 1721–1722.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge,: MIT Press.

Davidoff, J., Davies, I., & Roberson, D. (1999). Color categories in a stone-age tribe. *Nature, 398,* 203–204.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology, 46,* 621–646.

Everett, D. L. (2009). Pirahã Culture and Grammar: A Response to Some Criticisms. *Language, 85* (2), 405–441.

Fodor, J. A. (1981). *Representations: Philosophical essays on the foundations of cognitive science* (first MIT Press ed.). Cambridge, MA: MIT Press.

Fodor, J. A., Garrett, M. F., Walker, E. C., & Parkes, C. H. (1980). Against definitions. *Cognition, 8,* 263–267.

Frank, M., Gibson, E., Fedorenko, E., & Everett, D. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition, 108,* 819–824.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

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

Gordon, P. (2004b). Supplementary online materials to ''Numerical cognition without words: Evidence from Amazonia.'' [Electronic Version]. *Science Online*, from http://www.sciencemag.org/cgi/content/full/1094492/DC1.

Gordon, P. (2005). Response to letter concerning ''Numerical Cognition without Words.'' *Science, 307,* 1722.

Hammarström, H. (2008). *Small Numeral Systems and the Hunter-Gatherer Connection*. Paper presented at Language, Communication and Cognition, Brighton, UK.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93,* 10–20.

James, W. (1890). *The principles of psychology*. New York: H. Holt and Company.

Kant, I. (1848). *Critique of pure reason*. English translation by Francis Haywood. London: William Pickering.

Ladefoged, P. (1982). *A course in phonetics*, 2nd ed. New York: Harcourt, Brace Jovanovich.

Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages*. Malden, MA: Blackwell.

Lucy, J. A. (1992). *Language, diversity and thought. A reformulation of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. A., & Shweder, R. A. (1979). Whorf and his critics: Linguistic and non-linguistic influences on color memory. *American Anthropologist, 81*, 581–615.

Martin, L. (1986). Eskimo words for snow: A case study in the genesis and decay of an anthropological example. *American Anthropologist, 88*, 418–423.

Menninger, K. (1969). Number words and number symbols. A cultural history of numbers. Cambridge, MA: M.I.T. Press.

Nevins, A., Pesetsky, D., & Rodrigues, C. (2009). Piraha exceptionality: A reassessment. *Language, 85*, 355–404.

Piaget, J. (1952). *The child's conception of number*. London: Routledge and Kegan Paul.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian Indegene group. *Science, 306*, 499–503.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking.

Pullum, G. K. (1989). Topic ... Comment: "The great Eskimo vocabulary hoax." *Natural Language and Linguistic Theory, 7*, 275.

Pullum, G. K. (1991). *The great Eskimo vocabulary hoax, and other irreverent essays on the study of language*. Chicago: University of Chicago Press.

Quine, W. V. (1960). *Word and object*. Cambridge, MA: MIT Press.

Sapir, E. (1921). *Language, an introduction to the study of speech*. New York: Harcourt, Brace and Company.

Sapir, E. (1924). The grammarian and his language. *American Mercury, 1*, 149–155.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 38*, 179–211.

Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. *Infant Behavior and Development, 21*, 181–200.

Spelke, E. S. (2000). Core knowledge. *American Psychologist, 55*, 1233–1243.

Whalen, J., Gelman, I. I., & Gallistel, C. R. (1999). Non-verbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*, 130–137.

Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.

11 A WORLD OF RELATIONS
Relational Words

Julia Parish-Morris, Shannon M. Pruden, Weiyi Ma, Kathy Hirsh-Pasek, & Roberta Michnick Golinkoff

Children climb on jungle gyms, run around trees, and kiss their parents. The everyday world in which we live is fundamentally dynamic, with events defined in terms of the relations that objects have to actions. When we label these events and relations, we use relational words that come in all forms. They can be nouns such as *brother, island,* and *passenger* (Hall & Waxman, 1993; Keil & Batterman, 1984; Maguire, Golinkoff, & Hirsh-Pasek, 2006) or more prototypical verbs and prepositions such as *climb, run, on,* and *around,* each of which cannot be defined without specifying the referent in relationship to another object. For example, the verb *climb* refers to the relationship between an agent and something being ascended, and words such as *on* and *around* are defined by a ground object. Importantly, understanding these relational terms is not only central to a complete theory of language, but is also the key to understanding how children link words and grammar (Lidz, 2006). Until recently, we knew little about when and how children acquire *any* relational terms (see Hirsh-Pasek & Golinkoff, 2006, for a review). This chapter examines how children learn relational terms through the lens of verbs and prepositions. After reviewing evidence that verbs and prepositions are particularly difficult to learn, we examine several hypotheses to explain this disparity. We conclude that although children have the foundations necessary to learn these words, they have trouble mapping relational words to their referents.

RELATIONAL TERMS

Relational Terms Are Learned Later

Many have documented the fact that relational terms such as verbs and prepositions lag behind nouns in vocabulary acquisition (Braunwald, 1995; Choi, 1998; Choi & Bowerman, 1991; Choi & Gopnik, 1995; Clark, 1996 ; Fenson, Dale, Reznick, & Bates, 1994; Huttenlocher, Smiley, & Charney, 1983; Nelson, 1989; Smiley & Huttenlocher, 1995; Tardif, 1996; Tomasello, 1987; Tomasello & Kruger, 1992; but see Bloom, Tinker, & Margulis, 1994). These data come from two sources. First, *natural observation* studies reveal more concrete object words than verbs in children's vocabularies, a finding that generally holds across languages (Tomasello, Akhtar, Dodson, & Rekau, 1997; Gentner, 1982; Bornstein et al., 2004; but see Tardif, 1996). Second, *experimental paradigms* conducted in the laboratory also reveal that verbs are harder to learn than nouns (Imai, Haryu, & Okada, 2002; Imai, Haryu, Okada, Lianjing, & Shigematsu, 2006; Imai, Okada, & Haryu, 2005; Meyer et al., 2003; Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996).

Why Are Relational Words So Difficult to Learn?

In 1982, Gentner offered a framework to explain why verbs and other relational terms are generally harder to learn than words from other lexical classes such as nouns; Golinkoff and colleagues (1996) built on this framework. First, verbs are polysemous. They are more likely to have multiple meanings than nouns. For example, *Merriam-Webster's Dictionary* (1998) has over 72 entries for the verb *run*, but only 11 entries for the noun *ball*. Second, most objects can exist in a vacuum; that is, they can be labeled in the absence of any other referent. Actions, on the other hand, require the presence of either an agent or an object to make the action happen (e.g., something needs to traverse a path by a certain manner). Third, attaching (or mapping) a word to an object is guided by clear principles that are rarely violated. For example, the whole object assumption (Markman, 1991) holds that children as young as 12 months of age attach a novel label to entire objects rather than subparts (Hollich, Golinkoff, & Hirsh-Pasek, 2007). Children who hear labels in the presence of an agent or object engaged in an action are faced with the problem of determining whether the label maps to the *entity* or to the *action*, thereby increasing processing demands and slowing down the mapping process (Kersten, Smith, & Yoshida, 2006). Fourth, children must decide which element of the action is being labeled: Is it the *manner* (how the action is performed) of the action or the *path* (i.e., where the object is going)? Any number of event components could be packaged into a given verb, and the particular components that are packaged differ cross-linguistically. Fifth, verbs label actions, and actions are more likely to be ephemeral. Nouns, on the other hand, generally label concrete objects such as cars and cakes (Langacker, 1987; Slobin, 2001; Smith, 2000). Sixth, actions are often labeled before or after an action has taken place (Tomasello & Kruger, 1992). Finally, children learning different languages hear different proportions of nouns and verbs from their caregivers. English-speaking caregivers of 12- to 36-month-old children, for example (e.g., Blewitt, 1983; Ringler, Trause, Klaus, & Kennell, 1978; Shipley, Kuhn, & Madden, 1983), include more nouns at 18 months than at 27 months of age (Furrow, Nelson, & Benedict, 1979). This change is reflected in the proportions of nouns and verbs present in children's vocabularies from 15 to 35 months of age (Casasola et al., 2006; Sandhofer, Smith, & Luo, 2000). Cross-culturally, children in the United States tend to hear proportionately more nouns than verbs, whereas children in China hear an equal number (Tardif, 1996).

Taken together, these findings lead to three hypotheses for why it might be more difficult for children to learn verbs and prepositions than nouns. First, according to the *conceptual prerequisite hypothesis*, children might not have the conceptual foundations necessary to learn verbs and prepositions. That is, they might not be able to isolate *climbing* from the scene in which a child runs to a slide, climbs a ladder, and slides down. Second, according to the *mapping hypothesis*, children might experience difficulty connecting words to the events they isolate (Gentner, 1982; Gentner & Boroditsky, 2001). That is, they might be able to focus on the action of *climbing*, and even hear the word "climb," but may not be able to link the word to the world. Third and finally, concordant with the *differential input hypothesis*, children learning Western languages might be slower to learn relational terms such as verbs because they hear them less frequently, whereas children who are exposed to languages such as Mandarin, which have a higher proportion of verbs in the input, will learn verbs earlier (Tardif, 1996; Tardif, Gelman, & Xu, 1999; Tardif, Shatz, & Naigles, 1997). We are only beginning to investigate how the differential input hypothesis interacts with other factors to facilitate verb learning (Ma, Golinkoff, Hirsh-Pasek, McDonough & Tardif, 2009), so this chapter seeks to answer the question "Why are more-relational words such as verbs and prepositions harder to learn than less-relational words such as nouns?" by examining evidence in support of the first two hypotheses: the *conceptual prerequisites* and *mapping* hypotheses.

THE CONCEPTUAL PREREQUISITES HYPOTHESIS

Despite their relatively straightforward appearance, verbs and prepositions do not unambiguously label actions. Rather, they label a subset of what occurs when an action or spatial relation takes place—what we call "semantic components." These semantic components include *path* (the trajectory of the object or agent; e.g., come, approach, enter), *manner* (the way in which an agent moves; e.g., walk, dance, swagger, sway, stroll), *motion* (the general fact that motion is taking place), *figure* (the primary agent or object in the event; e.g., the dog/person/car), *ground* (the reference point for the event's path; e.g., a chair, the food dish, a parking lot), and *cause* (the cause of the figure's motion; e.g., being pushed; Talmy, 1985, 2000).

Relational words in languages across the world encode the semantic components of actions and events differently (Slobin, 2001; Talmy, 1985; see also Langacker, 1987; Slobin, 2001; Gentner & Boroditsky, 2001). Although the components *path* and *manner* are encoded by relational terms in many languages of the world, some languages emphasize one component over the other (Jackendoff, 1983; Langacker, 1987; Talmy, 1985). For example, English tends to package motion and manner in its verbs (as in *jump*) whereas Spanish tends to conflate motion and path in its verbs (as in *exit*), often not mentioning the *manner* by which a person leaves. These cross-linguistic differences mean that to learn their native tongue, infants must differentially attend to the ways in which semantic components are encoded in the ambient language. Consequently, infants may need considerable linguistic experience before they notice which components are encoded in the relational terms of their particular language, and are able to package these components into relational words.

Do prelinguistic infants have access to the semantic components that underlie relational language? Some speculate about the ability of infants to perceive and discriminate the components of actions and spatial relations, and suggest that children possess the conceptual prerequisites necessary for relational word learning from an early age. In fact, Gentner (1982; see also Gentner & Boroditsky, 2001) hypothesizes that such prerequisites might be in place during infancy. Gentner (1982) states that

> relations that act as predicates over objects are, I suspect, perceived quite early. Movement, change, directionality, and so on, seem quite interesting to infants. . . it is not perceiving relations but packaging and lexicalizing them that is difficult. (p. 326)

Others suggest that "the young child's conceptual repertoire may be rich and varied enough from the start" (Snedeker & Gleitman, 2004, p. 261). Additionally, some contend that children are equipped with an abundant conceptual base at a very young age (Jackendoff, 1983; Mandler, 1991, 1992, 2004). These hypotheses have only recently been put to the test.

Research suggests that infants are keenly aware of movement and use movement to individuate objects and actions (Wynn, 1996; Sharon & Wynn, 1998). Even if infants are able to perceive and discriminate the semantic components of actions and spatial relations that become the referents of relational words, this is not the end of the task for babies learning language. As Golinkoff and colleagues (2002) point out, children must not only discriminate between the semantic components of relations, but must also be able to form categories of actions to learn and extend relational terms. The analogy with object nouns is clear (Oakes & Rakison, 2003). Just as the noun "apple" refers to the apple I just ate, it also refers to an apple in the store and the apple on a sign and an apple in a storybook. Similarly, the verb "squish" can refer to squishing a bug, squishing blueberries, and squishing yourself into an overfilled Mini Cooper. Verbs, then, do not label single actions, but rather refer to categories of actions and events. Categorization is a useful heuristic device that allows language learners to approach the world in an organized fashion; that is, rather than requiring a new concept and a new label for every instance of "squish," infants who are adept categorizers can analyze an event and call up the proper category of similar events. For example, "squish" might call to mind a number of instances of overly tight relations under which the current situation may be evaluated and understood. Rather than processing the current instance of "squish" as a completely novel event, infants call upon their knowledge of categories—thereby reducing processing demands. This suggests that after children develop the ability to attend to actions and parse them into distinct action components, they must detect similarities across such actions and categorize them efficiently. Only then are they ready to attach a relational word to an action.

We will now evaluate the conceptual prerequisites hypothesis using evidence from research on how babies discriminate and categorize events and spatial relations. Four sets of relations have been studied most extensively: (1) containment, support, and degree-of-fit, (2) path and manner, (3) source and goal, and (4) figure and ground.

Containment, Support, and Degree-of-Fit

Three spatial relations, *containment*, *support*, and *degree-of-fit*, have commanded much attention in recent years (Aguiar & Baillargeon, 1999; Baillargeon, 1998, 2001, 2002, 2004; Baillargeon & Hanko-Summers, 1990; Baillargeon, Needham, & DeVos, 1992; Baillargeon & Wang, 2002; Casasola, 2005a,b; Casasola & Cohen, 2002; Casasola, Cohen, & Chiarello, 2003; Casasola & Wilbourn, 2004; Casasola, Wilbourn, & Yang, 2006; Choi & Bowerman, 1991; Choi, McDonough, Bowerman, & Mandler, 1999; Hespos & Baillargeon, 2001a,b, 2006; Hespos & Spelke, 2004; Wang, Baillargeon, & Brueckner, 2004). *Containment* is defined as "something in any fully or partially enclosed space... bounded space with an inside and an outside" (Mandler, 2004, p. 78). In English, this spatial relation is lexicalized by the word "in." Lexicalized by the English word "on," the spatial relation *support* is "when the figure is in contact with—typically supported by, attached to, or encircling—an external surface of the ground" (Choi et al., 1999, p. 247). *Degree-of-fit*, on the other hand, makes a distinction between interlocking surfaces (e.g., tight-fit) and noninterlocking surfaces (e.g., loose-fit; Bowerman & Choi, 2003).

Containment, support, and degree-of-fit are especially interesting spatial relations because they are packaged in very different ways across languages. English speakers lexicalize the distinction between containment and support relations, but do not consider degree-of-fit. For instance, English speakers use the word "on" to mark both an instance of tight-fitting support and loose-fitting support. In the Korean language, speakers lexicalize the distinction between tight-fitting and loose-fitting relations, collapsing across containment and support. That is, Korean speakers use the word "kkita" to denote both tight-fitting containment (e.g., fitting a peg *tightly* into a hole) and tight-fitting support (e.g., fitting one Lego® *tightly* onto another).

Discrimination of Containment, Support, and Degree-of-Fit

Abundant research suggests that infants have some understanding of containment and support relations at an early age. A classic series of studies by Baillargeon and colleagues was among the first to examine infants' basic knowledge about containment and support relations (Baillargeon, 1998, 2001, 2002; Baillargeon & Wang, 2002; Hespos & Baillargeon, 2001a,b; Wang, Baillargeon, & Brueckner, 2004). Recently, Hespos and Baillargeon (2001a) showed that 2.5-month-old English-learning infants have a rudimentary understanding of both containment and support relations. Furthermore, these very young infants can reliably discriminate between possible and impossible containment and support situations, despite having no prior habituation or training in the laboratory. This suggests that 2.5-month-old infants enter the laboratory with knowledge of containment and support events—relations lexicalized in most languages.

In one study, infants were shown two types of events. In the possible event condition, infants viewed a scene in which an object was lowered into an open container. In the impossible event condition, infants saw the same object lowered into a closed container. Infants looked reliably longer at the impossible event, indicating that they were surprised when their expectations about a containment event were violated (Hespos & Baillargeon, 2001b).

Recent research also documents the ability of infants to discriminate spatial concepts that are *not* typically codified in their native language (Hespos & Spelke, 2004). For example, 5-month-old infants exposed only to the English language appear to be sensitive to the degree-of-fit feature, a distinction that is not marked in their native language. During habituation, infants viewed either objects fitting tightly or loosely in a container. At test, infants were shown both the familiar relation they had seen during habituation and a novel relation. For example, infants shown tight-fitting containment during habituation viewed this same familiar relation (i.e., tight-fit) and a novel relation (i.e., loose-fit) at test. Infants showed a marked increase in looking at the novel relation during the test phase, which indicates that they discriminate between tight-fitting and loose-fitting containment (Hespos & Spelke, 2004). These findings suggest that infants may be predisposed to pay attention to the kinds of spatial relations that are relevant to learning *any* language in the world.

Categorization of Containment, Support, and Degree-of-Fit If infants have the capacity to recognize and discriminate the spatial relations containment, support, and degree-of-fit by 5 months of age (Baillargeon, 1998, 2001, 2002; Baillargeon & Wang, 2002; Hespos & Baillargeon, 2001a,b; Hespos & Spelke, 2004; Wang et al., 2004), when are they able to group these relations into language-specific categories? Recent evidence from Casasola and colleagues (2003) suggests that by 6 months of age infants can form categories that include multiple instances of containment. Using a habituation paradigm, infants were shown four different events in which a containment relation was depicted (e.g., candle *in* cookie cutter, peg *in* block, car *in* larger car, and monkey *in* basket). Two of the relations with which the infants were familiarized were *tight-fitting* and two of the relations were *loose-fitting*. Once habituated, infants were presented with four test trials that varied in the degree of familiarity that the child had with the objects and containment/support relations: (1) familiar objects–familiar relation (e.g., candle *in* cookie cutter; tight-fit), (2) amiliar objects–novel relation (e.g., peg *on* block; loose-fit), (3) novel objects–familiar relation (e. g., cup *in* bowl; loose-fit), and (4) novel objects–novel relation (e.g., turtle *on* other turtles; tight-fit). Infants showed increased attention to clips depicting the new spatial relation, *support*, regardless of whether the objects were familiar or novel and regardless of tight-fit or loose-fit. By successfully completing this task, infants not only demonstrated that they could categorize containment relations not only across varying objects but also across *degree-of-fit* (e.g., tight-fitting containment and loose-fitting containment; see McDonough, Choi, & Mandler, 2003).

Containment appears to be the easiest spatial relation for children to discriminate and categorize (Casasola & Cohen, 2002). Although it is also lexicalized in the English language, *support* appears to be much harder to categorize than *containment*. In fact, Casasola (2005a,b) found that infants did not demonstrate the ability to form categories of the support relation until 14 months of age, and even then, they were able to form the categories only if the different exemplars of *support* were perceptually similar to one another. For example, infants in one study were habituated to multiple perceptually similar exemplars of the *support* relation (e.g., a block on top of a table). After habituating to these events, infants were shown four test trials that varied in the degree of object and relation familiarity. Like the infants in the containment study (Casasola et al., 2003), infants increased their attention to test clips depicting a novel spatial relation (i.e., containment). These findings suggest that by 14 months of age infants can form categories of the spatial relation *support* in certain limited circumstances.

To explore whether adding spatial language facilitates infant categorization of *support* across perceptually dissimilar exemplars, English-learning 18-month-old children heard a spatial word (e.g., *on*) while being familiarized with members of the category *support* (Casasola & Bhagwat, 2007). Results revealed that when provided with spatial language, children were able to form a category of *support* that included perceptually dissimilar instances such as a *tight-fitting* ring on a post and a *loose-fitting* cup on a table. Casasola (2008) concluded that, "When a spatial category consists of perceptually variable exemplars, spatial language scaffolds infants' spatial categorization" (p. 24).

Categorization of *degree-of-fit* has received less attention by researchers, as it is not lexicalized in the English language. One of the few studies to address whether English-reared infants form categories of spatial relations not typically encoded in their native language was conducted by Casasola and Cohen (2002). Seventeen- and 19-month-old children viewed events in which objects depicted tight-fitting containment or tight-fitting support relations. Four different tight-fitting relations were shown to infants during habituation (e.g., candle *tightly in* cookie cutter, lego *tightly on* block, peg *tightly in* block, and round man *tightly on* car). Once habituated to these events, infants viewed four different test trials—each varying in the degree to which infants were familiar with the objects and the spatial relation depicted. Interestingly, 17- and 19-month-old children looked longer at the novel than at the familiar spatial relation, but *only when the objects were familiar*. Thus, although infants showed evidence of discriminating among the spatial relations in this study, they did not show evidence of having formed a category of *degree-of-fit* when the objects used to depict the relation varied (Casasola & Cohen, 2002).

In sum, a preponderance of studies suggests that preverbal infants can form nonlinguistic spatial categories of the semantic components *containment*, *support*, and *degree-of-fit*. Furthermore, the research by Casasola and colleagues (Casasola, 2005a,b; Casasola, 2008; Casasola & Cohen, 2002; Casasola et al., 2003) revealed that some spatial concepts are more accessible than others: *Containment* is discriminated and categorized by English-reared infants earlier than *support*. Finally, children's failure to categorize a relation not encoded in their native language suggests that our ambient language may play a role in what commonalities are attended to, a position favored by Bowerman and colleagues (Bowerman & Choi, 2003; Choi, 2006).

Path and Manner

Path and *manner* are semantic components of motion events that are perceived very early in life. *Path* is typically defined as the movement of "any object following any trajectory through space, without regard to the characteristics of the object or the details of the trajectory itself" (Mandler, 2006). *Manner* is the way in which a path is traversed or a motion is carried out (Talmy, 1985). Path and manner tend to be studied in conjunction because each object moving along a *path* must do so in some *manner*. Although *path* and *manner* are codified in the relational terms of many languages of the world (Jackendoff, 1983; Langacker, 1987; Talmy, 1985), little is known about when and how infants detect and discriminate these semantic components. Like the spatial relations *containment*, *support*, and *degree-of-fit*, *path* and *manner* are packaged differently across languages, suggesting that children may need the ability to detect these components prior to learning their native language.

Path is a semantic component that is fundamental to acquiring concepts such as animacy and causality, which are then recruited to learn relational language (Mandler, 2004). It has been argued that *path* is the most perceptually salient of all conceptual primitives (Mandler, 2004). In fact, research with both hearing and deaf populations shows that there is a path primacy in the production of relational terms (Naigles, Eisenberg, Kako, Highter, & McGraw, 1998; Zheng & Goldin-Meadow, 2002). For example, English-speaking 2-year-old children produce more language that emphasizes the *path* of an entity over the *manner* (Naigles et al., 1992). In addition, both American and Chinese deaf children produce more gestures for paths than for manners (Zheng & Goldin-Meadow, 2002).

Discrimination of Path and Manner Our laboratories explore the development of *path* and *manner* by studying the ability of prelinguistic infants to attend to changes and discriminate these semantic components in nonlinguistic, dynamic motion events (Pulverman & Golinkoff, 2004; Pulverman, Sootsman, Golinkoff, & Hirsh-Pasek, 2003; Pulverman et al., 2006). Using a habituation paradigm, English-reared 7-month-old infants viewed a video of an animated starfish moving along a single *path* in a single *manner* (e.g., the starfish *spinning over* the ball). After infants fully habituated to this event, they were shown four different test trials: (1) a control trial, in which the starfish moved along the same path by the same manner as seen in habituation (e.g., *spinning over*), (2) a manner change trial, in which the starfish moved along the same path by a new manner (e.g., *bending over*), (3) a path change trial, in which the starfish moved along a new path by the same manner (e.g., *spinning under*), and (4) a path and manner change trial, in which both a new path and new manner were shown (e.g., *flapping past*). Seven-month-old infants showed increased attention to all of the test trials except the control trial, suggesting that they noticed changes in these events.

Pulverman and colleagues (2003) also tested English-reared 14- and 17-month-old children using the same exact stimuli and methods in order to determine if the ambient language influenced infants' discrimination patterns. Like the 7-month-old infants, both 14- and 17-month-old children showed increased attention to path and manner changes in the test trials. Unlike the younger infants, however, Pulverman and colleagues found that 14- and 17-month-old English-reared children who possessed a relatively rich vocabulary were more attentive to *manner* changes than to *path* changes. Taken together, these results suggest that even preverbal infants notice those components of actions that are typically encoded by the world's languages, and that infants who attend more to those components that are relevant to their native language have larger vocabularies.

The results reviewed so far are based on infants' attention to animated events. Do these findings apply to naturalistic events as well? Casasola, Hohenstein, and Naigles (2003) showed English-reared 10-month-old infants naturalistic scenes using human agents. Consistent with the results of our studies in which infants could discriminate path and manner changes in animated scenes, Casasola and colleagues found that infants noticed and discriminated path and manner changes in naturalistic events.

To assess cross-linguistic differences in early attention to and discrimination of *path* and *manner*, an extension of our laboratory's research was conducted on site in Mexico. Pulverman, Golinkoff, Hirsh-Pasek, and Sootsman-Buresh (2008) used the same stimuli previously described to examine *path* and *manner* discrimination in 7-month-old and in 14- to 17-month-old Spanish-reared babies. Like English-reared infants, Spanish-reared infants also discriminated changes in both *path* and *manner* in nonlinguistic events. When looking behaviors were correlated with vocabulary, however, Spanish-reared infants showed a pattern of results different from their English-reared counterparts. Spanish-reared infants who had *small* vocabularies were more attentive to *manner* changes than to *path* changes, whereas Spanish-reared infants with large vocabularies did not attend more to any one specific element. These data suggest that if individual children focus on the components of events that the verbs in their language encode, they restrict the range of hypotheses they must entertain for what verbs might refer to. Thus, the English-reared children with large vocabularies who attend to *manner* changes are likely to learn more *manner* verbs. However, if individual Spanish-reared children attend to *manner*, they are attracted (for whatever reason) to the wrong component of events, thereby disadvantaging their learning of vocabulary (Pulverman et al., 2008).

A recent cross-linguistic study examined the effect of language-specific packaging on attention to *path* and *manner* in Greek-speaking and English-speaking adults (Papafragou, Hulbert, & Trueswell, 2008). This study is similar to the Spanish–English comparative study conducted by Pulverman and colleagues in that Greek is a *path*-oriented language and English is a *manner*-oriented language. Participants were shown an event and were asked to either memorize the event or prepare a verbal description of the event. An analysis of eye-gaze direction revealed that when participants freely examined the event, speakers of both languages attended equally to *path* and *manner*. When participants were asked to prepare verbal descriptions of the event they saw, however, they immediately focused on the particular event components that are emphasized in their native language (i.e., Greek speakers attended more to parts of the event that gave them *path* information, whereas English speakers attended more to parts of the event that provided information about *manner*). Similarly, the eye gaze of participants who were instructed to memorize the event reverted to language-specific patterns after the event ended and they attempted to commit the event to memory. Thus, event perception per se is not affected by differences in the speaker's native language, but when linguistic forms are recruited to the task (either via purposeful commitment of the event to memory or by preparation of a verbal description of the event), language-specific biases become apparent (Papafragou et al., 2008).

Categorization of Path and Manner A significant body of knowledge exists about when and how *path* and *manner* are discriminated by people both within and across languages. Much less is known about how infants abstract and categorize the semantic components *path* and *manner*.

Our laboratory's recent explorations into how children categorize *path* and *manner* used the same animated stimuli as Pulverman and colleagues (Pulverman & Golinkoff, 2004; Pulverman et al., 2003). In two studies we addressed whether infants could abstract an invariant path across multiple exemplars of manner (Study 1) and whether they could find the invariant manner across varying exemplars of path (Study 2; Pruden, Hirsh-Pasek, Maguire, & Meyer, 2004). Three age groups were tested: 7- to 9-month-old, 10- to 12-month-old, and 13- to 15-month-old infants. The youngest age group tested (7- to 9-month-old infants) was known to discriminate changes in both path and manner in dynamic events (Pulverman et al., 2003).

In Study 1, we addressed the question of whether infants could abstract an invariant path across varying manners (that is, form a category of *path;* Pruden et al., 2004). During the familiarization phase, infants viewed an animated starfish performing exactly the same path across four different manners (e.g., "flap *over,*" "side bend *over,*" "toe touch *over,*" and "spin *over*"). At test, two events were shown simultaneously on a split screen: (1) an in-category test event (i.e., same path and novel manner; e.g., "twist *over*") and (2) an out-of-category test event (i.e., novel path and novel manner; e.g., "twist *under*"). Infants' looking times to these two events were analyzed to determine if they had a preference for either event. Results revealed that by 10 to 12 months of age, infants show a significant preference for the in-category test event. This finding suggests that 10- to 12-month-old infants can abstract an invariant path across varying manners, which is a sign of the ability to categorize and ultimately conceptualize the referents of relational words.

In Study 2, we assessed whether infants could abstract an invariant manner across multiple exemplars of path (that is, form a category of *manner;* Pruden et al., 2004). During the familiarization phase of this study, infants were shown four exemplars in which the animated starfish moved in exactly the same manner, but varied its path (e.g., "*spin* around," "*spin* past," "*spin* in front," and "*spin* under"). During the test phase, infants viewed two events simultaneously: (1) an in-category test event (i.e., same manner and novel path; e.g., "*spin* over") and (2) an out-of-category test event (i.e., novel manner and novel path; e.g., "*side bend* over"). In contrast to Study 1, in which 10- to 12-month-old infants showed evidence of categorizing a *path* over various *manners*, only the oldest age group, the 13- to 15-month-old infants, showed evidence of the ability to abstract an invariant *manner* over multiple *paths*. Importantly, infants in both studies noticed path and manner per se and not just novelty or change, because both test scenes contained some novelty or change.

What happens to categorization when language is added to the familiarization that infants receive in these *path* and *manner* tasks? Results from studies of *containment* and *support* relations indicate that the addition of spatial language facilitates categorization (Casasola, 2005a; Casasola & Bhagwat, 2007). Is it possible that infant looking patterns will be affected by the addition of language to studies of *path* and *manner*? To answer this question, Pruden (2007) conducted a series of studies with 7- to 9-month-old infants. Stimuli were exactly the same, with the exception that the events in the familiarization phase were accompanied by a single word (e.g., *javing*). Results revealed that whereas infants younger than 10 months of age were not able to form a category of *path* in nonlinguistic studies, the addition of a word to the familiarization phase facilitated categorization of *path* such that 7- to 9-month-old infants were now able to abstract the invariant event component across multiple scenes. In contrast, although infants in the *manner* study heightened their attention to the familiarization phase when it was accompanied by a label, they did not show evidence of categorizing *manner* across multiple *paths*. This result is consistent with the robust finding that infants are able to discriminate and categorize *paths* earlier than they are able to discriminate and categorize *manners* (Pruden et al., 2004; Pulverman & Golinkoff, 2004; Pulverman et al., 2003, 2006).

Source and Goal Sources and goals represent a relatively new area of research concerning the prelinguistic foundations of relational language. Researchers recently found that *path* can be split into two subtypes: goal-oriented paths, in which an object moves toward or culminates in a goal or end point, and source-oriented paths, in which an object moves away from an origin or starting point (Lakusta & Landau, 2005). Exploration of children's prelinguistic understanding of source and goal has been both inspired and informed by multiple studies showing that a goal-bias in natural language is found across diverse ages and populations (Bowerman, 1996; Ihara & Fujita, 2000; Lakusta & Landau, 2005; Nam, 2004; Regier, 1996; Regier & Zheng, 2003; Zheng & Goldin-Meadow, 2002).

Discrimination of Source and Goal

Three studies conducted by Lakusta and colleagues (2007) revealed that 12-month-old infants are able to separate goals and sources when viewing events. In a goal-discrimination task, children were familiarized with a scene in which a duck waggled in place, accompanied by attention-getting audio, and then moved along a path to one of two possible goal objects (e.g., a green bowl, a red block). When the duck reached a goal object, it again stopped and waggled, accompanied by audio. At test, children saw the duck either move to the same goal in a different location, the same goal in the same location, a different goal in the same location, or a different goal in a different location. Variations in looking time indicated whether infants encoded the goal of the event. Increased looking to the "different-goal object" testing conditions but not the "same-goal object" conditions revealed that children preferentially encoded the goal of an action over its path or location and were surprised when the figure moved to a new goal but not when the figure moved to the same goal in a new location (Lakusta, Wagner, O'Hearn, & Landau, 2007; for a thorough review of the paradigm, see Woodward, 1998).

In a source-discrimination task also conducted by Lakusta and colleagues (2007), children viewed the same scenes, except that the *source object* of the action was manipulated. The same stimuli were used. For example, a duck waggled at one of two source objects (a bowl or a box) and then moved along a path and waggled at the end. During the test phase, infants saw the duck move from a different source object. An analysis of looking times revealed that 12-month-old infants did *not* reliably encode the source object of the action. In other words, they did not increase their looking time when the source of the duck's movement switched from the bowl to the box. In a follow-up study to determine whether 12-month-old infants could be *forced* to attend to source objects, infants were shown exceptionally salient source objects (e.g., a metallic blue block covered with sparkly pipe cleaners and a big orange bowl covered in puffs, bows, and sequins), and then viewed the same action sequence. Infants showed evidence of encoding "super-source" objects (Lakusta et al., 2007).

Do prelinguistic infants preferentially encode the goal of an action rather than the source, consistent with the bias toward encoding goals that is found in natural language (Bowerman, 1996; Ihara & Fujita, 2000; Lakusta & Landau, 2005; Nam, 2004; Regier, 1996; Regier & Zheng, 2003; Zheng & Goldin-Meadow, 2002)? To test this possibility, Lakusta and colleagues pitted these "supersources" against ordinary goals from the very first experiment and analyzed looking times to "different source/same goal" versus "same source/different goal" scenarios. Results revealed that 12-month-old infants preferentially encode goals over sources, even when the sources are very perceptually salient.

There is substantial evidence to suggest that the primacy of goal over source is directly related to infants' understanding of intentions or purposeful, rational behavior (Csibra, Biro, Koos, & Gergely, 2003; Csibra, Gergely, Biro, Koos, & Brockbanck, 1999; Gergely, Nadasdy, Csibra, & Biro, 1995). Infants as young as 12 months old successfully predict the rational ending point of an object's motion trajectory based on whether the action is intentional and "goal-directed," and likewise expect an agent to take the most efficient route to achieve a goal (Csibra et al., 2003; Wagner & Carey, 2005). Thus, at a very young age, children appear to have rational ideas about goal orientation, are able to distinguish between sources and goals, and are biased to encode goal objects over source objects (consistent with natural language usage).

Figure and Ground

A relatively unexplored pair of conceptual primitives that help form the foundation for relational language is *figure* and *ground* (Bornstein, Arterberry, & Mash, 2007). *Figure* is defined as the moving or conceptually movable entity, whose path, orientation, or site is variable (Talmy, 1985). Infant understanding of *figure* has been studied concurrently with *ground*, which is the reference entity providing a stationary setting with respect to the figure's path, site, or orientation (Talmy, 1985). Like *path/manner*, *containment/support*, and *tight-fit/loose-fit*, *figure* and *ground* are packaged differently across languages. In English, for example, satellites such as *over*, *into*, and *across* specify both a path that the *figure* follows and the spatial properties of the *ground* object. "Into" refers not only to a path along which the figure moves, but also mandates that the ground object be some kind of enclosure (Talmy, 2000). By the same token, "across" implies a relatively stable surface that can be traversed whereas "over" requires an obstacle to be scaled.

On the other hand, Japanese is a verb-framed language that conflates motion with *path* in the main verb and expresses *manner* in a subordinated verb. In Japanese, motion path verbs are classified as having one of two different semantic factors: directional path or ground path. *Directional Path* (DP) verbs define the direction of motion relative to a starting point or ending point. DP verbs do not restrict the *ground* on which motion occurs. The *figure* can be animate (self-moving and sentient) or inanimate, so there is no figure constraint with DP verbs (Muehleisen & Imai, 1997).

The other type of Japanese path verb is *Ground Path* (GP), in which the *ground* is incorporated into the verb's meaning. GP verbs are different from DP verbs in the sense that they incorporate information about the nature or shape of the ground along with the direction of motion. With GP verbs, a space is progressively and completely covered by the motion (Muehleisen & Imai, 1997). Thus, *figure* and *ground* present an interesting cross-linguistic example of different ways that verbs can be mapped onto events. Whereas English encodes ground in few of its path verbs and does not specify figure (that is, a vehicle, animal, or a person can "climb over" a mountain), in Japanese a subset of path + ground verbs (GP verbs) encodes both the nature of the ground and the figure that may perform the action as in *Jun wa yanna o nobotta* "Jun went up the mountain." In this case the verb *noboru* implies that Jun (a person) reached the top of the mountain by climbing up the slope from the bottom. However, we cannot use *noboru* for going up to the mountain by taking a cable car or helicopter. The existence of these different kinds of encodings allows us to compare how children package these primitives into the words of different languages.

Discrimination of Figure and Ground Given that languages differ in the way that they package *figure* and *ground* information into relational words, it is important to understand how infants at different ages detect changes of figure and ground, and when they distinguish them as separate components. Researchers in our laboratory have just begun to explore this issue.

Using the Preferential Looking Paradigm, 7- to 9-month-old, 10- to 12-month-old, and 13- to 15-month-old children were tested in either a figure discrimination or a ground discrimination task (Göksun, Hirsh-Pasek, Roseberry, & Golinkoff, 2008). Stimuli included four different figures (a woman, a man, a 6-year-old girl, and a 6-year-old boy) and six different grounds (railroad, road, narrow street, bridge, grass, and tennis court). Four of the grounds (railroad, road, narrow street, and bridge) were encoded by the same verb *wataru* (go across between points) in Japanese, and are characterized by clear starting and goal points. On the other hand, grass is not encoded by *wataru*, because it is an open space that does not have clear boundaries. Similarly, tennis court is not encoded by *wataru* in Japanese, because it does not have clear starting and goal points. Infants saw grounds from the same category (e.g., railroad and road) and from different categories (e.g., railroad and tennis court or road and grass).

Infants were familiarized with a scene in which one figure (e.g., woman) traversed one ground (e.g., railroad). In the *figure discrimination* task, they compared the old event (e.g., woman crossing railroad) with a new event in which only the figure differed from the event they saw during familiarization (e.g., *man* crossing railroad). In the *ground discrimination* task, infants saw the same old event (woman crossing railroad) paired with a new event that changed only the ground (e.g., woman crossing a *street*). The dependent variable was looking time; increased looking time to the novel event would suggest that children are able to discriminate between figures and grounds.

Results revealed that 7- to 9-month-old infants had no preference for novel events in either the figure or the ground task. Ten- to 12-month-old infants, on the other hand, were able to discriminate figures but not grounds. By 13 to 15 months of age, children were able to discriminate both figures and grounds (Göksun, Hirsh-Pasek, Roseberry, & Golinkoff, 2008), and, interestingly, discriminated grounds in a way that is consistent with the categorical distinction made in Japanese. Thus, it appears that figures may be more perceptually accessible or primitive than grounds for young infants, and that children learning both English and Japanese may initially discriminate grounds in the same way—and only later are shaped by the ambient language. This pattern of progressive entry mirrors findings from other research on the conceptual foundations of relational language, including studies of *containment/support*, *path/manner*, and *source/goal*.

Categorization of Figure and Ground Our laboratory has plans to conduct a series of experiments to determine how children categorize *figure* and *ground* in English and Japanese. Given the developmental decalâge found in English-learning babies' discrimination of *figure/ground,* we hypothesize that children will also categorize *figures* before *grounds* (T. Göksun, personal communication). Consistent with the finding that English-learning children discriminate grounds in accordance with Japanese categorical distinctions, preliminary data from our categorization studies suggest that children form the kinds of ground categories that are consistent with Japanese encoding (for example, grounds that have clear starting/end points fall into one category and grounds without clear starting/end points fall into another; T. Göksun, personal communication). It is as yet unknown how Japanese-reared children will categorize figures and grounds, and whether and when one component is shaped by the ambient language to become more primary.

The research presented here and in previous sections suggests that infants have the ability to perceive, process, and abstract semantic components that are encoded in languages across the world. Thus, infants bring to the language-learning task a set of concepts used to make sense of the world of events and spatial relations, as previously argued (Gentner & Boroditsky, 2001; Mandler, 2004). Given that infants appear to be equipped with the conceptual prerequisites needed to learn various relational terms at a young age, we turn to an evaluation of the second hypothesis, mapping, to explain why children have a harder time learning relational words such as verbs and prepositions.

THE MAPPING HYPOTHESIS

Evidence suggests that even prelinguistic children are equipped with the conceptual foundations necessary to learn relational words from a very early age. If conceptual readiness is not a problem for relational word learning, then perhaps relational words are harder to learn than less-relational words because children have difficulty packaging the semantic components of events so that relational words can be mapped. According to the "mapping hypothesis," children possess the relational concepts necessary for relational word learning but get "stuck" at the point of discerning how the language around them is linked to those concepts (Gentner, 1982; Gleitman, 1990; Maguire, Hirsh-Pasek, & Golinkoff, 2006). This hypothesis states that relational terms appear late in lexical time because it is difficult for children to discern how the language they are learning expresses the events observed in the world *in words*. Thus, the difficulty with relational language acquisition is not the result of an inability to form underlying concepts, but rather is specific to mapping words to these relations. Gleitman and colleagues write (2005):

> Why are words such as *know* harder for learners to acquire than words such as *dog* or *jump?* We suggest that the chief limiting factor in acquiring the vocabulary of natural languages consists not in overcoming conceptual difficulties with abstract word meanings but rather in mapping these meanings onto their corresponding lexical forms. (p. 23)

Even though infants form action categories very early in life, they fail to map a word to those same actions nearly 2 years later. Salkind, Golinkoff, Sootsman, and Hirsh-Pasek (2002) habituated 9- to 11-month-old infants to video clips of two different females performing the same jumping jack action. At test, infants saw three clips: (1) a *control* clip of the action that was identical to the habituation trials, (2) a *novel* actor performing the *familiar* action, and (3) the same *novel* actor performing a *novel* action. Results revealed that as early as 10 months of age, children with large receptive vocabularies relative to their peers could form abstract action categories that were independent of changes in actor.

If 9- to 11-month-old children can develop categories of action, then it is expected that children more than twice as old (with significantly larger vocabularies) should be able to learn a label for that same action. Using the Intermodal Preferential Looking Paradigm, Maguire et al. (2002) presented 18-, 24-, and 30-month-old toddlers with video clips of characters doing jumping jacks, while they heard a verbal description of the action *("Hey, she's blicking!")*. At test, the toddlers saw a new actor performing the familiar jumping jacks action on one side of the screen and a novel actor performing a new action on the other side of the screen. Results revealed that children *of all ages* watched the old and new actions to the same degree at test. In other words, they *failed* to map the verb onto the correct action. In light of the earlier studies revealing that children are capable of forming nonlinguistic categories of actions by 9 to 11 months old, this startling finding suggests that children still have difficulty mapping a verb to an *already formed* category nearly 2 years later (Salkind et al., 2002).

Is the mapping problem specific to *English?* It has been suggested that children learning Korean and Chinese, for example, have less of a problem with relational mapping than do English-reared children (Choi, 1998; Choi & Bowerman, 1991; Choi & Gopnik, 1995; Tardif, 1996). Research in English, Japanese, and Chinese, however, suggests that the mapping problem exists across languages (Imai, Haryu, & Okada, 2002; Imai, Haryu, Okada, Lianjing, & Shigematsu, 2006; Meyer et al., 2003). In a recent series of studies, English-speaking, Japanese-speaking, and Chinese-speaking 3- and 5-year-old children were shown video clips of a person engaged in a novel action with a novel object (Imai, Haryu, & Okada, 2002; Imai et al., 2006). Children were exposed to one of three conditions. In the "noun" condition, they were asked to "*Look at the blick*," encouraging a noun interpretation (that is, suggesting that the referent for the word "blick" is the object in the scene rather than the action being performed with it). In the "bare-frame verb" condition, children heard a novel verb in a bare syntactic frame ("*Look, blicking! Watch blicking!*"). In the "rich-syntax verb" condition, children were given additional syntactic information ("*Look, she's blicking it!*"). During test trials, children simultaneously saw the old object engaged in a new action on one side of the screen and the old action being performed with a new object on the other side. Children again heard either the noun ("*Where's the blick?*"), the bare-frame verb ("*Where's blicking?*"), or the rich-syntax audio ("*Where's she blicking it?*").

Across the three languages, 3- and 5-year-old children had no difficulty mapping a noun to an object in the noun condition. However, in *all three languages*, 3-year-old children were unable to map the verb to the action, and performed at chance levels. Not until 5 years of age was any consistency found in verb mapping. Thus, even 3-year-old children—who are language experts compared to 10-month-old infants—have difficulty mapping a word to an action, regardless of native language. These results suggest that although mapping to *verbs* is a problem, mapping itself is not; children readily mapped a novel noun to a novel object (Imai et al., 2002, 2005; Meyer et al., 2003).

A mapping problem specific to relational concepts has been found not only in children, but in adults as well (Gillette et al., 1999; Snedeker & Gleitman, 2004). In a simulated word learning paradigm, Gleitman and colleagues showed adults silent video clips of conversations between mothers and children. A tone or nonsense word was inserted exactly where a target word had been used. Adults were then asked to guess the target word. Results revealed that adults, who presumably have *no conceptual problems* understanding the objects and events represented on the tapes, still had more difficulty correctly guessing verbs than nouns. Interestingly, they performed better when asked to guess *concrete* verbs that described visible actions (e.g., *push* and *throw*) than abstract mental-content verbs (e.g., *think* and *love*). This suggests that mapping actions, especially mental actions, to words is a considerably more difficult task than mapping objects or concrete verbs to words (Gillette et al., 1999; Snedeker & Gleitman, 2004).

Why Is Mapping Relations So Hard?

There appears to be something more difficult about mapping to relational referents such as "run" than mapping to less-relational referents such as "cup." Regardless of age, conceptual ability, experience, or the specific language being learned, mapping words onto relations is hard even when the underlying nonlinguistic category is formed easily. However, research suggests that the mapping problem hinges on the *relationality* of the referent, not on form class membership (e.g., whether a word can be categorized as a verb, preposition, noun, etc.). Contrary to past theories, all nouns are not easier to learn than all verbs (Hall & Waxman, 1993; Keil & Batterman, 1984). Although children may begin to use the noun *brother* as a label for a specific person when they are young, they eventually grow to understand that *brother* actually denotes a *relation* between a male and his sibling. Children do not begin to use nouns such as *brother, island,* and *passenger* in flexible and extendable ways until relatively late in childhood (Hall & Waxman, 1993; Keil & Batterman, 1984).

Perceptual Salience Helps Children Map to Relational Referents If words from any form class are susceptible to the mapping problem, what characteristics make certain words easier to learn than others? Various studies of early vocabulary composition suggest that the actions labeled by early verbs and the entities labeled by early nouns share one important commonality: They are more perceptually grounded than other words (Ma, Golinkoff, Hirsh-Pasek, McDonough, & Tardif, 2009). Thus, whereas the noun *dog* may be one of a child's first flexibly extended words, the noun *truth* will probably not be. Similarly, the verb *kiss* might well be in the early vocabulary of a toddler (especially one who is frequently kissed), whereas the verb *think* will not be. *Dog* and *run* have perceptual instantiations; *truth* and *think* do not.

A number of theorists have suggested that perceptual factors affecting word learning are best described as variations in the *shape, individuability, concreteness,* and/or *imageability* of a concept (Bird, Howard, & Franklin, 2003; Gentner & Boroditsky, 2001; Gillette et al., 1999; Gilhooly & Logie, 1980; Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996; Landau, Smith, & Jones, 1998; Maguire, Golinkoff, & Hirsh-Pasek, 2006; Masterson & Druks, 1998; Snedeker & Gleitman, 2004). These theorists suggest that referents that are more perceptually salient, visible, countable, imageable, individuable, concrete, and/or uniquely shaped are easier to map to words than those that are not. Specifically, Gentner and Boroditsky (2001) suggest that the *key* mapping-relevant difference between verbs and nouns is that noun referents tend to be more individuated and less relational than verb referents. The referent of a noun is often something visible and countable, like a *dog*. The referent of a verb, on the other hand, is more likely to be intangible, less perceptual, and based on relations between objects, such as *tickle* or *run*. Therefore, the difficulty experienced in verb learning is not due to membership in the syntactic category "verb," but rather is due to difficulty abstracting and mapping complex, relational, semantic information (Gentner & Boroditsky, 2001; Smiley & Huttenlocher, 1995). A logical extension of this argument is that a highly individuated and less relational verb such as *jump* is more likely to be mapped correctly and extendably than a relational noun such as *island.*

Golinkoff, Hirsh-Pasek, Mervis, Frawley, and Parillo (1995) predicted that the earliest verbs in children's vocabularies would be those with a salient "shape" (like *dancing*) rather than those with a less visible "shape" (like *thinking*). Golinkoff, Jacquet, Hirsh-Pasek, and Nandakumar (1996) found that 37-month-old children exposed to static pictures with various Sesame Street characters performing actions extend a new verb to the characters whose bodies have the same shape (e.g., arms and one leg extended for an arabesque) but not to others with a different shape. Golinkoff and colleagues (2002) also found that 3-year-old children are sensitive to the body shape of actions displayed in "point lights," or small lights attached to the head and main joints of a person's body and filmed in the dark (Johansson, 1973). Children were shown four pairs of eight known actions (picking a flower, dancing, etc.) in point light displays on a split screen in the Intermodal Preferential Looking Paradigm (Golinkoff et al., 2002; Hirsh-Pasek & Golinkoff, 1996). The pair of videos was accompanied by a verb that labeled one of the actions. Despite the fact that children only saw lights moving about on a screen, they were able to find the match when asked to look at a particular action.

Another aspect of perceptual salience that affects which words are acquired by very young children is *imageability* (Bird et al., 2003; Gilhooly & Logie, 1980; Masterson & Druks, 1998). "Imageability" refers to the ability of a word to arouse a mental image (Paivio et al., 1968). Retrospective age of acquisition (AoA) studies have found that children's early words are more perceptually salient (or imageable) than words learned later (Bird et al., 2003; Gilhooly & Logie, 1980; Masterson & Druks, 1998).

McDonough, Song, Hirsh-Pasek, Golinkoff, and Lannon (under review) correlated imageability ratings with production data from the MacArthur Communicative Developmental Inventory (CDI; Fenson et al., 1994). After analyzing the 75 nouns and 44 verbs on the CDI that also had imageability ratings (Masterson & Druks, 1998), results revealed that AoA was significantly correlated with imageability. Words with higher imageability ratings, regardless of form class, were learned earlier than words with lower imageability ratings (McDonough et al., under review).

If the imageability construct predicts lexical acquisition, the findings from the work of McDonough et al. (2009) in English should extend to other languages, such as Mandarin Chinese (Ma, Golinkoff, Hirsh-Pasek, McDonough, & Tardif, 2009). An evaluation of early word imageability in Chinese indicated that early nouns have higher imageability ratings than verbs in both English and Chinese children's vocabularies. Furthermore, early Chinese verbs have higher imageability ratings than early English verbs (i.e., they tend to have very specific meanings), which may explain the finding that Chinese children learn more verbs at an earlier age than do English children (Ma et al., 2009).

An understanding of how various types of perceptual salience facilitate word–world mapping not only answers a long-standing question in the word learning literature (i.e., why more nouns than verbs are acquired in English but the proportion is roughly equal in Chinese), but also informs a general theory of word learning.

THE EMERGENTIST COALITION MODEL OF WORD LEARNING

There is significant evidence to suggest that the conceptual prerequisites for relational word learning are in place early. In seeming contrast to this conclusion, however, and despite the difficulty encountered by children when learning relational words such as verbs and prepositions in both naturalistic and laboratory settings, young toddlers nonetheless acquire a significant relational lexicon in their everyday lives. How can this be? We suggest that a unified theory of word learning can account for both relational and less-relational word acquisition. The Emergentist Coalition Model (ECM) of word learning holds that the process of relational language learning is multidimensional: Children draw on attentional, social, and linguistic cues differentially over developmental time to learn words (Hollich, Hirsh-Pasek, & Golinkoff, 2000; see also Lavin, Hall, & Waxman, 2006). Data from multiple studies conducted within the ECM framework suggest that to map words to the world, children first rely primarily on perceptual cues, followed by social and linguistic cues. Based on this model, we can make predictions about the first words learned by children, and explain why less-relational words (such as many nouns) predominate in early vocabulary.

The Role of Perceptual Cues

If the first strategy used by children acquiring language is to preferentially attend to referents that are highly perceptually salient, then words that map to perceptually salient objects or events should appear first. In addition, children in this first stage of language acquisition should be unwilling to map a word to an object that is low in perceptual salience. That is precisely what the data on noun learning suggest. In one study, children learned a label for a novel object at 12 months of age, but only if the object being labeled was the one they found most interesting (children always preferred the most perceptually salient object). Infants did not learn a word for an object that was not perceptually salient until 18 months of age (Hollich et al., 2000). A similar study by Pruden, Hirsh-Pasek, Golinkoff, and Hennon (2006) showed how powerful perceptual salience is at the start of word learning. They found that 10-month-old infants can also learn a name for a salient object. They further discovered, however, that if a boring object is labeled in the presence of an interesting object, babies will systematically *mismap a word onto the salient object!* This work coupled with the ECM and the evidence presented emphasizing the importance of perceptual salience suggests that increasing perceptual salience should help children accurately map words to relational referents.

The importance of perceptual salience to learning relational words has been confirmed across a number of paradigms. For example, one study found that young children (22 months of age) learned verbs when the actions being labeled also produced salient effects (e.g., a light turned on when the action was performed), but not when there was no outcome (Brandone, Pence, Golinkoff, & Hirsh-Pasek, 2007). It was not until 34 months of age that children could learn a name for an action that did not have a salient result. Thus, a salient result that heightens the perceptual salience of a relational referent also facilitates word–world mapping.

The Role of Social Cues

Although attention to perceptual salience is an important strategy for solving the problem of mapping relations to words, and can help young toddlers develop a basic vocabulary, it is often insufficient by itself. For example, the meaning of a word that has an imperceptible referent (e.g., *presume*) is not readily available from the perceptual world. In such cases, children come to rely on social cues (e.g., eye gaze, pointing) in addition to perceptual cues to word meaning.

One particular type of social information that children often use when determining the referent of a word is *intentionality.* An intention is "a mental state or plan that precedes the conduct of an action" (Behrend & Scofield, 2006, p. 290). A significant body of research suggests that infants understand intentions from an early age (Baird & Baldwin, 2001; Baldwin & Baird, 1999; Gergely, Nadasdy, Csibra, & Biro, 1995; Meltzoff, 1995; Woodward, 1998, 1999; Woodward & Somerville, 2000). Although intentions may be understood by very young children, the clearest evidence to date that children can harness their understanding of another's intentions to determine the meaning of a relational word has been found in 2 year olds (Poulin-Dubois & Forbes, 2002; Tomasello & Barton, 1994). In a behavioral reenactment task, toddlers saw a person perform two actions (Tomasello & Barton, 1994). Before the action began, they heard the actor say, "I'm going to *gorp* it." The actor then revealed a social cue to his intentions by performing one action in silence, but saying "oops" after performing the other action. If children understood that the actor intended to "gorp" and that the "oops" was a social signal that the intention was not successfully completed, then they should deduce that the verb "gorp" labeled the action that was not followed by "oops." When children were given the opportunity to "gorp" with the toys after observing the actor, they more often produced the action that was *not* followed by "oops" during training (i.e., the "intended" action) than the "oops" action (i.e., the unintended or mistaken action; Tomasello & Barton, 1994).

Can children utilize social cues to an actor's intentions to learn a novel verb? Poulin-Dubois and Forbes (2002) showed 21- and 28-month-old children pairs of novel action verbs on a video. The extent to which the actions appeared perceptually similar and were intentional versus nonintentional was manipulated (e.g., *topple* and *knock over* are perceptually similar but the intention of the actor distinguishes the verbs). Children were familiarized with each action and heard audio labeling the action. At test, a split screen showed two actions side by side, and children were asked to find the target verb. Results revealed that by

27 months of age, toddlers are able to use subtle social cues to intentionality to decide which of two perceptually similar actions is the correct referent of a novel verb.

Linguistic Cues

A third source of information that children can access when solving the problem of mapping word to world is linguistic in nature. Numerous studies have revealed that 2- to 5-year-old children interpret novel verbs depending on the syntactic structure in which they hear the verb (Fisher, 2002; Naigles, 1996; Naigles & Kako, 1993), a strategy known as "syntactic bootstrapping" (Gleitman, 1990). In addition to evidence showing that toddlers are able to use abstract linguistic information (e.g., the number of noun phrases in a sentence) to determine the meaning of a verb (Fisher, 1996; Fisher, 2002), 16- to 18-month-old children are able to use word order information to determine the correct scene described by a transitive sentence such as "Big Bird is tickling Cookie Monster" (Hirsh-Pasek & Golinkoff, 1996).

In conclusion, the ECM does not make predictions about how easily a word can be mapped onto a referent based on form class membership. Instead, the ECM provides a framework for looking at the trajectory of word mapping over time. Referents that are more relational will be more difficult to connect to words than referents that are less relational. That is, those words that refer to more perceptually available referents will be learned first, whether they are nouns, verbs, or prepositions. With less perceptual guidance, word learning must await support from social and linguistic cues to indicate word meaning. Because relational referents have ambiguous word–world mappings, children rely on a coalition of cues to solve the word learning problem.

CONCLUSIONS

For decades, researchers have sought a theory of vocabulary acquisition that accounts equally well for concrete words such as many nouns and relational words such as many verbs and prepositions. In this chapter, we used the noun–verb debate to inform a general theory of word learning. We explored evidence in support of two hypotheses to explain why some words might be learned earlier than others: the *conceptual prerequisites* hypothesis and the *mapping* hypothesis. After a thorough review of the literature on *containment/support, path/manner, source/goal,* and *figure/ground,* we concluded that children demonstrate primitive, language-relevant concepts from an early age (Mandler, 2004), and are able to form categories of concepts including action categories from a very young age (Pruden et al., 2004). Thus, the *conceptual prerequisites* hypothesis cannot fully account for the late appearance of relational words in children's vocabularies.

Our review of the literature on *mapping,* on the other hand, revealed that children learning English, Japanese, and Chinese seem to have similar levels of difficulty mapping verbs to action categories in the laboratory (Imai et al., 2006), and even adults have difficult mapping verbs to actions in ambiguous situations (Gillette et al., 1999). Thus, although children appear to have the conceptual foundations for relational word learning at a very early age, they have trouble connecting words to the world.

The ECM explains the mechanisms behind early word acquisition, the reasons why children acquire the words they do, and why some words are learned earlier than others. It not only offers a framework for understanding both relational and less relational terms, but also explains several themes in the literature. Why are nouns easier to learn than verbs? Because the majority of nouns are more perceptually accessible than are the majority of verbs. Why do verbs nonetheless exist in the earliest vocabularies of young children? We argue that the verbs appearing in children's earliest vocabularies are more perceptually rooted than verbs that are acquired later. In fact, even "broad verbs" that appear early in lexical development—such as *go* or *make*—might initially be understood in limited, perceptually governed ways (Theakston, Lieven, & Pine, 2002). Finally, why do children learning Mandarin have proportionally more verbs in their early vocabularies than children learning English? Perhaps because the verbs used in Mandarin are more perceptually rich and contextually bound than the verbs used in English. The ECM offers a viable and general model for explaining word development by accounting for how both relational and less-relational words are mapped to the world, and thereby provides answers to long-debated questions in the field of language acquisition.

Understanding the processes that underlie the acquisition of words from all form classes not only is important to the field of language development in general, but also provides more information on the etiology of atypical language. For example, given that perceptual salience is highly important for initial word learning, might language-delayed children benefit from an intervention that focuses on learning words for salient objects and actions first, to give them a foothold into vocabulary? Children diagnosed with autism have communication problems (American Psychiatric Association, 1994), and may have particular trouble with relational language (Ishikawa & Uda, 1996; Takata & Okuma, 1986; Watake, 1996). Given the importance of access to social and grammatical cues to relational word meaning, and the well-established variability of skill demonstrated by children with autism in those areas (Baron-Cohen, 1995), our theory might explain any difficulty with relational language found in this population. Indeed, recent research has demonstrated that children with autism are able to learn words for perceptually salient objects (similar to young typically developing children, Hollich et al., 2000) but have difficulty using social cues to determine word meaning (Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff, & Tager-Flusberg, 2007).

Based on our review of what is known about conceptual primitives and the processes underlying relational language acquisition, where should the field direct its future research efforts? Talmy (2000) listed numerous conceptual primitives that underlie the concepts ultimately packaged into categories and codified by language. Exploring each and every posited primitive has the potential to further illuminate our understanding of how both typical and atypical language develops in children across the world. This chapter reviews the first steps that have been taken toward understanding these primitives, advances the field by viewing this evidence through the framework of a unified theory of word learning, and suggests a clear direction for research that can guide us toward a better understanding of our most fundamental human faculty: language.

ACKNOWLEDGMENTS

The research reported in this chapter was supported by NSF Grants SBR-961-5391, SBR-990-5832, and BCS0642529 to K. Hirsh-Pasek and R. Golinkoff and by NIH Grant 1-R01-HD-050199-0182 to A. Chatterjee, K. Hirsh-Pasek, and R. Golinkoff. We thank Barbara Malt and Phillip Wolff for their many helpful comments on earlier versions of this chapter. We would like to thank all of the parents and children at Temple University and the University of Delaware who participated in our studies.

References

Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology, 39*, 116–157.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Baillargeon, R. (1998). Infants' understanding of the physical world. In M. Sabourin & F. Craik (Eds.), *Advances in psychological science, Volume 2: Biological and cognitive aspects* (pp. 503–529). Hove, UK: Psychology Press/ Erlbaum.

Baillargeon, R. (2001). Infants' physical knowledge: Of acquired expectations and core principles. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 341–361). Cambridge, MA: MIT Press.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 47–83). Malden, MA: Blackwell Publishers.

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science, 13,* 89–94.

Baillargeon, R. & Hanko-Summers, S. (1990). Is the top object adequately supported by the bottom object? Young infants' understanding of support relations. *Cognitive Development, 5*(1), 29–53.

Baillargeon, R., Needham, A., & DeVos, J. (1992). The development of young infants' intuitions about support. *Early Development & Parenting, 1*(2), 69–78.

Baillargeon, R., & Wang, S. (2002). Event categorization in infancy. *Trends in Cognitive Science, 6,* 85–93.

Baird, J.A. & Baldwin, D.A. (2001). Making sense of human behavior: Action parsing and intentional inference. In: Intentions and intentionality: Foundations of social cognition. Malle, Bertram F. (Ed.); Moses, Louis J. (Ed.); Baldwin, Dare A. (Ed.); Cambridge, MA, US: The MIT Press, pp. 193–206.

Baldwin, D.A. & Baird, J.A. (1999). Action analysis: A gateway to intentional inference. In: Early social cognition: Understanding others in the first months of life. Rochat, Philippe (Ed.); Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 215–240.

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press.

Behrend, D.A. & Scofield, J. (2006). Verbs, Actions, and Intentions. In: Action meets word: How children learn verbs. Hirsh-Pasek, Kathy (Ed.); Golinkoff, Roberta Michnick (Ed.); New York, NY, US: Oxford University Press, pp. 286–307.

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: the importance of being imageable. *Journal of Neurolinguistics, 16*(2/3), 113–150.

Blewitt, P. (1983). Dog versus collie: Vocabulary in speech to young children. *Developmental Psychology, 19*(4), 602–609.

Bloom, L., Tinker, E., & Margulis, C. (1994). The words children learn: Evidence against a noun bias in children's vocabularies. *Cognitive Development, 8,* 431–450.

Bornstein, M. H., Arterberry, M. E., & Mash, C. (2007). Infant perception and categorization of object-context relations. In *Proceedings of the IEEE 6th International Conference on Development and Learning* (pp. 99–104). New York: IEEE.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Sung-Yun, P., Pascual, L., Pêcheux, M., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75(4),* 1115–1139.

Bowerman, M. (1996). Learning how to structure space for language: A cross-linguistic perspective. In P. Bloom, M.A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and Space* (pp. 385–436). Cambridge, MA: MIT Press.

Bowerman, M. & Choi, S. (2003). Space under construction: Language-specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 387–427). Cambridge, MA, US: MIT Press.

Brandone, A., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action speaks louder than words: Young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development, 78,* 1322–1342.

Braunwald, S.R. (1995). Differences in the acquisition of early verbs: Evidence from diary data from sisters. In: Beyond names for things: Young children's acquisition of verbs. Tomasello, Michael (Ed.); Merriman, William Edward (Ed.); Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc, 1995. pp. 81–111.

Casasola, M. (2005a). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology, 41,* 183–192.

Casasola, M. (2005b). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development, 76,* 279–290.

Casasola, M. (2008). The Development of Infants' Spatial Categories. *Current Directions in Psychological Science, 17*(1), 21–25.

Casasola, M., & Bhagwat, J. (2007). Does a novel word facilitate 18-month-olds' categorization of a spatial relation? *Child Development, 78,* 1818–1829.

Casasola, M., & Cohen, L. B. (2002). Infant categorization of containment, support, and tight-fit spatial relationships. *Developmental Science, 5,* 247–264.

Casasola, M., Cohen, L. B., & Chiarello, E. (2003). Six-month-old infants' categorization of containment spatial relations. *Child Development,* 74, 1–15.

Casasola, M., Hohenstein, J., & Naigles, L. (2003). Ten-month-old infants' discrimination of manner and path in motion events. In M. Casasola (Chair), *From infancy to adulthood: Exploring the effect of linguistic input of the discrimination of manner and path in motion event.* Symposium presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.

Casasola, M., & Wilbourn, M.P. (2004). 14-Month-Old Infants Form Novel Word-Spatial Relation Associations. *Infancy, 6*(3), 385–396.

Casasola, M., Wilbourn, M. P., & Yang, S. (2006). Can English-learning toddlers acquire and generalize a novel spatial word? *First Language, 26 (2),* 187–205.

Choi, S. (1998). Verbs in early lexical and syntactic development in Korean. *Linguistics, 36*(4), 755–780.

Choi, S. (2006). Preverbal spatial cognition and language-specific input: Categories of containment and support. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41,* 83–121.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22,* 497–529.

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. M. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development, 14,* 241–268.

Clark, E.V. (1996). Early verbs, event types, and inflections. In: Children's language, Vol. 9. Johnson, Carolyn E. (Ed.); Gilbert, John H. V. (Ed.); Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc, pp. 61–73.

Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27,* 111–133.

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition, 72*(3), 237–67.

Fenson, L., Dale, P. S., Reznick, J. S., & Bates, E. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55–64.

Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: some simple relationships. *Journal Of Child Language, 6*(3), 423–42.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj, II (Ed.), *Language development. Vol. 2: Language, thought, and culture* (pp. 301–334). Hillside, NJ: Lawrence Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). New York: Cambridge University Press.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165–93.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12(4),* 395–427.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73,* 135–176.

Gleitman, L. (1990). Structural sources of verb meaning. *Language Acquisition, 1,* 3–55.

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words. *Language Learning and Development, 1(1),* 23–64.

Göksun, T., Hirsh-Pasek, K., Roseberry, S., & Golinkoff, R. M. (2008). Processing events and learning relational terms: Figures are more prominent than grounds. Paper presented at S. Pruden & T. Göksun (chairs), *Conceptual primitives for processing events and learning relational terms.* Symposium at the XVIth International Society on Infant Studies, Vancouver, Canada.

Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., Maguire, M. J., & Hennon, E. A. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology, 4,* 604–615.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L., & Wenger, N. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28,* 99–108.

Golinkoff, R. M., Hirsh-Pasek, K., Mervis, C. B., & Frawley, W. B. (1995). Lexical principles can be extended to the acquisition of verbs. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 185–221). Hillsdale, NJ: Lawrence Erlbaum.

Golinkoff, R. M., Jacquet, R., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development, 67,* 3101–3119.

Hall, D. G., & Waxman, S. R. (1993). Assumptions about word meaning: Individuation and basic-level kinds. *Child Development, 64,* 1550–1575.

Hespos, S. J., & Baillargeon, R. (2001a). Knowledge about containment events in very young infants. *Cognition, 78,* 207–245.

Hespos, S. J., & Baillargeon, R. (2001b). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science, 12,* 140–147.

Hespos, S.J. & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition, 99* (2), B31-B4.

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

Hirsh-Pasek, K. & Golinkoff, R. (1996). How children learn to talk. The World Book Encyclopedia, Health and Medical Annual. Chicago: World Book, Inc. 92–106.

Hirsh-Pasek, K. & Golinkoff, R.M. (Eds.). (2006). Action meets word: How children learn verbs. New York, NY, US: Oxford University Press, xvi, 588 pp.

Hollich, G., Golinkoff, R., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology, 43,* 1051–1062.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (With Hennon, E., Chung, H. L., Rocroi, C., Brand, R. J., & Brown, E.). (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65* (3, Serial No. 262).

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90*(1), 72–93.

Ihara, H., & Fujita, I. (2000). A cognitive approach to errors in case marking in Japanese agrammatism: The priority of the goal -ni over the source -kara. In A. Foolen & F. van der Leek (Eds.), *Constructions in cognitive linguistics: Selected papers from the Fifth International Cognitive Linguistics Conference* (pp. 123–140). Amsterdam: John Benjamins.

Imai, M., Haryu, E., & Okada, H. (2002). Is verb learning easier than noun learning for Japanese children?: 3-year-old Japanese children's knowledge about object names and action names. *Proceedings of the 26th Annual Boston University Conference on Language Development, 1,* 324–335.

Imai, M., Haryu, E., Okada, H., Lianjing, L., & Shigematsu, J. (2006). Revisiting the noun-verb debate: A crosslinguistic comparison of novel noun and verb learning in English-, Japanese,- and Chinese-speaking children. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs,* New York: Oxford University Press.

Imai, M., Okada, H., & Haryu, E. (2005). Mapping Novel Nouns and Verbs Onto Dynamic Action Events: Are Verb Meanings Easier to Learn Than Noun Meanings for Japanese Children? *Child Development, 76*(2), 340–355.

Ishikawa, Y. & Uda, M. (1996). The use of everyday situations to teach 2-word sentences: Parent participation in the training. *Japanese Journal of Special Education, 33*(5), 65–71.

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

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

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal learning and Verbal Behavior, 23,* 221–236.

Kersten, A. W., Smith, L. B., & Yoshida, H. (2006). Influences of object knowledge on the acquisition of verbs in English and Japanese. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs.* New York: Oxford University Press.

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

Lakusta, L., Wagner, L., O'Hearn, K., & 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., Smith, L., & Jones, S. (1998). Object shape, object function, and object name. *Journal of Memory and Language, 38*(1), 1–27.

Langacker, R. W. (1987). Nouns and verbs. *Language, 63,* 53–94.

Lavin, T.A., Hall, D.G., & Waxman, S.R. (2006). East and west: A role for culture in the acquisition of nouns and verbs. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

Lidz, J. (2006). Verb Learning as a Probe Into Children's Grammars. In: Action meets word: How children learn verbs. Hirsh-Pasek, Kathy (Ed.); Golinkoff, Roberta Michnick (Ed.); New York, NY, US: Oxford University Press, pp. 429–449.

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language, 6*(2), 405–423.

Maguire, M.J., Golinkoff, R. M., & Hirsh-Pasek, K. (2006). It's not about nouns and verbs. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

Maguire, M.J., Hennon, E.A., Hirsh-Pasek, K., Golinkoff, R.M., & Slutzki, C. B. (2002). Mapping words to actions and events: how do 18-month-olds learn a verb? In B. Skarabela, S. Fish, & A. H. J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language*, Vol. 1. Somerville, MA: Cascadilla Press.

Maguire, M.J, Hirsh-Pasek, K., & Golinkoff, R. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

Mandler, J. M. (1991). Prelinguistic primitives. In L. A. Sutton & C. Johnson (Eds.), *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society* (pp 414–425). Berkeley, CA: Berkeley Linguistics Society.

Mandler, J. M. (1992). Foundations of conceptual thought in infancy. *Cognitive Development, 7*, 273–285.

Mandler, J. M. (2004). A synopsis of The Foundations of Mind: Origins of conceptual thought. *Developmental Science, 7(5),* 499–505.

Mandler, J.M. (2006). Actions organize the infant's world. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word* (pp. 111–133). New York: Oxford University Press.

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 72–106). New York: Cambridge University Press.

Masterson, J., & Druks, J. (1998). Description of a set of 164 nouns and 102 verbs matched for printed word frequency, familiarity and age-of-acquisition. *Journal of Neurolinguistics, 11,* 331–354.

McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology, 46,* 229–259.

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (in prep.). An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning.

Meltzoff, A.N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*(5), 838–850.

Merriam-Webster. (1998). Merriam-Webster's Collegiate Dictionary, 10th edition.

Meyer, M., Leonard, S., Hirsh-Pasek, K., Imai, M., Haryu, R., Pulverman, R., & Addy, D. (2003). *Making a convincing argument: A cross-linguistic comparison of noun and verb learning in Japanese and English*. Poster presented at the Boston University Conference on Language Development, Boston, MA.

Muehleisen, V., & Imai, M. (1997). Transitivity and the Incorporation of Ground Information in Japanese Path Verbs. In: Keedong Lee, Eve Sweester and Marjolijn Verspoor, eds., *Lexical and Syntactical Constructions and the Construction of Meaning,* 329–346. Amsterdam: John Benjamins.

Naigles, L.R. (1996). The Use of Multiple Frames in Verb Learning via Syntactic Bootstrapping. *Cognition, 58*(2), 221–51.

Naigles, L.R., Eisenberg, A.R., Kako, E.T., Highter, M., & Mcgraw, N. (1998). Speaking of Motion: Verb Use in English and Spanish. *Language & Cognitive Processes, 13*(5), 521–549.

Naigles, L.G. & Kako, E.T. (1993). First Contact in Verb Acquisition: Defining a Role for Syntax. *Child Development, 64*(6), 1665–1688.

Nam, S. (2004). *Goal and source: Asymmetry in their syntax and semantics*. Paper presented at the Workshop on Event Structure, Leipzig, Germany.

Nelson, K. (1989). *Narratives from the crib.* Cambridge, MA: Harvard University Press.

Oakes, L. M., & Rakison, D. H. (2003). Issues in the study of categories and concepts in early development: An introduction. In D. H. Rakison & L. M. Oakes, (Eds.), *Early concept and category development: Making sense of the blooming, buzzing confusion*. New York: Oxford University Press.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108(1)*, 155–184.

Parish-Morris, J., Hennon, E., Hirsh-Pasek, K., Golinkoff, R. M., & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intent in word learning. *Child Development, 78(4)*, 1265–1287.

Poulin-Dubois, D. & Forbes, J.N. (2002). Toddlers' attention to intentions-in-action in learning novel action words. *Developmental Psychology*, *38*(1), 104–114.

Pruden, S.M. (2007). Finding the action: Factors that aid infants' abstraction of path and manner. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 68(3–B), 1964.

Pruden, S.M., Hirsh-Pasek, K., Golinkoff, R.M., & Hennon, E.A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development, 77(2)*, 266–280.

Pruden, S. M., Hirsh-Pasek, K., Maguire, M. J., & Meyer, M. A. (2004). Foundations of verb learning: Infants form categories of path and manner in motion events. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 461–472). Somerville, MA: Cascadilla Press.

Pulverman, R., & Golinkoff, R. M. (2004). Seven-month-olds' attention to potential verb referents in nonlinguistic events. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (pp. 473–480). Somerville, MA: Cascadilla Press.

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Sootsman-Buresh, J. L. (2008). Manners matter: Infants' attention to manner and path in non-linguistic dynamic events. *Cognition, 108*, 825–830.

Pulverman, R., Hirsh-Pasek, K., Golinkoff, R. M., Pruden, S. M., & Salkind, S. J. (2006). Conceptual foundations for verb learning: Celebrating the event. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. New York: Oxford University Press.

Pulverman, R., Sootsman, J. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2003). The role of lexical knowledge in nonlinguistic event processing: English-speaking infants' attention to manner and path. In *Proceedings of the 27th Annual Boston University Conference on Language Development* (pp. 662–673). Somerville, MA: Cascadilla Press.

Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press.

Regier, T., & Zheng, M. (2003). *An attentional constraint on spatial meaning.* Paper presented at the Proceedings of the 25th Annual Meeting of the Cognitive Science Society, Boston, MA.

Ringler, N., Trause, M., Klaus, M., & Kennell, J. (1978). The Effects of Extra Postpartum Contact and Maternal Speech Patterns on Children's IQs, Speech, and Language Comprehension at Five. *Child Development, 49*(3), 862–865.

Salkind, S.J., Sootsman, J. L., Golinkoff, R.M., Hirsh-Pasek, K., & Maguire, M.J. (2002). *Lights, camera, action! Infants and toddlers create action categories*. Poster presented at the biennial meeting of the International Conference on Infant Studies, Toronto, Canada.

Sandhofer, C.M., Smith, L.B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language, 27*, 561–585.

Sharon, T., & Wynn, K. (1998). Individuation of actions from continuous motion. *Psychological Science, 9*, 357–362.

Shipley, E.F., Kuhn, I.F., & Madden, E.C. (1983). Mothers' use of superordinate category terms. *Journal of Child Language, 10*(3), 571–588.

Slobin, D.I. (2001). Form-function relations: How do children find out what they are? In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 406–449). New York: Cambridge University Press.

Smiley, P., & Huttenlocher, J. (1995). Conceptual development and the child's early words for events, objects, and persons. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 21–61). Hillsdale, NJ: Lawrence Erlbaum.

Smith, L.B. (2000). Learning how to learn words: An associative crane. In R.M. Golinkoff, K. Hirsh-Pasek, L. Bloom, A. Woodward, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Breaking the word learning barrier: What does it take?* (pp. 51–80). New York: Oxford University Press.

Snedeker, J., & Gleitman, L. R. (2004). Why is it hard to label our concepts? In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 257–293). Cambridge, MA: MIT Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description* (Vol. 3, pp. 57–149). New York: Cambridge University Press.

Talmy, L. (2000). *Toward a cognitive semantics.* Volume I: *Concept structuring systems* (pp i–viii, 1–565). Cambridge, MA: MIT Press.

Takata, H. & Okuma, H. (1986). Building verbal behavior for one adolescent autistic boy: Some magnet movable equipments for teaching him many verbs. *Kyushu Neuro-psychiatry, 32*(3–4), 341–348.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32,* 492–504.

Tardif, T., Gelman, S.A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development, 70,* 620–635.

Tardif, T., Shatz, M., & Naigles, L.R. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24,* 535–565.

Theakston, A.L., Lieven, E.V.M., & Pine, J.M. (2002). Going, going, gone: The acquisition of the verb 'go'. *Journal of Child Language, 29(4),* 783–811.

Tomasello, M., Akhtar, N., Dodson, K., & Rekau, L. (1997). Differential productivity in young children's use of nouns and verbs. *Journal of Child Language, 24*(2), 373–387.

Tomasello, M. & Barton, M.E. (1994). Learning words in nonostensive contexts. *Developmental Psychology, 30*(5), 639–650.

Tomasello, M., & Kruger, A. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19,* 311–333.

Wagner, L., & Carey, S. (2005). 12-month-old infants represent probable endings of motion events. Infancy, *7(1)*, 73–83.

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition, 93* (3), 167–198.

Watake, S. (1996). A case study of an autistic boy who used finger writing for communication. *Japanese Journal of Special Education, 33*(5), 33–39.

Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34.

Woodward, A.L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior & Development, 22* (2), 145–160.

Woodward, A.L. & Sommerville, J.A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11*(1), 73–78.

Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science, 7,* 164–169.

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition, 2,* 145–175.

12 LEARNING A LANGUAGE THE WAY IT IS

Conventionality and Semantic Domains

Eve V. Clark

Acquiring a vocabulary is an undertaking essential to the learning of every language. But consider the challenge presented to children by the many ways in which languages partition the world. Languages differ in how they divide up domains, how many terms they include in each domain, and how these terms are related to each other. In addition, cultures differ in how they view children and hence the circumstances under which adults talk to them. How do children learn which words belong in each domain or work out how the words in a domain are related to one another? How do they acquire specific semantic domains? Where do they get information about word meanings and semantic relations from? The focus in this chapter is on how children acquire and learn to use a language, and in particular how the language addressed to them influences their uptake of new words and their organization of those words into semantic domains.

While children make use of perceptual and cognitive factors in their early hypotheses about possible word meanings, their acquisition of the vocabulary of a language relies critically on the usage they are exposed to. Adult speech is a primary source of information for children about the forms of a language—words, phrases, and whole utterances—and about their usage. Children take in words from the speakers who talk to them, and the uses of the words they hear shape the way they set up the relevant lexical domains in that language.

As children build up a vocabulary, they also need to find out how to *use* each of the terms being acquired. Learning a list of forms is not enough: They must learn which meanings are customarily, conventionally, carried by each form, and how people generally expect each one to be used in the speech community. After considering how much vocabulary children have typically learned by the time they reach adulthood, and how they might find out what meanings an unfamiliar word conveys, I take up conventions: what conventions are, the basic role they play in communication, and how they are learned. I then consider some of the ways languages differ from each other, and whether these differences pose any problems in the setting up of semantic domains as children learn the vocabulary of their first language(s). In the second half of the chapter, I look at some data from English-speaking parents: how they flag new words as new, the kinds of added information they supply along with such words, and how both these sources of information could inform children's setting up of semantic domains. I conclude with some discussion of how children learn the specific semantic packaging characteristic of their language, and how this is related to the cognitive and perceptual information they also draw on as they set up initial possible meanings.

BUILDING A VOCABULARY

Some Numbers

Consider first the size of the vocabulary children are expected to learn by adulthood in many parts of the Western world. By the age of 2, children can produce on average between 100 and 600 words. By the age of 6, they are estimated to understand around 14,000 words, with somewhat fewer in production. Then, from age 6 up to 18 years (while in school), they typically add about 3000 words a year to their repertoire, for an accumulated total of around 50,000 words. In addition, from age 12 to 17 years, they are exposed to some 10,000 stems and affixes a year for comprehension, from school textbooks alone. Educated English-speaking adults typically have an estimated vocabulary of 50,000 to 100,000 words (Anglin, 1993; Clark, 1993; Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994).

The size of children's vocabularies appears to depend critically on how much speech is addressed to them during early development. On average, from 13 to 36 months of age, children acquiring English hear some 1440 words per hour addressed directly to them. (This measure does not include overheard speech from conversations among the people around them.) In the course of an hour, children hear some 340 utterances, including 90 questions, 62 directives, and 105 assertions. They also hear, on average, 17 offers of approval and 7 prohibitions per hour (Hart & Risley, 1995).

But these numbers differ with socioeconomic status. Professional families address 2153 words per hour to their children, working class families 1251 words, and welfare families 616 words. These differences accumulate over the first 4 years of children's language use as shown in Table 12.1 (based on Hart & Risley 1995), and are related to differences in the children's growth of vocabulary from 10 months to age 3 years (Hart & Risley, 1995, Appendix B). The higher socioeconomic (SES) children averaged around 1100 words in production at age 3 years, the middle/lower SES ones about 750 words, and the welfare children a little over 500 words. Amount of parental speech during their first few years was highly correlated with children's vocabulary size on entry to school, as well as over the next few years.

Finding Meanings for Unfamiliar Words

In the earliest stages, nearly all the words children hear are unfamiliar: They have yet to learn their meanings and how to use them. As time goes by, children hear many unfamiliar words and generally have to assign them some meaning on the fly. But adults also offer some new words overtly flagged as new words. And because children attend closely to such offers (Clark, 2007a), these provide important clues to the general process of uptake during acquisition. The attention children pay suggests that they quickly learn to identify the cues that identify a new word as new, so they can begin to make inferences about (some of) the meaning it usually carries. That is, they must decide what the new word most probably refers to on that occasion, come up with a reasonable preliminary assignment of meaning, retrieve any other (already familiar) word meanings it might be connected to, and identify its word class (Clark, 1993). Of course, they must do this with all the unfamiliar words they encounter, whether or not they have been flagged as new by adult speakers. Where the flagging is overt, though, it is easier to observe the process of uptake in children.

When adults offer children new words directly, flagging them as new, they also often

TABLE 12.1. Cumulative Exposure to Language[a]

Status	One Week	One Year	Four Years
Higher socioeconomic status	215,000	11 million	44 million
Middle/lower socioeconomic status	125,000	6 million	24 million
Welfare	62,000	3 million	12 million

[a]Measured in number of words directed to the child, by social class, extrapolated to a 100-hour week, a year, and 4 years.

FIGURE 12.1. A California Thrasher.

license a variety of inferences, in context, about the meanings of those words. They do this by adding information relevant to identifying the intended referent on each occasion. For example, an adult might first identify a bird by providing a label for the species—for instance, *That's a California Thrasher*—and then distinguish that bird from others by drawing the child's attention to the shape and length of the beak, which is relatively long with a downward curve, as shown in Figure 12.1.

If adults consistently supply additional information about distinguishing features or properties for members of a domain, this would help children set up a network of connected but distinct entities for all the types of birds they have encountered so far. Such a network would form the starting point for organizing their domain of birds. On other occasions, adults might fill in details of how to talk about other domains such as tools or boats, containers or trees; activities such as hitting, kicking, or throwing; movement along a path or in a particular manner; characteristic sounds, properties, or substances; habitats; or spatial locations. Much or all such information could serve to (1) distinguish the first members of a domain, and then (2) provide dimensions for organizing information pertinent to all members of that domain. The added information adults supply would license inferences about the target referent on each occasion, and so help children both identify the referent of the unfamiliar word and work out the meaning intended by the speaker in each occasion.

Much, and perhaps most, of children's new-word uptake is actually through indirect offers, where speakers presuppose that child-addressees already know the words, idioms, or constructions being used (see Clark, 2007a). On those occasions, children should be even more dependent on the surrounding information, linguistic and nonlinguistic, that is given along with any unfamiliar words: They need to use it to make inferences about the most probable new references and meanings in context. They can check any inferences they have made against subsequent uses they hear on later occasions. The process of acquiring a meaning is a drawn-out affair, and can last for years: Even adults have only partial meanings for many of their words in domains where they have had little reason to learn more than the minimum needed for basic communication.

With both direct and indirect offers of new words, children can draw on whatever additional information is available after their initial mapping of word to world (Clark, 2002, 2010). They observe the usage of target terms by others, alongside their own knowledge of any already-familiar neighboring terms and expressions. For children, the social incentive for attending to adult speech, hence to adult *conventions* on usage, supports their impulse to communicate effectively with the people around them. As a result, the partitioning of the world that children adopt in learning the conventions of each language is the one they hear from other speakers. Adult conventions provide the model of usage for children as they learn to think for speaking.

CONVENTIONS

Conventions are regularities in how people do something, whether driving (which side of the road, which traffic rules?), eating (which implements, which hand?), or talking about sailing (which terms?). These regularities can be defined formally as in Lewis (1969: 42):

A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a convention if and only if, in instances of S among members of P,

- everyone conforms to R;
- everyone expects everyone else to conform to R;
- everyone prefers to conform to R on condition that the others do, since S is a coordination problem and uniform conformity to R is a coordination equilibrium in S

There are many regularities in our behavior that fit Lewis' definition of convention. Conventions can hold for just two people (e.g., when and where to meet for coffee on Mondays), for larger groups (when to hold orchestra rehearsals), or for whole communities (when to have the local weekly market).

One domain in which we all rely on conventions is in our daily use of language. There, the conventions consist of a set of agreements, observed by the whole speech community, on how to use that language. The agreements cover vocabulary (which terms are normally to be used for which meanings), syntactic constructions (which combinations of terms convey meanings such as transitive causative action, locative motion towards a goal, the role of agent, and so on), pronunciation (the sound system for that dialect of the language), and all other general patterns of grammatical usage for conveying long agreed-on meanings (e.g., the inflectional system for person and number, case marking for grammatical roles, aspect and tense, evidentials, and gender).

Whereas some conventions are set up explicitly (which side of the road to drive on, for instance), others may evolve from tacit agreements that from then on are transmitted further by example. This appears to be the general case for language. We rarely stop to ask why a specific word has come to have a particular meaning, or ponder why it is carried by a particular sequence of sounds. Instead, we tend to take the conventions for granted, and, as children, simply learn what each word denotes and how to use it by observing more expert speakers.

This is not to say that the conventions of a language are immune to change. The vocabulary of a language, for example, is constantly evolving. Speakers coin new words when these are needed for new inventions, new technologies, new social movements, and new fields of knowledge (consider the impact of aviation in the early twentieth century, or the advent of techniques for studying molecular biology). Some coinages are adopted rapidly and then transmitted to the next generation of speakers; others fail to be taken up so readily or so widely (e.g., Clark, 1993; Clark & Clark, 1979). Speakers also abandon words for which there is no longer any use. Changes in the speakers' world can lead to changes in the conventions that govern some words too, so their meanings may shift considerably with time (as did terms such as *fossil* and *car*). There are also changes in phonology from one generation to the next, and even from one group to the next, as speakers shift their pronunciation to reflect solidarity, for instance. And there are changes in the syntactic constructions favored for specific meanings just as there are in the word-formation patterns speakers choose for coining new words (see Eckert, 2000; Joseph & Janda, 2005).

When children acquire a language, then, they must acquire the conventions that govern current usage in their community. Many speakers acquire the conventions for several dialects (one of which may be regarded as the standard) and for more than one language. For example, a child might hear both *tap* and *spigot* with the same referent but from different speakers (a dialect difference), or *hibou* and *owl*, again with the same referent (here, a language difference). The sources of such conventions are the people who speak to young children—parents, caretakers, siblings, and teachers.

Early in their exposure to language, children appear to recognize and rely on two complementary pragmatic principles (Clark, 1983, 1987, 1993):

Conventionality

For certain meanings there is a form that speakers expect to be used in the language community.

Forms can be words, word endings, idiomatic expressions, syntactic constructions, or any combination of these. Note, though, that conventions work effectively for communication only when different forms are governed by different conventions. Consider what it would be like if we had two different forms with exactly the same meanings for talking about *X*. (These would be true synonyms.) Speakers could in principle choose either one of them, but in fact would end up using one more frequently than the other. This in turn would make the more frequent of the two more accessible than the other one, with the result that it would become the more preferred and hence the more often used of the two. What happens then to the other term? Speakers typically either abandon it or assign it a somewhat different meaning—one that is more specialized, or even more general. Within a language, it turns out to be very difficult to maintain any true synonyms (Clark, 2007b). Overlaps in meaning, however, cause no problem because they do not involve full synonyms—hence the presence in English of such closely related terms as *big* and *large*, *tall* and *high*, and *wide* and *broad*, all in the domain of dimensionality (see Bierwisch, 1967). Close examination shows that the meanings of such terms differ in subtle ways, and each has distinct patterns of use despite areas of extensive overlap.

To capture the treatment of different forms as signaling differences in meaning, adults and children alike rely on the principle of contrast in conjunction with conventionality:

Contrast

Speakers take every difference in form to mark a difference in meaning.

Different words have different meanings. A single word can, of course, have several meanings, but two different words cannot have exactly the same meaning. Conventionality and contrast work together. When children register that a word or construction is unfamiliar, they can assume that it must have a meaning different from that of any forms they already know.

LEARNING CONVENTIONS

Children do not arrive with language already installed and at their disposal for use. And they do not just acquire language out of the blue. Rather, they draw on what they already know. They also draw on what they hear used in the community around them, and they combine their observations of such usage with any available conceptual categories when they set up initial hypotheses about possible meanings (e.g., Clark, 1983, 2005a). As they begin to set up conceptual categories of objects and actions, children attend to how objects are organized and how they move. They draw on perceptual information about canonical orientation (cups stand with their openings up, trees are oriented vertically, table-tops are horizontal), and canonical relations (people walk on the ground, plates sit on flat surfaces, apples can be held in the hand or heaped in containers, coats hang on hooks), adding all this information to their conceptual representations (e.g., Malt, Sloman, & Gennari, 2003). This range of conceptual knowledge, when based on similar ranges of experience across children, will be much the same within and across cultures. But the words children hear used to talk about them can differ considerably from one language to the next.

Each language presents an overlay on top of whatever conceptual categories children have already set up. But each language organizes these overlays differently in each semantic domain because each has different numbers of words available, and also often presents different arrays of grammatical distinctions to be observed. Different languages also have different conventions for how to convey grammatical meaning—through inflections, case marking, and word order, for instance. In learning a language, children must learn what the overlay is for that language, and hence which distinctions to attend to whenever they use that language. This all requires that they master the conventions of use for that language.

Where do children get the conventions from? They must draw on adult speech and the usage they observe there. In fact, by the age of 2, and even earlier, they start to ask what the word is for *X*, with sometimes interminable *What's that?* questions (e.g., Clark, 1983). This suggests that they have already grasped that there is an accepted (conventional) term for each object- and event-type they can observe around them. But to find out how each term is used, they must also track adult usage in the speech addressed to them on each occasion. That is, they need to attend to the precise terms used on each occasion and how the usage of each term differs from the usage of any other term from the same domain.

In the early stages, children are helped in their initial attempts at mapping by the fact that adults offer new words in context, when adult and child are jointly attending to the same entity or event (Estigarribia & Clark, 2007; Veneziano, 2002). These offers are frequently supplemented with additional information linking the new word to others in the same domain, by specifying the semantic relations linking the meanings, and with information about the referent entity—its parts, properties, and functions (e.g., Callanan 1985, 1990; Clark & Wong, 2002; Shipley & Kuhn, 1983; see also Kako, 2005). Sources of information like these provide children with primary material for establishing preliminary meanings for the terms they are learning, for example, that *dog* picks out dogs and that dogs differ from cats, *squirrel* picks out squirrels, *sled* sleds, and *throw* acts of throwing. But identifying further details and dimensions relevant to the conventional meanings of such terms can take months or years. Additional information offered in context will typically license further inferences relevant to the adult meanings.

Children attend to the terms and constructions they hear from the adults talking to them. For instance, in word formation they pick up first on productive options, before relatively unproductive ones, in the coining of new agentive nouns (Clark & Berman, 1984). In syntax, they pick up first on those constructions in parental speech used most commonly with specific verbs, or, for example, those constructions produced most frequently in asking questions (Chenu & Jisa, 2006; de Villiers, 1985; Estigarribia, 2006). In learning to talk about specific actions, they attend closely to the kinds of objects affected by each action as exemplified in adult speech, and make use of that information as they build up meanings for verbs (Bowerman, 2005; Wilkins, 2002) and for nouns (Goodman, McDonough, & Brown, 1998; Kako, 2005).

Within a domain, the effect of adult usage shows up in the terms children pick up early on. Consider relations in space such as "in," "on," "attached to," or "separated from." In English, children start out using various locative particles like *in, on,* and *off* before they combine these with the relevant verbs, usually *put* or *take*. In Korean, they start out with the verbs for "put-in (close-fit)," "put-in (loose-fit)," "put-on," "take-off," "take-out," "attach," or "detach" (see Bowerman, 1996; Choi & Bowerman, 1991; also Choi, McDonough, Bowerman, & Mandler, 1999; Casasola, 2002). Korean maps the meanings of such actions directly onto a set of verbs, where each verb captures a specific locative relation, while English generally combines a single verb, *put,* with a variety of locative particles. This difference in how spatial relations are expressed results in different patterns of lexical acquisition for young children hearing English versus Korean.

Differences in the patterns of uptake can also be seen in children's acquisition of verbs of motion where in some languages, motion is combined with the path followed, as in French *entrer* "go-in" or *sortir* "go-out," while in others, motion is generally combined with manner, as in English *run, stroll, walk,* or *skip,* with the path marked with an added particle, as in *run* ***out*** or *walk* ***in*** (Slobin, 1996b, Slobin & Hoiting, 1994; see also Chenu & Jisa, 2006; Pulverman, Sootsman, Golinkoff, & Hirsh-Pasek, 2002). Adult usage displays to children the terms that must be mapped onto each conceptual domain. Working out the details of this mapping may take time, but once established it provides patterns of use for each form in that language whenever children, like adults, are *thinking for speaking* (Slobin, 1996a), as they plan each utterance.

That is, when speaking, people appear to cut up the world in different ways and make use of different, though at times overlapping, grammatical distinctions. For example, in some languages, speakers mark each noun with a case ending to identify its role—as agent, experiencer, or location, say—and its grammatical status—as subject or object. Others lack such case marking, and speakers rely instead on word order to do much the same work. In others still, speakers use both case and word order to distinguish such roles and grammatical relations. Some languages require that speakers indicate gender on all nouns and on any elements that agree with nouns in gender (e.g., French, Greek). Others require that speakers always indicate whether an action has been completed or not (e.g., Polish, Russian). Others still require that speakers indicate whether they know of a fact or an event from personal experience or from some form of hearsay (e.g., Guaraní, Turkish).

Languages differ in which and how many grammatical distinctions they make, whether and where they mark such distinctions as aspect, tense, gender, and number (Bybee, 1985); whether they exhibit clear patterns of preferred argument structure (DuBois, 1987), and what parts of speech they use (Dixon, 1982; Sadock, 1986). Such typological differences require that speakers habitually attend to whichever facets of experience receive obligatory grammatical encoding in their language (Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996). As Whorf ([1940] 1956) pointed out:

> Users of markedly different grammars are pointed by their grammars towards different types of observations and different evaluations of externally similar acts of observation. ([1940] 1956: 221)

Grammatical categories in a language are obligatory: They are "those aspects of language that *must* be expressed" (Boas, 1938, 127), and speakers have to make use of them whenever they speak. But the set of obligatory grammatical elements differs from one language to another, so speakers of different languages describing a specific event, for example, have to be selective in which details they actually encode, and they therefore represent to someone else only part of what they have actually observed. They are in effect thinking for speaking (Slobin, 1996a).

This point becomes clearer when we consider the relation between the terms available in any one semantic domain for a specific language and the conceptual domain these words map onto. As Slobin (1979: 6) pointed out,

> Language evokes ideas: it does not represent them. Linguistic expression is thus not a straightforward map of consciousness or thought. It is a highly selective and conventionally schematic map.

This selectivity can be seen in the different ways languages cut up semantic domains, assigning different numbers of nouns to the domains for baskets, say, or camels, rice, trees, or containers (e.g., Clark & Clark, 1977; Malt, Sloman, & Gennari, 2003). The vocabulary that the speakers of each language draw on reflects the current inventory of the many uses to which it has been put over time, accumulated over the cultural history of the people using that language. Words needed in the desert differ from those needed by seafarers, farmers, or factory workers. And in each domain, experts rely on a larger and more nuanced vocabulary than novices (e.g., Boster & Johnson, 1989; Johnson & Mervis, 1997). That is, adult speakers of the same language can differ in how fully they have mastered the conventions on word meanings, yet still manage to communicate with each other with little difficulty (see Clark, 2005b; Wolff, Medin, & Pankratz, 1999). Languages differ, and their speakers make use of sometimes very different vocabularies and grammars, and so appear to attend to different ranges among the many possible distinctions that could be drawn.

The task for children trying to master the conventions they need in order to use the language they are learning should therefore be viewed from Slobin's perspective. That is, children need to acquire that "highly selective and conventionally schematic map" of the pertinent conceptual domain for their language, whether this happens to be the domain of buildings, containers, vehicles, or furniture or of animals, plants, body parts, or spatial relations. The map of the domain, in each case, is constituted from the terms and grammatical distinctions available *in that language*. Children must learn those terms, with at least some of the conventional meaning assigned to each one, if they are to master ways of talking about that domain. In addition to the established vocabulary for a domain, they must also master any options available for coining further words, with yet other, contrasting meanings, as needed (see further Brennan & Clark, 1996; Clark, 1993; Malt & Sloman, 2004). One of their first tasks, then, in the process of acquisition, is to identify established words and expressions and work out their conventional meanings in the language spoken around them. Just how children might go about doing this is explored in the next section.

WHEN ADULTS OFFER CHILDREN NEW WORDS: SOME FINDINGS FROM ENGLISH

One source of information about language conventions can be found in adult offers of new words. These offers are typically made with a specific object or event at the locus of their joint attention, during adult–child exchanges (Clark, 2007a; Clark & Wong, 2002). Because the intended referent is the locus of their joint attention, this reduces any indeterminacy about what the adult is referring to on that occasion. These new words provide children with conventional devices for referring to objects or events for which they lack appropriate referring expressions. Adults typically follow up these offers with additional information relevant to the target referent, information that highlights contrasts in meaning with nearby terms and that points to semantic relations linking the new word with various neighbors and connecting it to what is already known. This added material provides children with the basis for further inferences about the meaning of the new word they have just heard.

What leads children to attend to new words? Consider a book-reading task where I asked parents to read a picture book and, when possible, to use the word printed at the top of each page. This allowed me to track how parents went about introducing words from different parts of speech, to observe the specific form in which they introduced members of each word class, and to identify any ancillary information they offered about the intended referent (Clark, 2010).

Formulaic Frames

One signal for children can be found in the forms of the offers they hear. Adults present new words in fixed, formulaic frames that offer information about the referent and the word class of the term used for it. Objects are typically selected with nouns, actions with verbs, properties with adjectives, and relations with prepositions or verbs. The syntactic frames adults favor tend to be fixed for each word class for individual speakers, with each parent exhibiting considerable consistency (Clark, 2010). The syntactic frames in (1) through (4) are typical in adult introductions of new nouns, verbs, adjectives, and prepositions, respectively, in English:

(1) Nouns
  a. The front sail is called the JIB sail.[1]
  b. That's a TERN.

(2) Verbs
  a. What's that owl doing? He's KNITTING a sweater.
  b. The bird is– CHIRPING [+ chirping noises]

(3) Adjectives
  a. That's a man climbing, climbing, climbing, and it's very– STEEP.
  b. Some things are rough 'n some things are–
    —"*Not rough.*"
    Not rough? Another word for "not rough" is SMOOTH.

(4) Prepositions
  a. They're sitting next to each other, BESIDE each other.

b. It's under the umbrella or it's–
—"*In the sand.*"
In the sand, it's BELOW– the umbrella.

Emphatic Stress

Adults rely on two other devices to mark new words as new in English: stress and word order. First, they add emphatic stress to the new word between 60% and 80% of the time, depending on the word class, as shown in Figure 12.2. That is, they place additional stress on the new word, even when it is not in final position. Final position is where speakers of English place new information, usually marked with added stress. (They never place stress on given or known information.) Even though verbs and prepositions do not normally carry stress, unless they carry contrastive information, adults add stress to them as well, 62% and 60% of the time, respectively. And they do this for all four age groups (Clark, 2010).

Utterance-Final Position

Second, adults place new information, in this instance, new words, in utterance-final position, as shown in Figure 12.3. New nouns are readily placed in final position as bearers of new information in English, and adults did this 80% of the time in their offers of new nouns. Predicate adjectives are also often produced in clause- or utterance-final position, and adults did this 56% of the time. But they also went to some lengths to place verbs (54%) and even prepositions (28%) in utterance-final position, as can be seen in Figure 12.3. For example, they would introduce transitive verbs as if they were intransitive, so they would come last, and then repeat them immediately in transitive form with a following direct object. With prepositions, they would pause right after saying the preposition (with stress), then produce the prepositional phrase in full, as in (4b). Note that given information can occur in final position provided that the new information in the utterance, marked by stress, is nonfinal.

These ways of presenting new words often lead adults to repeat the target word more than once, and, in doing this, they often ratify any attempts by the children to say the new word as well (see Clark, 2007a, 2010). Finally, the devices favored by adults in these

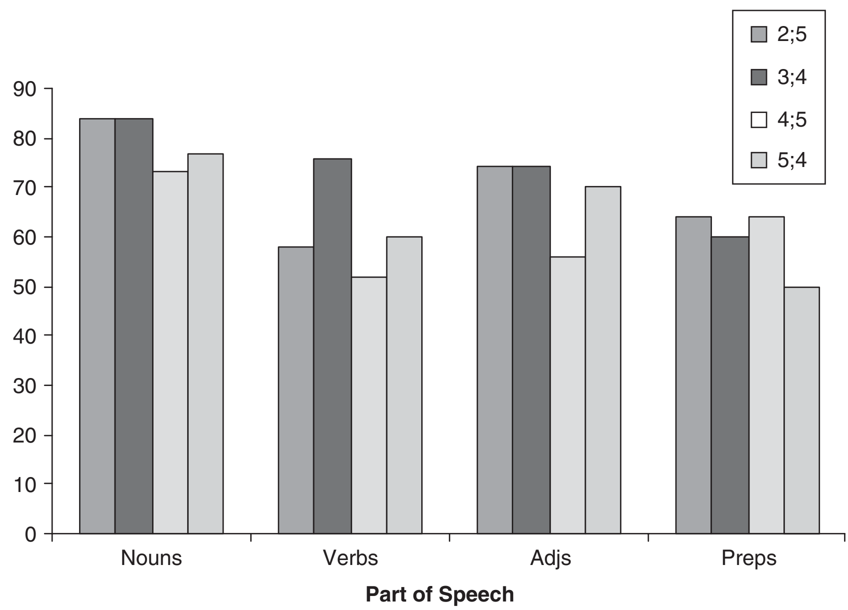


FIGURE 12.2. Adults offer new words with contrastive stress, regardless of position in the utterance.

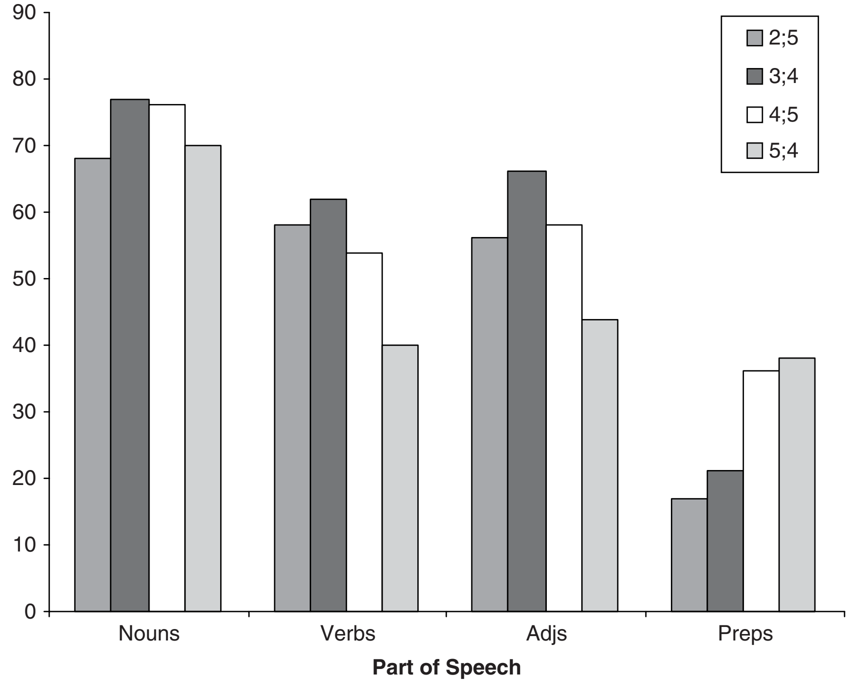


FIGURE 12.3. Adults offer new words in utterance-final position, regardless of part of speech.

book-reading exchanges are fully consistent with those observed for new-word offers in everyday conversation (Clark & Wong, 2002).

Would adults use the same devices to introduce new words in other languages? Not necessarily, since not all languages use both stress and word order to mark new or prominent information. But they would very likely use whatever devices are available for highlighting new information so children would be more likely to attend to what was being offered.

Additional Information

Adults accompany their offers of new words with additional, often distinguishing, information. They mention parts, properties, and relations, and, for example, point out such elements as fins on a sting-ray or wheels on a truck or tractor (see also Ninio, 1983; Masur, 1997), again with considerable consistency in the frames they use. Typical frames for presenting such information are illustrated in (5). In both (5a) and (5b), parents mention the whole object first and then introduce the target part:

(5) a. Mother (to Sarah, 2;3.19[3]): that's his foot.
Sarah: *foot*.
Mother: there's his toes. [Brown/Sarah 3:419][4]
b. Father (to Naomi, 2;7.16): those are cobblestones. that's a street made out of stones. [Sachs 68: 197]

Adults also mention patterns (e.g., stripes or spots) and colors where these are properties of the target-referent. They talk about where an object is in space, and point out characteristic actions and functions. For instance, they talk about whether animals bite or not, how they move, and what kind of noise they make. With objects, they tend to talk about their usual functions, pointing out that tongs are for picking things up, say, or that a sieve is used for straining (Clark & Estigarribia, 2009; Clark

& Wong, 2002). Again, a typical frame for such information is shown in (6a), with a more discursive account of function in (6b):

(6) a. Mother (to Adam, 3;2.0): that's a knife for cutting chicken. [Brown/Adam 24:442]
b. Parent (to child, 2;10), tracing fingers around the inner edge of a strainer): \ and I would maybe put it– maybe some tea leaves here and pour some hot water through . . . so the water comes through with some tea leaves in it but the leaves stay in there. (Clark & Estigarribia, 2009)

Parents of young children appear to offer different types of information about animals (or toys standing in for the real thing) compared to household objects or utensils. For example, they typically mention parts for animals three times as often as for household objects, and they talk about twice as many functions for household objects as for animals (Clark & Estigarribia, 2009). Follow-up talk about the target referent appears to be designed to help children link the unfamiliar object to others that are already familiar, and also helps them link the new word for that object to known words for other objects from the same domain.

Making Use of Contrast

When adults offer added information about referents, they both identify semantic relations and select properties that distinguish near neighbors from the target entity in the relevant semantic field (Akhtar, 2002; Callanan, 1985, 1990; Clark & Wong, 2002; Shipley & Kuhn, 1983). This highlights local contrasts within the pertinent semantic domain. Consider the exchange in (7) between a mother and her child, aged 1;8.12, who is looking at a picture of owls in a new book:

(7) Child: *duck, duck*.
Mother: yeah those are birds. (looks at picture)
they're called owls. (points at picture)
owls, that's their name. owls. (looks at Child)
Child: *birds*.
Mother: and you know what the owl says? (points at picture again) the owl goes "hoo," "hoo."
Child: *owl*.
Mother: that's what the owl says.
Child: *hoo*. (smiles)
Mother: that's right. [NewEng, NE20:0571]

An exchange like this licenses a series of inferences (indicated in boldface) on the part of the child, as shown in (8), (9), and (10) (see Clark, 2002):

(8) "Yeah those are birds. They're called owls . . ."
⇒ Ducks are birds (yeah . . .); *X*s are also birds; *X*s are not ducks but owls; therefore ducks and owls must both belong in the set of birds
► **An owl is a kind of bird**

Effectively, the parent here approves the child's attempt at naming the new type of bird, but does so by using the superordinate term first (*bird*) and then immediately introducing a term (*owl*) that contrasts with the child's term, *duck*. This indicates, first, that ducks are birds, and that the unfamiliar object is also a type of bird, although not a duck. So the inference licensed here is that an owl is a kind of bird.

The parent then goes on to emphasize the new term, *owl*, which is offered simultaneously as a new term and as a repair to the child's proposal.

(9) "They're called owls. Owls, that's their name, owls"
⇒ offer of a contrasting term (as a repair of the child's proposal of *duck*)
► **This is an owl (not a duck as first proposed)**

The parent then proceeds to offer identifying information that can be used to distinguish owls from ducks, namely the characteristic sound made by owls:

(10) "You know what the owl says? The owl goes 'hoo,' 'hoo.'"
⇒ a property of owls that distinguishes them from ducks (and other birds . . .)
► **An owl goes "hoo"**

By highlighting differences between objects from the same domain, for instance a duck and an owl (both birds), a donkey and a horse (both four-legged hoofed mammals), or a cup, mug, and glass (all drink containers), adults focus children's attention on those properties that differentiate one (sub)kind from another. Such comparisons contain direct local contrasts, and these in turn help children as they organize both their conceptual categories and the conventional terms for those categories. The same goes for adult uses of terms for differences in typical types of motion, in characteristic sounds, and in functions (see, e.g., Akhtar, 2002; Kemler Nelson, Chan-Egan, & Holt, 2004; Shipley & Kuhn, 1983). In many exchanges, the contrasts are quite explicit, as in (11) (from Gelman, Coley, Rosengren, Hartman, & Pappas, 1998).

(11) Mother: you know what that one is?
Child (2;11): *Ummm*.
Mother: I don't know if you know what that one is.
Child: *It's a snake*.
Mother: It looks like a snake, doesn't it? It's called an eel.
It's like a snake only it lives in the water. And there's another one.

In this exchange, the parent takes up the child's proposal (*It's a snake*) and adds a warning (*It looks like* . . .) as the prelude to offering a different term, *eel*, and then immediately goes on to provide additional information that distinguishes eels from snakes, namely that eels live in water. This combination of a new word (eel) and distinguishing information that the child can use in the context of the offer, recurs again and again in the new-word offers adults make to young children.

In short, adults "place" new words in semantic fields by presenting children with close neighbors of the target words, as well as with other terms relevant for identifying semantic relations. This provision of semantic neighbors should be particularly helpful as children assign preliminary meanings to new terms and add these terms to domains they have already identified.

Note that this is another place where we may well see differences emerge among languages, in particular in the specific sets of contrasts instantiated in the lexicon. The semantic neighbors adults provide help identify the relevant dimensions of contrast within a domain. Take the meanings of motion verbs: Where these combine with path or direction, contrasting verbs should also pick out combinations of motion with path, say, and mark manner separately or not at all, as in French or Hebrew. But in languages where motion generally combines with manner, contrasting verbs should also pick out manner, and, in these languages, mark path separately.

Semantic Domains

One common activity in which adults provide such information is when they are reading to (very) young children (Clark, 2009; Murase, Dale, Ogura, Yamashita, & Mahieu, 2005; see also Fernald & Morikawa, 1993). Consider some typical examples of how adults in fact offered new words along with closely associated semantic neighbors in English, in (12)–(15) in the book-reading task mentioned earlier. In (12), the parent of a 3-year-old child offered one of the target nouns, *jib*.

(12) Parent to child (3;6.20), picture of a sailboat:
Now this is the back sail [points], this is the front sail [points], and this [points at the front sail again] has a special word to describe it.
D'you know what it is?
Child: *I don't know*.
Parent: It's called a JIB.

The commonest near neighbors offered with this word were first the general noun *sail*, and then the more specific *front-sail, mainsail, back sail*, and *spinnaker*. So the target word *jib* was clearly linked by adults to other words for sails. In addition, adults introduced a variety of other terms associated with sailing, boats, and water, in reference to this page in the book (Clark, 2010; see also Table 12.2).

In (13), the parent of a 4-year-old child introduces a target verb, *carve*, contrasting it directly with the more general verb *cut*:

(13) Parent to child (4;9.0), picture of a hand carving a piece of wood:
This must be some wood [points], and he's using that [points to tool], to CARVE it. So that's another word for cut.
Child: *Cut (the) wood!*
Parent: Cut wood.
Child: *Yeah.*
Parent: CARVING is cutting wood.

The commonest near neighbors offered with *carve* were the verbs *cut* and *shape*. Many parents also talked about instruments used for this activity—*tool, knife*, and *chisel*—as well as the material affected, namely *wood*.

When parents offered unfamiliar adjectives, they tended to select as near neighbors terms for other entities to which the adjective could also apply. This can be seen in (14) where one parent introduces the adjective *steep*.

(14) Parent of child (2;8.11), picture of someone walking up steep slope:
They're–, they're– he's climbing up a hill. It's a STEEP hill because it goes . . . high! It goes high. It goes really high really fast. So it's STEEP. Can you say STEEP? STEEP.
Child: *Steep!*
Parent: STEEP. Do you like STEEP hills? Do you remember– Do you know we walked up some STEEP hills this morning on our walk– when we were looking for rocks?

The commonest near neighbors here were the verb *climb*, along with the nouns *mountain, hill*, and *slope*, and the adjective *flat*. Parents often alternated between *steep* and *flat* when talking about surfaces where one could walk. In (14), the parent also appeals to an episode experienced by the child (*Do you remember–*) in trying to characterize the meaning of the new adjective. A number of parents did this in following up their initial introductions of new words.

With prepositions, adults generally offered near neighbors that were close in meaning to the target term, as shown in (15).

(15) a. Parent of child (2;5.13), picture of an umbrella and ball on beach:
This is a ball, BELOW. . . an umbrella. If you're under something it means you're below it.
b. Parent of girl (3;8.16), same picture: Okay so, the ball is where in–<repair> compared to the umbrella? You know?
Child: *Under*.
Parent: Under, and another word for under is . . . BELOW.

The commonest neighbors here were *under, on the ground*, and *low*. Adults tended to present much less added information here than for the referents of new nouns, verbs, and adjectives. The frequency with which adults used different terms as they talked about the domain and the relations between the new word and other terms in that domain is represented graphically in Tables 12.2 through 12.4, with the more commonly used neighbors given in larger type. The older the child, the more words parents tended to use from the pertinent domain. These words also tended to be more closely linked to the conventional meaning of the target word. This should not be surprising since older children generally knew at least some of the relevant words already. Parents could therefore draw on further expressions and terms that they would be less likely to use with 2- or 3-year-old children (Clark, 2010).

TABLE 12.2. Jib (Noun)

2s	3s
• SAILBOAT, BOAT, FRONT, SAIL • SHIP, SAILING, MAINSAIL, BACK, WATER, CAPTAIN HOOK	• SAILBOAT, FRONT, SAIL • BOAT, WIND, BACK-SAIL • TRIANGLE, SAILING, SHIP, MERMAID • BACK, MAINSAIL, CATCH, BOW, STERN, RUDDER, SHAPE, STRONG, MOTOR
4s	**5s**
• SAIL, FRONT, SAILBOAT • BOAT • MAINSAIL, SAILING, STEER, POLE • RUDDER, BACK, HULL, BOTTOM, SQUARE, WIND, BLOW, WATER, GAFF-RIG	• FRONT, SAIL • SAILBOAT • BOAT, MAINSAIL, SPINNAKER, ROPE, SAILING • BACK, WIND, BLOW, LET OUT, SHEET, RATCHET, MAIN, POOPDECK, NAUTICAL

Table 12.2 summarizes the frequencies for common terms and expressions associated with the noun *jib*, Table 12.3 for those associated with the verb *carve*, and Table 12.4 for those associated with the adjective *steep*. The type size on each line in the Tables offers an approximation of the relative frequencies with which the parents for that age group used the associated term. Larger type indicates that more of the parents used the word or expression, with fewer parental uses shown in successively smaller type sizes. The associated terms are divided into four or five tiers, depending on the frequency with which parents produced them. The patterns shown here are representative of the terms from each word class as a whole. The noun *jib* (Table 12.2) elicited general sailing terms, with *boat* the term that was used most often, then terms for other kinds of sail, all local contrasts for *jib*, then terms for other boat parts (*rudder*, *bow*), and then other factors involved in sailing (*wind*, *gaff-rig*, *motor*). Notice that most of these contrasting terms are nouns.

In Table 12.3, the verb *carve* presents a similar picture, except that several of the closely associated terms consist of near synonyms, for instance, several other verbs that overlap with *carve* in meaning—such as *cut*, *sculpt*, *scoop*, *shave off*, and *smooth out*. The other terms used are mainly terms for the tools used in carving and for the material that is carved.

Table 12.4 lists the terms frequently offered with the adjective *steep*. These were fairly

TABLE 12.3. Carve (Verb)

2s	3s
• WOOD • CUT, SCULPTURE • TOOL, KNIFE, CHISEL, PLADOH, CLAY • WOOD-CARVER, SAND, EDGE, SHAPE, STATUE, STONE	• WOOD • TOOL, CLAY • PLADOH, KNIFE, CUT • SCULPTURE, PICTURE, SHAPE, CARVINGS, DESIGNS, SCULPT, SCOOP, INSTRUMENT, TURKEY
4s	**5s**
• WOOD, CLAY, TOOL • PLADOH, CUT, INSTRUMENT • SHAPE, CHIP OUT/OFF, STATUE, ROCK • DECORATION, STONE, SAND, SIDE, EDGE, SCOOP OUT, SMOOTH OUT, SHOVEL, CHISEL, CARPENTRY, POTTERY, ROCK	• WOOD • TOOL, KNIFE, CHISEL, STATUE, CUT, SHAVE OFF, SHAPE, TURKEY, PUMPKIN • SCRAPE, LINES, SCULPTURE, LOAF

TABLE 12.4. Steep (Adjective)

2s	3s
• UP, MOUNTAIN, HILL • CLIMB (UP) • HIGH • FLAT, BIG, MOUNTAIN CLIMBER	• HILL, UP, CLIMB (UP) • HIGH • LADDER • MOUNTAIN, BIG, TALL, HIKE, FLAT, ROCK, STAIRS, SLIDE • ROCK CLIMBING, EASY, LOW, FALL
4s	**5s**
• UP • CLIMB, MOUNTAIN • HILL • HIGH, HARD • WALK, EASY, BIG, FLAT, DOWN TALL, HIKE, MOUNTAIN CLIMBER	• CLIMB UP, MOUNTAIN • FLAT, HARD, ANGLE • HILL, SLANTED, DEEP, INCLINE, GRADE • HIGH, SLOPE, SIDE, OCEAN, ROPE, SLIDE, UPHILL, SHARP, HANDS AND KNEES, ALL FOURS, TWO HANDS, TWO FEET

limited in number for the 2- and 3-year-old children for whom this adjective was accompanied primarily by the nouns *hill* and *mountain*, and the verb *climb* (or *climb up*). The number of neighboring adjectives (e.g., *flat, big, tall, high, low*) that parents offered increased with the age of the children, presumably reflecting the elaboration, by 4- and 5-year-olds, of the domain of dimensional and spatial adjectives (e.g., Clark, 1972; Rogers, 1978), as well as their more extensive knowledge about terms for physical properties of the landscape and activities like hiking and climbing.

In summary, when adults offer children new words directly, they typically offer information about the phonological form of the target word. (They sometimes emphasize the form in follow-up repeats, after the child first repeats it to try it out.) They offer information about the intended referent (whatever is currently at the locus of joint attention) and hence some part of the meaning of the word on that occasion. And, in using a new word, they also present children with information about its part of speech (shown by its syntactic frame and inflections). They follow all of this up with additional information in the form of a selection of adjacent terms—neighbors from the relevant semantic domain that help delimit the probable meaning of the new term, and with identification of other relevant domains through uses of further relevant terms in their added general talk about the topic currently "on stage."

In summary, the initial complex of phonological, lexical, and syntactic information about a new word helps children set up a preliminary entry in memory. Any added information about the referent allows them, in addition, to link the new word to whatever they already know about the pertinent semantic domain.

Learning Conventions Takes Time

Learning a first meaning for an unfamiliar word may take only a little time, and from that moment on, children can use the word. If they do so in contexts similar to the one in which they first heard it, they will often appear to do so appropriately; in other contexts, it may be more apparent that they do not yet know the full conventional meaning. In fact, the process of acquiring terms in any one domain can last many years (see, e.g., Gropen, Pinker, Hollander, & Goldberg, 1991; Mazurkewicz & White, 1984). In one study of terms for a variety of different types of cup and glass, Andersen (1975) found that children's usage did not fully match adult usage until the children were between 9 and 12 years old. In another study of the acquisition of the names of common household objects, Ameel, Malt, and Storms (2008) found that once children aged 5 to 14 years had picked up the relevant terms, they spent several years organizing them before their usage matched that of adult speakers. Acquisition of adult-like usage is, in part, a matter of learning to attend to the appropriate features in distinguishing among the meanings of neighboring terms. Adult-like acquisition also depends on learning any idiosyncratic mappings specific to the language. Although instances identified as core exemplars are often the same for the speakers of different languages, the extensions of the relevant terms vary considerably both within and across languages (see also Berlin & Kay, 1969; Kronenfeld, 1996).

This is perhaps most easily illustrated with the case of color terms: Children have to fix the reference of each term, so they can appropriately apply *green*, say, to a range of different greens differing in hue, saturation, and brightness. The same holds for *red, blue, yellow*, and other basic color terms. But children also have to learn that there are conventional ways of labeling the colors for certain things. For speakers of English, grapes are *red* (but can range from a dark purple to a palish red-brown) or *white* (generally a pale green); lentils are *green* (actually a darkish green-brown), *red* (a pale carrot-like orange), or *brown* (a light brown and larger than "green" lentils); cabbage is *green* (pale green with white stems) or *red* (purplish-red with white stems), and so on. Such ascriptions of color may bear relatively little relation to the actual colors. In fact, the conventions for talking about the colors of food and drink vary from one language to the next, with white wines, for example, identified as white, green, or gray, depending on the language (Clark, 2006). In addition, the specific color term used on any one occasion may depend on the contrasts present in context. As Kronenfeld (1996) pointed out, a red shirt will be called *red* if it is next to another shirt that is closer than the first to orange, but it will be called *orange* if it is contrasted with other shirts that are a much deeper shade of red. Usage depends in part on how speakers use context to make references that will be successful for their addressees.

Do children have to learn the full conventional meaning of a term in order to use it for communicating with others? Adults offer them many different words and expressions along with clues in context to their conventional meanings. But working out all the details of a conventional meaning can take a very long time. The initial fast mapping children do in assigning some preliminary meaning is only the first step. And, for some domains, even adults appear quite content to learn only some minimal amount about the meaning of a term, and yet use it anyway. They may, for example, be familiar with terms for trees like *ash, poplar, hornbeam,* and *elm,* yet know nothing more about the meanings of these words, much less how to recognize actual instances of such trees or what distinguishes one from another. Much the same applies for any domain in which speakers are not experts—terms for birds, baskets, or types of lace or for activities such as bookbinding, cooking, or blacksmithing. Speakers may know the forms of some words and know the domain they belong to, but nothing more. For most purposes, knowing that the term *hornbeam* refers to some kind of tree is all is needed. But if you work with trees, you will need to know the full conventional meanings of all the relevant terminology (Clark, 2005b). In many domains, speakers learn as much as they need of the conventional meanings for everyday use. For these terms, speakers need to become conversant with the full conventional meaning, but for terms in more specialized domains, speakers may defer learning many of the details and still manage quite adequately with only partial meanings in place.

Since, in many domains, people never learn the full meanings of many of the terms available in the community, the course of acquisition may not just be protracted, it may never be completed. We rarely notice this because there is generally enough overlap in the meanings known to the conversational participants for them to communicate quite effectively (Clark, 2005b). For instance, knowing that the terms *gingko* and *rowan* both designate trees, or that *wagtail* and *thrasher* both designate birds, may be all that nonexperts know, or need to know, about these terms. That appears to be enough, on most occasions, to get by with. Such a level of partial knowledge in adults is analogous to the knowledge of young children who have just assigned a preliminary meaning to some word and have yet to learn all the other details included in the adult meaning. The overlap between preliminary and full meaning in context is often enough for the child to assign some interpretation of the word in context, and to produce it with just its partial meaning. For child and adult alike, the partial meaning at that moment *is* the meaning (Clark, 2005b; see also Wolff, Medin, & Pankratz, 1999).

Transmitting Conventions

Children are born into speech communities that use a specific language or languages. And these languages are the ones children first hear. Adults talk to young children and so expose them to the language of the community. In doing this, they display the conventions that govern their uses of words and other forms of the language. They present children with the conventional terms for designating objects, actions, relations, and properties, and they combine such terms, again following the conventions of the language, in constructions used to refer to the events around them. Adults make sure they have established joint attention when they talk to young children (e.g., Estigarribia & Clark, 2007), and, not surprisingly, some 90% of their talk to young 1-year-olds focuses on physical copresence, the here and now (e.g., Veneziano 2002). This is what is accessible in joint attention.

Adults' direct offers of conventional words, along with the many others they offer less directly, form the linguistic "overlay" available in each language for speakers to use in talking to others about their representations of events. This overlay is made up of the words and grammatical relations in each language, and it is presented to children from the start, in each instance, as *the way to talk about X*. In the case of bilingual children, of course, they are offered two distinct overlays, with distinct conventions associated with the use of each one. In many bilingual families, for example, parents will guide very young children to use one language with one parent and the other language with the other. In the case of standard versus nonstandard dialects, children are again given guidance on who to use each dialect with on each occasion. And in the case of differences in vocabulary that accompany register choice, adult usage indicates that certain forms are informal while others are formal, for instance, or that some forms mark familiarity or intimacy, while others mark social distance, and so on. In every case, children have to learn the conventions for forms and their customary meanings, so that they in turn can use these forms appropriately. Their major source for this information is the patterns of use they observe in the speech addressed to them.

New speakers—mainly children—need to learn the conventional meanings of everyday words, expressions, and constructions: Without them, they are unable to use language to convey their intentions. It is this linguistic knowledge that, transmitted from one generation to the next, keeps languages viable for communication. For young children, adults are the initial source of this knowledge. They expose children to it as they talk to them, thereby showing them how to use terms in context, which other terms each one cooccurs with, and the constructions each term can appear in.

Children acquire the conventions for their language—how to use the forms and their meanings—from the people who talk to them. As they themselves become more adept at using language, they can also attend more to other speech, not necessarily addressed to them, and take in additional information from usage there as well. As they learn the conventions for the language they are growing up with, they must accumulate the appropriate forms to use, along with the conventional meanings of those forms, in that community.

CONCLUSIONS

What conclusions can we draw from these findings? What do they imply about speakers' reliance on conventions in language and about their transmission to children? To what extent does child-directed speech influence children's acquisition of lexical structure—whether words and the semantic relations that link them or the details of the semantic packaging favored in the language being acquired? And do these findings have any implications for how children learn their first language?

Conventions and Communication

Learning the conventions of a language is crucial for its effective use for communication. And children acquiring a language are learning to communicate with language. They attend to child-directed speech, typically tailored to their level of comprehension (Snow & Ferguson, 1977). And they attend in settings where they share the locus of attention with the adult who's speaking. Joint attention allows them to focus on the referents of new terms introduced by the adult speaker, and to make pertinent inferences about possible meanings in context (Clark, 2002).

Joint attention, in fact, is a prerequisite in conversation. It licenses both the establishment and the accumulation of common ground for the adult and child speaker. New words, offered in context, and taken up by the child, are thereby added to common ground (Clark, 2007a). The same holds for the uptake of constructions, which are often associated first with specific lexical items. Adult usage is therefore a crucial ingredient in the process of acquisition as adult speakers introduce words and constructions in appropriate contexts.

Exposure and frequency both play a role here. Children are exposed to language in that adults talk to them directly, to varying degrees (Hart & Risley, 1995). This speech, as well as speech children overhear as they get older and can make some sense of it, provides the primary material from which children construct the language for themselves. Moreover, when they make errors, adults typically reformulate what the children appeared to be trying to say, but in conventional form (Chouinard & Clark, 2003). These reformulations provide children with additional evidence about the conventions of the language since they are reproduced immediately after and in contrast to children's erroneous utterances for the same intentions.

Children soon learn the commonest patterns exemplified in adult usage. They learn first those terms and constructions they are exposed to most frequently (e.g., Cameron-Falkner, Lieven, & Tomasello, 2003; Estigarribia, 2006). In acquiring verbs, for instance, they first learn the constructions that cooccur with each verb most often in parental speech, and display individual differences where the adults they hear most differ in their usage (e.g., Chenu & Jisa, 2006; de Villiers, 1985). Similarly, with word formation, they pick up first on those forms that are the most productive and hence most frequent in adult speech (e.g., Clark & Berman, 1984). The language children hear, therefore, displays the conventions of that language in that language community. And these are the conventions children must adopt to communicate effectively.

Lexical Structure and Semantic Packaging

Conventional adult usage of the lexicon necessarily exposes children to the lexical structure of their language and to the semantic relations that link the terms within a domain (e.g., Lyons, 1977; Cruse, 1986; Levin, 1993; Berlin, 1992; Murphy, 2003). From the start, they hear objects, actions, and relations picked out with terms that overlap in meaning (e.g., *sail, jib*; *cut, carve; under, below*), sometimes at several levels (*animal, dog, spaniel*; *plant, flower, tulip*), and sometimes from alternative perspectives (*dog, pet; buy, sell; attack, defend*; *waste, recycling*). Lexical structure comprises a variety of relations among word meanings within particular domains, and speakers exploit this in their lexical usage to reflect the perspectives they choose for representing an object or event to someone else (Clark, 1997).

Children as young as 2-and-a-half years come to understand and make use of a relation such as inclusion in noun hierarchies, and, for familiar domains, readily accept that the same entity may be identified as an *animal* on some occasions and a *sheep* on others (e.g., Clark, 1997). And they recognize that inclusion can mark certain differences in perspective. Children make use of this in coining new terms for subtypes as in *fire-dog* (type of dog found at the site of a fire) or *spear-page* (a page with a picture of people holding spears) as young as age 2 (Clark, Gelman, & Lane, 1985). In summary, built-in constraints on the kinds of meanings children attribute to new terms do not in fact limit children's uses of reference to apply just one term to one referent type (e.g., Markman, 1989). At the age of 2 they readily accept and use more than one term for the same referent type in many settings, both in elicitation experiments and in their spontaneous speech (e.g., Clark & Svaib, 1997; Clark & Grossman, 1998; see also Waxman & Hatch, 1992; Deák & Maratsos, 1998). And they do this because they can both take and present different perspectives on the same entity using the lexical resources at their disposal.

When children start to set up semantic domains, they are presumably aided by the fact that adults often make use of contrasting terms in context. For instance, in describing the dimensions of a toy truck, they will talk not only about its *height*, but also about its *length* and *width*. They also make functional comparisons, pointing out, for example, that it is *too high* to go under this bridge, but *low enough* (*not too high*) for that one; that it is *too wide* to fit through that gap, but can easily go down this way, and so on (Rogers, 1978; Ebeling & Gelman, 1994; see also Shipley & Kuhn, 1983). By supplying local contrasts as well as adjacent terms, adults help children group related terms together as they build up semantic domains and also set up the relations that link those terms in different ways (Clark, 1995).

Languages differ in how meanings are "packaged." In some, information about singular versus plural, for example, is conveyed with numeral and classifier combinations, whereas in others, number is added directly to the noun in the form of an affix for "more than one." Similarly, in some languages, information about path and direction is bundled with information about motion, in the verb, but in others, path and direction are expressed with particles like English *out* in *He ran out*. In those languages, motion is typically combined with manner, as in *run, walk, saunter, hop*, or *stride* (see Talmy, 1985). As children make inferences in context about the possible meanings of unfamiliar words, they gradually uncover the characteristic "semantic packaging" of the language spoken around them. And they do this in part from how adults use local and general contrasts in the lexicon. That is, they find out where, for example, there are systematic inclusion relations, which verbs contain information about manner versus direction, which verbs contain information about the doing versus undoing of an action, and so on, as well as how the language marks such meanings as person (first vs. second vs. third), number (singular vs. plural), or time (present vs. nonpresent), along with other grammatical distinctions (Bybee, 1985).

In child-directed speech, adults not only offer appropriate terms but also identify near neighbors of those terms, and provide many of the semantic relations that link them to each other in the pertinent domain. In the case of *jib*, for example, common local contrasts were *sail* and *mainsail*, and all the terms for sails were presented in the context of *boat, water*, and the term for the activity, *to sail* (e.g., Bowerman, 1985; Choi & Bowerman, 1991; Callanan, 1985; Rogers, 1978; Shipley & Kuhn, 1983). Adult usage, then, is the child's guide to semantic typology in terms of how the meanings in each domain are connected.

Collocations within a language—what occurs with what—are another tool for children in tracking the conventions on words in the lexicon. Children make use of the patterns of cooccurrence in adult speech as they work out the meanings of verbs, for example. They take into account which nouns for objects commonly occur with which verbs, whether these are similar to each other, and which other terms also occur with those verbs. Consider transitive uses of the verbs *to eat* and *to open*. *To eat* is consistently used with nouns for kinds of food—cheese, apple, bread, while *to open* is used with nouns for containers with lids or barriers to access: box, tin, jar, door, window. When presented with an unfamiliar term collocated with *eat*, children can infer that it probably refers to some kind of food even when there is nothing physically present to support this inference (see Goodman et al., 1998; Bowerman, 2005; Kako, 2005; Wilkins, 2002). That is, while physical copresence offers a critical basis for inferences about possible meanings, the surrounding language presents an equally important source for establishing conventional meanings.

Finally, the conceptual categories infants establish in their first year present another important tool in matching meanings to forms. In fact, these have been proposed as the initial opening wedge into language for children, and they are clearly one source of early meanings for words in context (Clark, 1983, 2005a; Slobin, 1985). But when the conventionality of language is taken into account, it is clear that any conceptual information must interact with whatever words or expressions adult speakers propose in context to designate particular kinds of objects, actions, or relations. That is, children must arrive at the conventional linguistic "overlay" for each language they learn, based on the usage they observe in adult speakers (see Gentner & Goldin-Meadow, 2003). It is only by learning the linguistic conventions of a language—conventions that govern its use—that they will arrive at how to think for speaking within their community (Slobin 1996a). And it is in learning the conventions for the vocabulary of a language that children come to master both meaning and use.

ACKNOWLEDGMENTS

Some of the research reported here was supported in part by grants from the National Science Foundation (SBR97-31781), The Spencer Foundation (199900133), and the Center for the Study of Language and Information, Stanford University. I am indebted to Caroline Bertinetti, Meghan Everett, Elizabeth Liebert, and Natalie W.-M. Wong for help with some of the data collection and analysis, and to Michelle M. Chouinard, Bruno Estigarribia, and Barbara F. Kelly for much discussion along the way.

Notes

1. Capitals are used to indicate contrastive or sentential stress on the word.
2. Child utterances are given in italics.
3. Ages are given in years;months.days (3;4.1), or just years;months (2;5).
4. The information here identifies the corpus in the CHILDES Archive, by name, file number, and line.

References

Akhtar, N. (2002). Relevance and early word learning. *Journal of Child Language* 29, 677–686.

Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory & Language, 58,* 262–285.

Andersen, E. S. (1975). Cups and glasses: Learning that boundaries are vague. *Journal of Child Language, 2,* 79–103.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development,* 58 (Serial No. 238).

Berlin, B. (1992). *Ethnobotanical classifications: Principles of categorization of plants and animals in traditional societies*. Princeton, NJ: Princeton University Press.

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

Bierwisch, M. (1967). Some semantic universals of German adjectivals. *Foundations of Language, 5,* 153–184.

Boas, F. (1938). Language. In F. Boas (Ed.),, *General anthropology* (pp. 124–145). New York: Heath.

Boster, J. S., & Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist, 91,* 132–142.

Bowerman, M. (1985). What shapes children's grammars? In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition,* Vol. 2 (pp. 1257–1319). Hillsdale, NJ: Lawrence Erlbaum.

Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Bowerman, M. (2005). Why can't you 'open' a nut or 'break' a cooked noodle? Learning covert action categories in action word meanings. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time* (pp. 209–243). Mahwah, NJ: Lawrence Erlbaum.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22,* 1482–1493.

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development, 56,* 508–523.

Callanan, M. A. (1990). Parents' descriptions of objects: Potential data for children's inferences about category principles. *Cognitive Development* 5, 101–122.

Cameron-Falkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27,* 843–873.

Casasola, M. (2002). Exploring the relationship between language-specific semantic spatial categories and infants' nonlinguistic spatial categories. In E. V. Clark (Ed.), *Space in language—Location, motion, path, and manner: Proceedings of the 31st Stanford Child Language Research Forum* (pp. 1–10). See online publications: http://csli-publications.stanford.edu.

Chenu, F., & Jisa, H. (2006). Caused motion constructions and semantic generality in early acquisition of French. In E. V. Clark & B. F. Kelly (Eds.), *Constructions in acquisition* (pp. 233–261). Stanford, CA: CSLI Publications.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition, 41,* 83–121.

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. (1999). Early sensitivity to language-specific spatial terms in English and Korean. *Cognitive Development, 14,* 241–268.

Chouinard, M. M., & Clark, E. V. (2003) Adult reformulations of child errors as negative evidence. *Journal of Child Language, 30,* 637–669.

Clark, E. V. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning & Verbal Behavior, 11,* 750–758.

Clark, E. V. (1983). Meanings and concepts. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology*, vol. 3: *Cognitive development* (pp. 787–840). New York: John Wiley.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge, UK: Cambridge University Press.

Clark, E. V. (1995) Later lexical development and word formation. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 393–412). Oxford: Blackwell.

Clark, E. V. (1997) Conceptual perspective and lexical choice in acquisition. *Cognition, 64,* 1–37.

Clark, E. V. (2002). Making use of pragmatic inferences in the acquisition of meaning. In D. Beaver, S. Kaufmann, B. Clark, & L. Casillas (Eds.), *The construction of meaning* (pp. 45–58). Stanford, CA: CSLI Publications.

Clark, E. V. (2005a). Semantic categories in acquisition. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 459–479). London: Elsevier.

Clark, E. V. (2005b). *A gradualist view of word meaning*. International Workshop on Word Meaning, Concepts, and Communication, Cumberland Lodge, Windsor Great Park, UK.

Clark, E. V. (2006). Color, reference, and expertise in language acquisition. *Journal of Experimental Child Psychology, 94,* 339–343.

Clark, E. V. (2007a). Young children's uptake of new words in conversation. *Language in Society, 36,* 157–182.

Clark, E. V. (2007b). Conventionality and contrast in language acquisition. In M. Sabbagh & C.Kalish (Eds.), *Right thinking: The development of conventionality* (New Directions in Child & Adolescent Development, Vol. 115) (pp. 11–23). San Francisco, CA: Jossey-Bass.

Clark, E. V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, 30.

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word-formation. *Language, 60,* 542–590.

Clark, E. V., & Clark, H. H. (1979). When nouns surface as verbs. *Language*, 55, 767–811.

Clark, E. V., & Estigarribia, B. (2009). Introducing new words to children with speech and gesture. Manuscript, Stanford University.

Clark, E. V., Gelman, S. A., & Lane, N. M. (1985). Noun compounds and category structure in young children. *Child Development, 56,* 84–94.

Clark, E. V., & Grossman, J. B. (1998). Pragmatic directions and children's word learning. *Journal of Child Language, 25,* 1–18.

Clark, E. V., & Svaib, T. A. (1997). Speaker perspective and reference in young children. *First Language, 17,* 57–74.

Clark, E. V., & Wong, A. D.-W. (2002). Pragmatic directions about language use: Words and word meanings. *Language in Society, 31,* 181–212.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language*. New York: Harcourt Brace Jovanovich.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge, UK: Cambridge University Press.

Deák, G. O., & Maratsos, M. P. (1998). On having complex representations of things: Preschoolers use multiple labels for objects and people. *Developmental Psychology, 34,* 224–240.

de Villiers, J. G. (1985). Learning how to use verbs: Lexical coding and the influence of input. *Journal of Child Language, 12,* 587–595.

Dixon, R. M. W. (1982). *Where have all the adjectives gone?* The Hague: Mouton.

Du Bois, J. W. (1987). The discourse basis of ergativity. *Language, 63,* 805–855.

Ebeling, K. S., & Gelman, S. A. (1994). Children's use of context in interpreting "big" and "little." *Child Development, 65,* 1178–1192.

Eckert, P. (2000). *Linguistic variation as social practice*. Oxford: Blackwell.

Estigarribia, B. (2006). *Variation in input and production of canonical and non-canonical Yes-No questions in child English.* Paper presented at the Annual Meeting of the Linguistic Society of America, Albuquerque, NM.

Estigarribia, B., & Clark, E. V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). *Variability in early communicative development. Monographs of the Society for Research in Child Development, 59* (Serial No. 242).

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64,* 637–656.

Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., & Pappas, A. (1998). *Beyond labeling: The role of maternal input in the acquisition of richly structured categories. Monographs of the Society for Research in Child Development, 63* (Serial No. 253).

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

Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development, 69,* 1330–1344.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language, 18,* 115–151.

Gumperz, J. J., & Levinson, S. C. (Eds.). (1996). *Rethinking linguistic relativity*. Cambridge, UK: Cambridge University Press.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: H. Paul Brookes.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126,* 248–277.

Joseph, B. D., & Janda, R. D. (Eds.). (2005). *The handbook of historical linguistics*. Oxford: Blackwell.

Kako, E. (2005). Information sources for noun learning. *Cognitive Science, 29,* 223–260.

Kemler Nelson, D., Chan-Egan, L., & Holt, M. (2004). When children ask 'what is it?' What do they want to know about artifacts? *Psychological Science, 15,* 384–389.

Kronenfeld, D. B. (1996). *Plastic glasses and church fathers: Semantic extension from the ethnoscience tradition*. Oxford: Oxford University Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Lewis, D. K. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.

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

Malt, B. C., & Sloman, S. A. (2004). Conversation and convention: Enduring influences on name choice for common objects. *Memory & Cognition* 32, 1346–1354.

Malt, B. C., Sloman, S. A., & Gennari, S. P. (2003). Universality and language specificity in object naming. *Journal of Memory & Learning, 49,* 20–42.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Masur, E. F. (1997). Maternal labeling of novel and familiar objects: Implications for children's development of lexical constraints. *Journal of Child Language, 24,* 427–439.

Mazurkewicz, I., & White, L. (1984). The acquisition of the dative alternation: Unlearning overgeneralizations. *Cognition, 16,* 261–283.

Murase, T., Dale, P. S., Ogura, T., Yamashita, Y., & Mahieu, A. (2005). Mother-child conversation during joint picture book reading in Japan and the USA. *First Language, 25,* 197–218.

Murphy, M. L. (2003). *Semantic relations and the lexicon*. Cambridge, UK: Cambridge University Press.

Ninio, A. (1983). Joint book reading as a multiple vocabulary acquisition device. *Developmental Psychology, 19,* 445–451.

Pulverman, R., Sootsman, J. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2002). Infants' non-linguistic processing of motion events: One-year-old English speakers are interested in manner. In E. V. Clark (Ed.), *Space in language—Location, motion, path, and manner: Proceedings of the 31st Stanford Child Language Research Forum* (pp. 11–20). See online publications: http://csli-publications.stanford.edu.

Rogers, D. (1978). Information about word-meanings in the speech of parents to young children. In R. N. Campbell & P. T. Smith (Eds.), *Recent advances in the psychology of language: Language development and mother-child interaction* (pp. 187–198). New York: Plenum.

Sadock, J. M. (1986). Some notes on noun incorporation. *Language, 62,* 19–31.

Shipley, E. F., & Kuhn, I. F. (1983). A constraint on comparisons: Equally detailed alternatives. *Journal of Experimental Child Psychology, 35,* 195–222.

Slobin, D. I. (1979). The role of language in language acquisition. Invited Address, 50th Annual Meeting of the Eastern Psychological Association, Philadelphia.

Slobin, D. I. (1985). Crosslinguistic evidence for the Language-Making Capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition*, Vol. 2 (pp. 1157–1249). Hillsdale, NJ: Lawrence Erlbaum.

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

Slobin, D. S. (1996b). Two ways to travel: Verbs of motion in English and Spanish. In M. S. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form ad meaning* (pp. 195–220). Oxford: Clarendon Press.

Slobin, D. I. , & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. In *Proceedings of the twentieth annual meeting of the Berkeley Linguistics Society* (pp. 487–505). Berkeley, CA: University of California.

Snow, C. E., & Ferguson, C. A. (1977). *Talking to children: Language input and acquisition*. Cambridge, UK: Cambridge University Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. E. Shopen (Ed.), *Language typology and syntactic description,* Vol. 3: *Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge University Press.

Veneziano, E. (2002). Displacement and informativeness in child-directed talk. *First Language, 21,* 323–356.

Waxman, S. R., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language, 19,* 153–166.

Whorf, B. L. (1956). Linguistics as an exact science. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 220–232). Cambridge, MA: MIT Press. (Original work published in 1940, *Technology Review, 43,* 61–63, 80–83.)

Wilkins, D. (2002). *On being "open": An ethnosemantic description of the English verb 'open' based on adult-child interactions*. Paper presented at the Symposium in honor of Melissa Bowerman, MPI for Psycholinguistics, Nijmegen, NL.

Wolff, P., Medin, D. L., & Pankratz, C. (1999). Evolution and devolution of folkbiological knowledge. *Cognition, 73,* 177–204.

13

LANGUAGE STRUCTURE, LEXICAL MEANING, AND COGNITION

Whorf and Vygotsky Revisited

John A. Lucy

Words move, music moves
Only in time; but that which is only living
Can only die. Words, after speech, reach
Into the silence. Only by the form, the pattern,
Can words or music reach
The stillness, as a Chinese jar still
Moves perpetually in its stillness.

—T. S. Eliot[1]

Psychologists often speak about language and meaning in terms of individual "words" that label "things" in the world. Such expressions emerge not only in casual speech, but also pervade the scholarly literature. Influential works such as Roger Brown's (1958) *Words and Things* lend support to this tendency in their titles, if not always in their detailed arguments. Speaking in this way perpetuates the illusion that language and speaking are primarily about individual words, words conceived of as sturdy self-sufficient forms with straightforward referential meanings, forms that we take out into the world to attach to various objects we encounter in our experience, much the way a hunter might take a snare out into the forest to capture rabbits (or other "natural kinds"). This view, in turn, strongly entails another, namely, that word meanings are merely derivative of experience, simply labeling and thus ultimately responding to objects, rather than having their own internal value and logic that can play a dynamic role in the constitution and conceptualization of experience. This derivative view of word meaning, now serving as a proxy for all of language, leads quite naturally to a view of human cognition in which language plays a secondary, handmaiden role at best. In this way, our own "fashion of speaking" about language in terms of "words" shapes our scientific understanding of the significance of language for thought.

Yet over a century of systematic linguistic research shows that this view of words and hence word meaning is unsustainable on several counts. First, what qualifies as a word form can be difficult to determine and is often the product of analysis rather than its starting point, not only for scientists but also for children learning language. Although this point can be difficult to appreciate for speakers of languages (such as English) that happen to have large numbers of minimal, freestanding forms, it is transparently obvious in other languages in which most forms are internally complex or bound to other forms. Second, many word forms do not actually refer to "things" (or "objects") but rather to events or properties, or even to nothing external at all, having essentially a grammatical function (e.g., English *of*). And those words that do refer to objects may do so in remarkably different ways, selecting this or that feature as the basis for reference (compare Quine, 1960). Finally, perhaps most importantly, even referential word meanings depend heavily on the place a given form occupies not only in a specific utterance but also in an overall grammatical structure of meaningful forms or categories, structures that differ dramatically across languages. Failure to recognize the structure-dependent quality of word meanings makes it difficult to understand their dynamic nature, both psychologically, as they come to play a constitutive role in cognitive development, and culturally, as they come to shape diverse conceptualizations of experienced reality.

We need, therefore, to replace this understanding of words and things with a structure-oriented understanding of lexemes and their meanings in order to grasp the full potential for language to influence thinking. The first section of this chapter articulates a view of lexemes and lexical meaning that provides a salient place for grammatical and semantic structure. Though hardly novel, this view needs firm emphasis if we are not to be misled by the casual, uninformed understandings described above. The second section links this view of lexical meaning to traditional theories about the relation of language and thought in the work of Whorf and Vygotsky. The discussion makes clear that both theorists saw language structures as the principal locus of linguistic influences on thought and argues that this provides a key connection that allows comparative and developmental theories to be joined into a unified position. The third section presents an empirical case study that compares two language groups, American English and Yucatec Maya, showing empirically how structural factors shape the referential value of individual lexical forms and how those referential values then influence cognition both developmentally and comparatively. The concluding discussion rethinks the course of language development and human development more generally in light of these language effects.

THE ROLE OF LANGUAGE STRUCTURE IN LEXICAL MEANING

Since the term *word* is typically used to refer to a unit of language whose form and meaning are given independently of grammatical analysis, the term *lexeme* will be used here instead precisely to highlight the dependence of such units on a structural analysis. Other terms are often used to similar purpose (e.g., *lexical item, lexical unit*) and there is no intent here to advance or defend a particular view of lexemes within the array of competing theories in linguistics. All that is essential in the present context is that lexemes be understood as abstract units of language emerging from structural analysis. They are composed of one or more morphemes (minimal units of semantic meaning) that express a meaning and a morphosyntactic category relevant to predication.[2] Language structure shapes each lexeme in two distinct, but interrelated ways. First, most *lexemes* can be identified as *forms* only with reference to an analysis of the overall language structure, rather than solely by reference to their own independent phonological and referential properties. Second, the substantive *meaning* of each lexeme then depends intimately on its structural–functional role in the language. Both of these issues have been widely ignored or deliberately evaded in traditional approaches to understanding words, approaches that focus on the phonological and referential values of words at the expense of the structural values of lexemes (Lyons, 1968, Ch. 5).

Lexical Form

A lexeme is, first of all, a constituent of the language, a lexical form that serves as an identifiable functional unit within the larger structural whole. Despite this, many psychologists treat lexemes as if they were *structurally autonomous words*, that is, as forms whose phonological and referential properties can be defined and aligned independently of structural considerations to produce minimal, freestanding referential units. Although some lexemes may have such properties, this is not generally the case. Often the relevant properties do not apply, or they do not align as expected, or they apply and align only when supplemented by other structural information. In such instances, we are forced to recognize that the identification of most lexemes is the product of grammatical analysis, not independent of it.

And this raises serious questions about accounts of language acquisition built on regarding words as freestanding primitive building blocks.

Phonological features alone cannot be used to identify the lexical forms in a language. A phonological word is a segment of speech that constitutes a single unit of pronunciation. The phonological criteria for a word vary from language to language depending on the sound patterns characteristic of it. In English, a word typically contains one dominant stress; in other languages, such as Turkish, processes such as vowel harmony play a role in defining a word (Lyons, 1968, p. 205; Trask, 2004, p. 2). But whatever the criteria, such a phonological unit need not align exactly with the meaningful lexical forms in a language as defined on other grounds. For example, the English utterance "You're here" contains two phonological words: *you're* composed of two lexemes (*you* and *be* [~*are*]) and *here* composed of one lexeme. In Yucatec Maya, the comparable utterance "*wayaneche'* " places the entire sentence into a single phonological word, composed of four morphemes (*way-yàan-ech-e'* 'here-exist-you-deictic'). Conversely, in the English utterance "Did you guys really break up?" the verb *break up* serves as a single lexeme, even though it is composed of two phonological words. To complicate things, the two phonological words (*break* and *up*) can serve in other contexts as separate lexemes in their own right and, as a noun, the combination *break up* forms a single phonological word! Likewise, in our Yucatec example, the morphemes *way . . . -e'* function as a single lexical form, which sometimes occurs as such (e.g., *waye'* 'here'), but where the *way* sometimes stands alone as a phonological word and sometimes binds to other intervening material, but always requires the *-e'* or a similar form to complete it further along in the utterance. More generally in Yucatec, as well as other languages such as Nootka (Whorf, 1941/1956) and Yup'ik (Trask, 2004), whole sentences appear as single phonological words. In short, there is no reliable relationship here between a phonological word and a lexeme in the language. Or, as Trask (2004, p. 2) puts it, "Phonological words are important in the study of pronunciation, but they are irrelevant to the study of grammar." Phonological criteria alone provide a poor guide to the meaningful lexical forms in a language, and theories of acquisition based on such an assumption are not generally viable. Indeed, the ability to segment words as a metalinguistic act comes fairly late in development, usually as part of literacy skills (Bialystok, 1986).

Nor can denotational referential values alone be used to identify the lexical forms in a language. The traditional view is that a word denotes an object (or type of object) in the world. But there is great variation in which objects show up as lexemes in languages. Where one language may have a single lexeme, another will require a construction, and vice versa. For example, Yucatec has a lexeme *chúuh* to designate a type of gourd with a figure-eight shape used to carry liquids. English, by contrast, uses a construction describing the functional shape and material, *bottle gourd*. Spanish more often simply uses the term for gourd, *calabaza*, perhaps joining it if necessary with an indication of the typical function (of carrying wine), *calabaza vinatera*. Conversely, English has a single word *boy* designating a male child. In Yucatec, the equivalent expression would be *xí'pal*, a compound form explicitly indicating male child (*xìib'-pàal* 'male-child'). There is an equivalent form in Spanish, *muchacho*, but notice that it is composed of a stem modified with a gender affix, neither of which can stand alone (*muchacho* 'boy, child' vs. *muchacha* 'girl'). So an object designated by a single form in one language may require two or more forms to be joined in another language; these forms in turn may be independent or dependent. The same argument holds when we extend the analysis beyond objects, to semantic features. For example, whereas the feature of 'masculine' in the previous example is designated by a separate morpheme in Yucatec, it appears as a dependent suffix in the Spanish, and is fused into the lexeme in English. And to complicate matters still further, each of these languages is capable, in other contexts, of using each of these three techniques to indicate gender.[3] Finally, some lexemes simply do not designate any single object or referential feature. For example, English lexical forms such as *of*, *if*, and *respectively* do not refer to experience so much as they tell us how to interpret relationships among elements in a construction; likewise, the Yucatec "prepositional" form *ti'* 'to, at, in, on, by, for, from, [etc.]' covers such a range of meanings that it approaches purely relational status, deriving its specific meaning from context. Or, as Lyons (1968, p. 200) puts it, "We must conclude that semantic considerations [alone] are irrelevant in the definition of the word [unit]. . . ." In sum, denotational reference alone is a poor guide to the meaningful lexical forms in a language, and we should be wary of theories of acquisition that presume some natural one-to-one correspondence across languages.

Even when phonological form and referential meaning combine, there are other ways in which the traditional notion of a word creates difficulties for analysis. Should we regard the nouns *boy* and *boys* as one word or as two? Should we regard the verb forms *stink*, *stinks*, *stank*, *stunk*, *stinking* as one word or five? And should we recognize the relationship among *be*, *is*, and *are*, or among *go*, *went*, and *gone*? To see these as single lexical forms in the language, we have to identify an underlying commonality of meaning across variable sound forms and of sound forms across shifts in referential meaning. Such an analysis requires an appeal to structural patterns at every step. Note, too, that in those languages without neutral, uninflected base forms, one or another form will have to be regarded as the citation form for the lexeme. For example, in Spanish, the infinitive (e.g., *hablar* 'to speak') has traditionally been selected as the citation form from an array of some 46 inflected forms for the verb. For the native speaker, most especially new learners such as children, this very structural dependence and absence of an obvious base form may make it difficult to consciously recognize and discuss lexical forms, even though they are used fluently in speech.

In short, the traditional notion of a word, a phonologically freestanding base form with clear reference to objects in the world, even when enriched to include characteristic conjunctions of form and meaning, fails to recognize the tacit structural analysis that the speakers must conduct in order to recognize and deploy the lexemes in their language. To capture these phenomena, we need to distinguish between phonological words and referential values on the one hand, and *lexemes* as functional units in the language warranted by structural analysis on the other hand. It is true that phonological and referential values are essential to the language learner in inducing the lexemes in a language, but they are not sufficient. Structural analysis and, ultimately, structural meanings must also be invoked at each stage of the induction process.

Lexical Meaning

A lexeme, like any other form in language, conveys a variety of structural meanings each time it is used. If it has denotational referential value, that is, a regular correspondence with some element of experience, then of course it signals that meaning. The focus and scope of such denotational meanings may vary across languages.[4] But there will always be language internal structural contributions to the meaning as well. Structural meaning of this sort arises both *paradigmatically*, from its selection from an array of available forms with contrasting values, and *syntagmatically*, from its placement with other forms in utterances. This selection-and-placement, this structural aspect, complete and transform the referential meaning of every individual form in an utterance. We can say then that lexical meaning has two faces: one to the outer environment of the experienced world and the other to the inner environment of the code structure itself. We are concerned in what follows with the contribution of the latter to overall meaning.

Each lexeme occupies a place in a *paradigm* of similar and contrasting forms such that part of its meaning comes from its position in the paradigm. For example, English has three lexemes for siblings, a general term *sibling* and two more specific terms contrasting in gender: *brother* 'male sibling' and *sister* 'female sibling'. Yucatec has four lexemes covering the same referential range: *láak'* 'other, sibling', *'íits'in* 'younger sibling', *sukú'un* 'older male sibling', and *kìik* 'older female sibling'. Note that when we refer to someone with the English term *brother*, this conveys that the person is a sibling and, because it was chosen instead of *sister*, it also signals male gender information; but it does not convey information about age. In contrast, when we refer to the same person with the Yucatec term *sukú'un*, this too conveys that it is a sibling and, because it was chosen instead of *'íits'in*, it signals age information—that this is an older sibling. Simultaneously, because it was chosen instead of *kìik*, it also signals gender information within the set of older siblings. Figure 13.1 displays the paradigmatic contrasts of age and gender within the set of sibling terms. As should be clear, I cannot directly translate the English term *brother* into Yucatec without first considering the relative age of the sibling, ignoring gender if he or she is younger and obligatorily marking it if he or she is older. But at least in this case, going in either direction, we can clarify the sense with an adjective signifying gender or age.

The situation can quickly become more complicated, however, when there is no easy combinatorial equivalent. For example, the Yucatec word *pàach* can refer to a person's back, the back of a hand, etc. and so is often glossed as 'back'. But it can also refer to the skin of a fruit, the shell of a nut, the bark of a tree, or the hide of an animal—where there is clearly no contrasting notion of 'front' in play as there is in our English notion of 'back'. Furthermore, *pàach* can even be used to refer to the outside portions of a house (and adjacent portions of the yard) away from the two entrances of the oval structure, as shown in Figure 13.2. From an English vantage, the term *pàach* joins our meaning of 'back' (as opposed to 'front'), with other meanings such as 'outside' (as opposed to 'inside') and 'unimportant part' (as opposed to 'important part'). Although there are clear cases of denotational referential overlap in which *back* and *pàach* can both refer to "the same thing," they are in fact quite different in their overall structural potential and hence not equivalent in meaning. And this then is the more general lesson: Referential overlap does not ensure structural (or semantic) equivalence in meaning if the overlapping forms come from distinct paradigms of forms.

Each lexeme can also enter into an array of *syntagmatic* or morphosyntactic relations with other elements such that part of its meaning comes from this placement. For example, the English lexeme *pig* can be an agent ("The pig bit the dog") or a patient

<table>
<tr><th colspan="2">English</th><th colspan="2">Yucatec</th></tr>
<tr><td rowspan="2">brother</td><td rowspan="2">sister</td><td>sukú'un</td><td>kìik</td></tr>
<tr><td colspan="2">'íits'in</td></tr>
</table>

FIGURE 13.1. Comparison of English and Yucatec Maya kin terms for siblings.

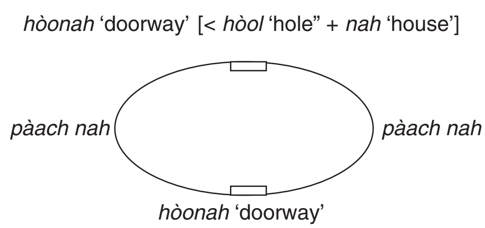


FIGURE 13.2. Range of application of the Yucatec Maya term *pàach* in relation to the regions of the traditional oval Mayan house or *nah*.

("The dog bit the pig") depending on its placement in a sentence. The capacity to enter into both roles is part of the meaning of *pig* (as opposed to *apple* or *stone*) and the placement activates the relevant role—so the meaning lies in a lexical–grammatical interaction within an overall construction (see Goldberg, 2003; Lucy, 1994, 2000). In much the same way, the English lexeme *up* has a variety of spatial meanings ("John went up to his room") and a variety of temporal (aspectual) meanings ("John and Sarah broke up") each evoked by placement with other elements; in some cases, both readings are possible ("When Sarah came in, John stood up"). Sometimes the syntactic possibilities are predictable from referential features (so pigs can bite because they are alive and have mouths with teeth), but this is not always the case. Note that an English speaker has no trouble saying "We had chicken for dinner," but it is odd to say "We had pig for dinner." It is difficult to make sense of this on referential grounds. What matters is that there is another lexeme *pork*, which substitutes paradigmatically in the latter case. But sometimes, even when there is no lexical competitor, a syntactic possibility can be blocked. For example, within English the terms *red* and *green* both denote colors and we can say that objects can "turn red" or "turn green"; but though we can say those objects *redden* we cannot say that they *greenen*. So the capacity to take the inchoative suffix *-en* is part of the structural potential of *red* but not of *green*, which means we must resort to a phrase such as *turned green* to convey a similar meaning. Furthermore, what is referentially predictable also differs across languages. In Yucatec, the word *k'éek'en* 'pig' can be used equally well as a singular or a plural. In fact, *k'éek'en* can be used equally well to describe a live pig, a dead pig, or a piece of pork, hence the ambiguity in the expression *túnkonik k'éek'en*, which can mean 'he is selling pig/pigs/pork'. So the need to distinguish between 'pig' and 'pigs' or between 'pig' and 'pork' is part of the structural meaning of English *pig* in a way it that it is not part of the structural meaning of Yucatec *k'éek'en* (Lucy, 1992a). The structural contrasts among verbal predicates are even more marked and diverse (see Levin, 1993, for examples). In short, we have to consider the syntactic potential of a lexeme as part of its meaning and to consider its specific meaning in use as a product of that potential interacting with a local syntactic arrangement.[5]

In general then, a lexeme is a functional unit within the overall language structure. A structural analysis is required to demarcate lexical forms, using the available phonological and referential patterns, and each such unit naturally incorporates aspects of meaning based on its place in the overall structure. These structural meanings include not only the conventional referential scope of the item, but also the array of forms with which a lexeme shares and contrasts in meaning, the array of morphosyntactic potentials that it has, and the specific meaning values evoked by specific structural placements. The lexicon is therefore deeply structure dependent. Whether and how these lexemes appear as "words" will vary by language. Traditional efforts to treat "words" as if they were structure-independent forms that correspond in a straightforward way with language-independent "things" in the world simply cannot be sustained in the face of these analyses. We must instead adopt an approach that considers structure central to language form and meaning from the outset and not as something added later. And when we look for possible effects of language on thought, it is to these same structural factors that we should turn our attention. Although each language is structured when the child encounters it, it may take years for that child to become sensitive to the full significance of that structure. And sensitivity to these structural regularities will provide the fulcrum by which language-internal forces can come, ultimately, to (re)shape the child's view of reality.

THE INFLUENCE OF LANGUAGE STRUCTURE ON THOUGHT: THEORETICAL PERSPECTIVES

Despite the important contribution of language structure to meaning, including word meaning, investigations of the influence of language on thought rarely consider structural factors. Even comparative studies exploring the impact of different languages on thought exhibit the same neglect of structure. To some extent this neglect stems from disciplinary factors—many of the psychologists and anthropologists interested in this area understand little about language structure or understand it in very partial ways, so they are not likely to consider it in their approaches (Lucy, 1992a, 1997a). But the neglect also has roots in general limits to *metalinguistic awareness* that lead people to foreground readily segmentable forms at the expense of broader structural patterns (Silverstein, 1981; Sapir, 1927/1949) and in more language-specific projections (or transfers) of our own structural patterns and understanding onto other languages; this is often called *transfer* (Jarvis & Pavlenko, 2008) and I have elsewhere called it *semantic accent* (Lucy, 2004, 2006). Because structural factors rarely enter into the linguistic analysis, they also rarely form the target of the language and thought interface in research.

Instead, current research on language and thought manages to circumvent or minimize attention to structure in several ways. The first and most basic approach of course is to focus on words in isolation, that is, without any reference to their grammatical or other systematic properties. This is especially common in studies of key words referring to culturally salient mental states, emotional complexes, or social values. The researchers act as if such terms can be described and understood without any reference to their grammatical status. A second approach involves isolating a small set of words on some notional or domain-based criterion without any effort to articulate their grammatical unity or heterogeneity. So there are studies of body parts that do not report whether all the terms are obligatorily possessed, or studies of motion verbs that fail to tell us whether they all denote activities or state changes or are divided in this respect, or studies of color terms that fail to take account of the reality that the terms in some languages fall into diverse parts of speech, some nouns, some verbs, and some adjectives, and have varying syntactic options within each of these. Other approaches do not neglect structure completely, but subordinate it to other factors and draw on it only insofar as it is convenient. So a third pattern is to attend to grammatical structure, but highly selectively. Here we see studies that extract a formal class such as "spatial" prepositions, but then ignore all the nonspatial uses of the forms as, for example, when the same prepositions also mark temporal aspect. A fourth pattern works to compare functionally similar morphological categories across languages, but fails to take account of the differing formal structural properties characterizing these categories in different languages. Here we see studies of spatial frames of reference that lump together obligatory verb inflections in one language with optional lexical modifiers in another as comparable. (In a sense this approach merges the domain centrism of the second approach with the structural selectivity of the third.) And a fifth pattern effectively ignores structure by being methodologically blind to it, that is, by tacitly importing it into the analysis. Here are the various elaborate efforts to define the lexemes of a language with sentences composed of other, purportedly more basic lexemes, but without specifying the syntax being used to construct those definitional sentences. And here too are all the efforts to interpret the forms of other languages in terms of our own, as, for example, when we project the individuating properties of our "count nouns" onto lexemes in a language in which there is no warrant for doing so (Lucy, 2000).[6] The pattern should be clear: Contemporary approaches offer manifold ways to avoid attending to the impact of structural meanings on the relation of language and thought.

Yet there are many reasons to believe that structural meanings might offer the greatest impact on thought. Structural patterns are ubiquitous, leading speakers to monitor constantly (or habitually) certain relevant features of meaning across a wide array of forms. They are often obligatory, in which case speakers are actually compelled to attend to and signal them. And they are typically systematic, such that the various components of structural meaning are interrelated, leading to cross-cutting entailments among semantic categories that reinforce and interact with one another. Their ubiquity, obligatoriness, and systematicity thus make them good candidates to provide direction to interpretations of the world beyond language, especially under conditions of uncertainty.

In exploring the potential impact of structural meaning on thought, at least two distinct levels of the problem emerge (Lucy, 1996).[7] The first, more general level concerns the significance for human thought having any language. Here we ask when in development, in what contexts, and in what ways does thought become sensitive to or dependent on language? The second, more specific level concerns the significance for speakers having a particular language. Here we ask whether the language and the specific lexicogrammatical structures a person uses matter. Of course these two levels are intertwined: General mechanisms always operate through specific structures, and those specific structures in turn realize their effects through general mechanisms. Hence an adequate approach must address both levels in a unified way. Often research does not link these two levels, and even when it does the linkage is through "words" conceived of as culturally supplied cognitive tools without any relation to the rest of the language (e.g., Brown & Lenneberg, 1954). In light of the discussion here, we need a robust approach that takes account of structural meaning at *both* levels of analysis.

The contrast between these two levels can be seen clearly in the work of Whorf (1941/1956) and Vygotsky (1934/1987), two of the most influential theorists studying the relation between language and thought. Whorf argued that diverse languages can provide very different analogical guides to habitual thinking, but he said very little about how or why such analogical effects develop psychologically, that is, how or why cognition comes to draw on language categories at all. Vygotsky argued that language was essential to the psychological development of conceptual thought and higher levels of intellectual development, but he had little to say about the impact of cross-linguistic variation, that is, how different language categories might affect thought dependent on them.[8] A more or less similar division of labor characterizes most contemporary approaches to the relation between language and thought. Yet the two approaches complement each other in important ways and each would gain power if joined to the other. But how can this be done? The same two theorists provide an important clue *in their shared appeal to language structure as the key element of the language and thought relationship*. The effort here to join their two approaches in a unified account aims to initiate a much broader engagement between these two levels of the language and thought problem precisely by emphasizing the pivotal role of language structure.

Whorf's (1941/1956) views on the importance of structural patterns for thought are well known. He repeatedly emphasized that the overall patterns of relations were more powerful factors determining meaning and influencing thought than were individual forms. He makes this clear in his seminal "Habitual Thought" paper when, after some preliminary examples, he introduces the main argument by focusing on patterns across grammatical categories, parts of speech, and "surface" formal status:

> The linguistic material in the above examples is limited to single words, phrases, and patterns of limited range. One cannot study the behavioral compulsiveness of such material without suspecting a much more far-reaching compulsion from large-scale patterning of grammatical categories, such as plurality, gender, and similar classifications (animate, inanimate, etc.), tenses, voices, and other verb forms, classifications of the type "parts of speech," and the matter of whether a given experience is denoted by a unit morpheme, an inflected word, or a syntactical combination.[9] (1941/1956, p. 137)

In the body of the paper, he shows in detail how complex meanings emerge out of structural relations (or categories) in the grammar and how by a system of *analogical projections* and secondary reanalyses based on such relations, speakers are led to quite different habitual understandings of experience (for a full discussion, see Lucy, 1985, 1992b). In this way, human languages as particular systems of meaning can guide speakers into one direction of thought rather than another. And finally, in his concluding remarks he invokes an even more expansive notion of structure, arguing that our concepts are shaped by very large-scale structures of meaning that draw on and cross-cut multiple systems in the language:

> Concepts . . . depend upon the nature of the language or languages through the use of which they have been developed. They do not depend so much upon ANY ONE SYSTEM (e.g., tense, or nouns) within the grammar as upon the ways of analyzing and reporting experience which have become fixed in the language as an integrated "fashion of speaking" and which cut across the typical grammatical classifications, so that such a "fashion" may include lexical, morphological, syntactic, and otherwise systemically diverse means coordinated in a certain frame of consistency. (Whorf, 1941/1956, p. 158; emphasis in the original)

As should be clear, subsequent work seeking to evaluate Whorf's ideas empirically by research on individual words or small sets of words, without reference to their place in the language as a whole, simply miss the central insight of his work. In the present context, by contrast, his insight about the importance of language structure becomes our focus.

Vygotsky's (1934/1987) views on the importance of language structure in intellectual development are perhaps less widely recognized. This may be because his key unit of analysis has been translated into English as *word meaning*—which is then interpreted to have to do with free-standing "words" as discussed earlier. But word meaning for Vygotsky has an important structural element.[10] According to Vygotsky's theory, in the second year of life human intellect (as the capacity to generalize) joins with human sociality (as the capacity to communicate) in the form of word meaning conceived of as "*a unity of thinking and speech*" or, more broadly, "*a unity of generalization and social interaction*" (1934/1987, p. 49; italics in original). Crucially, in the course of subsequent development, word meaning develops, and what develops are the structural or systematic relations to other word meanings, which in turn enable true concepts and mature language.

His argument takes the following form. He begins by emphasizing that the key problem in conceptual development is the formation of a system of relations among concepts:

> We turn now to the central problem of our research, the problem of the system.
>
> There is no question that any concept is a generalization. Up to this point, however, we have been dealing with separate, isolated concepts. We must now ask *what kinds of relations there are between concepts*. How is the individual concept—this stitch that we tear away from a living integral fabric—intertwined and interwoven with the system of concepts present in the child? Only within such a system can the concept arise, live, and develop. . . . Moreover, without well-defined relationships to other concepts, the concept's existence would be impossible. In contrast to what is taught by formal logic, the essence of the concept or generalization lies not in the impoverishment but in the enrichment of the reality that it represents,[11] in the enrichment of what is given in immediate sensual perception and contemplation. However, this enrichment of the immediate perception of reality by generalization can only occur if complex connections, dependencies, and relationships are established between the objects that are represented in concepts and the rest of reality. By its very nature, each concept presupposes the presence of a certain system of concepts. Outside such a system, it cannot exist. (Vygotsky, 1934/1987, p. 224; italics in original)

He then clarifies that these systematic relationships among concepts mediate our view of the empirical world:

> Outside a system, the only possible connections between concepts are those that exist between the objects themselves, that is, empirical connections. This is the source of the dominance of the logic of action and of syncretic connections of impressions in early childhood. Within a system, relationships between concepts begin to emerge. These relationships mediate the concept's relationship to

the object through its relationship to other concepts. A different relationship between the concept and the object develops. *Supra-empirical connections between concepts become possible*. (Vygotsky, 1934/1987, p. 234; italics in original)

And finally, he concludes the discussion by linking the emergence of concepts, and the structures within which they exist, to the maturation of word meaning:

the central point—the main thought—of our entire work. . .[is] the notion that the development of the corresponding concept is not completed but only beginning at the moment a new word [form] is learned. The new word is not the culmination but the beginning of the development of a concept. The gradual, internal development of the word's meaning leads to the maturation of the word itself. Here, as everywhere, the development of the meaningful aspect of speech turns out to be the basic and decisive process in the development of the child's thinking and speech.[12] (Vygotsky, 1934/1987, p. 241)

In short, for Vygotsky the development of the structural aspect of the word (i.e., its full lexical meaning in the sense explained earlier) represents the central driving force in the relationship between language and thought. This general developmental process may, in turn, be supplemented (or pushed) by explicit training to exploit such structural relationships via institutional practices such as formal schooling. Despite his emphasis on the importance of the development of word meaning, that is, on the development of the structural aspect of meaning, most research inspired by Vygotsky continues to neglect it. In the present context, his emphasis on structure provides a crucial point of connection to more comparatively oriented theories such as Whorf's.

The two theories connect through their appeals to the role of language structure in influencing thought, thereby providing one way to synthesize research on the different levels of the language and thought problem. Together the theories suggest that language use in thought makes possible the development of higher order thinking, that is, thinking in (true) concepts; but this can happen only if the speaker commits to the specific categories of a given language, that is, to the structural relations instantiated in its morphosyntactic categories. More specifically, as children develop, they can achieve power and generality in their thought only by exploiting the structural properties of language, as described by Vygotsky, but this necessarily entails commitment to the locally available structural properties, that is, to the language-specific structural means available, with all that that entails in terms of the sorts of limits described by Whorf (Lucy, 1985, 1996; Lucy & Wertsch, 1987).

This synthesis opens up a new approach to empirical research. The approach breaks with the traditional focus on "words" in favor of a focus on lexical meaning and grammatical categories. It effectively joins the cognitive-developmental and language-comparative aspects of the problem, recognizing their mutual interdependence. And the integration of the two levels creates new methodological opportunities as well. Developmental sequences can be used to clarify causal relations that might be unclear in the correlational studies characteristic of comparative work; the existence of different developmental endpoints across languages allows new precision in the diagnosis of ongoing language and thought interactions in development (Lucy & Gaskins, 2001). The case study that follows serves to exemplify the features of this approach.

THE INFLUENCE OF LANGUAGE STRUCTURE ON THOUGHT: A CASE STUDY

The following study illustrates a structure-centered approach to the relation of language and thought by examining whether the structural differences between American English and Yucatec Maya, a language indigenous to southeastern Mexico, lead to distinctive effects on habitual cognition. The discussion first describes some salient contrasts between the two languages, then proposes possible cognitive effects, and finally assesses whether speakers exhibit the expected effects and when such effects arise in development. The study thus integrates Whorf's concerns about the effects of language differences on thought with Vygotsky's concerns about the effects of language development on thought. The focus in this article will be less on the substantive details, which have been reported elsewhere (Lucy, 1992a; Lucy & Gaskins, 2001, 2003a), and more on the intimate relation between lexical meaning and grammatical structure and on the fruitfulness of synthesizing Whorf's and Vygotsky's approaches.

Language Contrast: Number Marking Semantics

American English and Yucatec Maya differ in their number-marking patterns for nouns (Lucy, 1992b, pp. 56–83). First, the two languages differ in the way they signal plural for nouns. English speakers obligatorily signal plural for nouns semantically marked as referring to discrete objects (e.g., *car*, *chair*) but not for those marked as referring to amorphous materials (e.g., *sugar*, *dirt*). Yucatec speakers are never obliged to signal plural for any referent, although they often do mark plural for animate referents. Second, the two languages differ in the way they enumerate nouns. For English nouns marked as having semantically discrete reference, numerals directly modify their associated nouns (e.g., *one candle*, *two candles*); for nouns not so marked, an appropriate unit, or unitizer, must be specified by a form that then takes the number marking (e.g., *one clump of dirt*, *two cubes of sugar*). Yucatec requires that constructions with numerals be supplemented by a special form, also a unitizer but usually referred to in the literature as a numeral classifier, which typically provides crucial information about the shape or material properties of the referent of the noun (e.g., *'un ts'íit kib'* 'one long-thin candle', *ká'a ts'íit kib'* 'two long-thin candle').

These grammatical differences correspond to a difference in lexical semantics having to do with quantification. In essence, all nouns in Yucatec are semantically unspecified as to quantificational unit, almost as if they referred to unformed substances. So, for example, the semantic sense of the Yucatec word *kib'*, which is typically translated as 'candle' (as in the previous example), is better translated into English as 'wax' (i.e., 'one long-thin wax')—even though, when occurring alone without a numeral modifier, the word *kib'* can routinely refer to objects with the form and function that we would call candles (as well as to other wax things). Again, this pattern is general and is not confined to this specific lexeme. So, for example, as discussed, *k'éek'en* can refer equally well to 'pig', 'pigs', or 'pork' and *há'as* indifferently to 'banana plant, banana leaf, banana fruit, etc.' in both singular and plural. If the context does not make the intended reference clear, it can be specified further with the optional plural marker or by using contrasting unitizers.[13] Because *all* Yucatec lexical nouns are neutral in regard to quantification, *all* of its nouns require such obligatorily marking, in contrast to English, which requires such marking only for some of its nouns (e.g., *dirt*, *sugar*).[14] Given the quantificational neutrality of the Yucatec noun it becomes clear why it is important to specify a unit when counting: Expressions such as 'one wax' or 'one beef' do not make quantificational sense without some unitizer (e.g., 'one stick [of] wax' or 'one head [of] beef'). In contrast, those nouns in English that do include the notion of quantificational 'unit' (or 'form') as part of their basic lexical meaning can simply take the numeral directly without any unitizer (e.g., *one candle*, *one pig*, *one banana*). Furthermore, even outside of such numeral constructions, those same English nouns including quantificational unit in their basic lexical meaning typically require obligatorily plural marking for reference to multiple tokens, whereas in Yucatec such marking is always optional. These complementary patterns of plural marking and numeral modification form part of a unified number-marking pattern that is evidenced typologically across many languages (Lucy, 1992b, pp. 61–71).

The crucial point in the present context is that apparently similar forms in the two languages can have very different underlying lexical meanings. More specifically, denotational overlap, that is, having two forms that select the same object in the world in a given context, does not amount to semantic equivalence, that is, having the same lexical meaning. In context, both *kib'* and *candle* can refer to candles, but *kib'* accomplishes this by reference to the material and *candle* by reference to other characteristics such as shape, composition, and use.[15] Crucially, this meaning difference is not an isolated, idiosyncratic one, characteristic only of this translation pair, but rather forms part of a systematic pattern of lexical meaning for all nouns, which in turn forms part of an overall number-marking schema of plural and enumeration in the language.[16] The meaning of lexical nouns in Yucatec depends on their participation in the broader system of quantification and item-by-item translation in terms of referential overlap with English will likely miss it completely. Not only will the individual forms be mistranslated, but the language itself may seem unduly arbitrary and haphazard from an English point of view, lacking vital distinctions and insisting on dispensable ones.

Cognitive Hypotheses and Predictions

To assess whether traces of these contrasting verbal patterns influence speakers' cognitive activities more generally, we need first to elucidate the potential implications of these grammatical patterns for the general interpretation of experience. If we consider the denotational meaning of nouns referring to discrete concrete referents, that is, *stable objects* that maintain their physical appearance over time, then certain regularities appear from which cognitive implications can be drawn. The quantificational unit presupposed by English nouns referring to objects of this type is frequently the shape of the object, where shape here refers to a set of dimensional, boundary, integrity, and arrangement properties. Hence use of these English lexemes routinely draws attention to the shape of a referent as the basis for incorporating it under some lexical label and assigning it a number value. Yucatec nouns referring to objects of this type, lacking such a specification of quantificational unit, do not draw attention to these aspects of shape and, in fact, fairly routinely draw attention to the material composition of the referent as the basis for incorporating it under some lexical label. If these linguistic patterns translate into a general cognitive sensitivity to these properties of referents of the discrete type, then we can draw the following prediction: Yucatec speakers will attend relatively more to the material composition of stable objects (and less to their shape) than will English speakers, whereas English speakers will attend relatively less to the material composition of stable objects (and more to their shape) than will Yucatec speakers.

We can develop a second prediction about material referents. Any concrete material referent must appear at any given moment in time with some spatial configuration, that is, in some shape or arrangement. We will confine our attention here to those materials that retain their contiguity without the assistance of a container (e.g., a squeeze of toothpaste), what we can term *malleable objects*. For these referents, a temporary (or accidental) shape is available at the moment of reference, but it could be otherwise for it is highly contingent on the current state of affairs. Because both Yucatec and English nouns referring to such material referents lack a presupposed quantificational unit, their lexical semantics should ignore the temporary shape and, in fact, should routinely draw attention to the material composition of a referent as the basis for incorporating it under a lexical label. If the linguistic patterns translate into a general cognitive sensitivity to these properties of referents of the material type, then Yucatec and English speakers should not differ from each other in their cognitive responses to malleable objects.

The two sets of predictions can be brought together into a unified prediction for these two types of objects. English and Yucatec speakers should differ in their cognitive response to stable objects in line with the differences in their grammatical treatment of them, but they should agree in their response to malleable objects in line with the similarity in their grammatical treatment of them. Alternatively, looking within each language, we can predict that English will show a cognitive split vis-à-vis the two types of objects, whereas Yucatec will show cognitive continuity across them. Note that these predictions are relative rather than absolute, that is, they contrast two patterns, not absolute values. Note also that neither pattern of classification can be described as inherently superior to the other. And note finally that these predictions are based on general structural patterns in the language, not on the analysis or selection of particular individual lexemes.

Cognitive Contrast with Adults: Shape versus Material Preference

The cognitive predictions were tested with adult speakers from both languages (Lucy, 1992a, 2004; Lucy & Gaskins, 2001, 2003a). Twelve speakers in each group were shown various triads of objects. Each triad consisted of an original *pivot* object and two *alternate* objects, one of the same shape as the pivot and one of the same material as the pivot. So, for example, speakers were shown a plastic comb with a handle as a pivot and were asked whether it was more like a wooden comb with a handle or more like a plastic comb without a handle. The expectation was that English speakers would match the pivot to the other comb with a handle whereas the Yucatec speakers would match it with the other comb made of plastic.

Informants were shown one group of such triads involving stable objects that, across the stimulus set, controlled for size, color, function, wholeness, and familiarity (see Lucy & Gaskins, 2001, for examples). The predicted classification preference was strongly confirmed, with adult English speakers choosing the material alternate only 23% of the time and adult Yucatec speakers favoring it 61% of the time. That is, Yucatec speakers were over two-and-a-half times more likely to choose the materially similar alternative.

A second set of informants was shown a different group of such triads involving malleable objects, that is, foams, creams, gels, pastes, powders, particles, or granules, each formed temporarily into distinctive shapes (see Lucy & Gaskins, 2003a, for examples). In this case, as expected, both groups made a substantial number of material choices, with adult Yucatec speakers favoring material choices 53% of the time and adult English speakers favoring them 34% of the time. Although the direction of contrast is similar to that found for stable objects, the group difference narrows considerably such that the Yucatec speakers are making only about 50% more material choices than are the English speakers. The difference is no longer statistically reliable, exactly in line with the language-based prediction.

Clearly the two adult groups in this comparison do construe and classify these objects differently and in line with the expectations based on the underlying lexicogrammatical structures of the two languages. Even when English and Yucatec speakers both have a lexeme capable of designating an object in the world, the cognitive construal of that object differs as a function of the underlying structure-based meaning value of the lexemes. Furthermore, the difference in orientation does not only reside locally as a fact about each individual lexeme, but also forms part of a global structural pattern in the language, part of a coherent "fashion of speaking" that splits the lexicon in one case and does not in the other.

Cognitive Contrast with Children: Shape versus Material Preference

A second line of research explored the development of these cognitive differences in childhood (Lucy & Gaskins, 2001, 2003a). This effectively joins the Whorfian question about the impact of language differences of the sort just described with a Vygotskian one about the impact of language on the development of thought. And it provides the synergistic methodological benefits outlined earlier. On the one hand, developmental research allows us to address the question of which comes first, the language pattern or the cognitive pattern. Although many factors suggest that the language categories must be the leading force here (see Lucy, 1992a), direct developmental evidence can provide an important confirmation. On the other hand, once we have an adult contrast, we can use it to help us diagnose when and how language and thought interact in development. One of the perennial problems in studies of the development of language and thought is that the two are so entangled that it can be difficult to determine which is influencing which. But when there are distinct developmental endpoints, we can use the sequential appearance of these differences as a diagnostic tool to determine the order and nature of the developmental process. Thus, in terms of the different types of language and thought effects we described earlier, this second line of research links the structural and semiotic levels of the language and thought problem. That is, the research uses the effects of diverse languages to explore the general processes by which language and thought interact and uses the properties of the general process to help clarify the effects of language diversity.

Pilot work indicated that the distinctive cognitive patterns were appearing at around age 8 years. By age 7 years children in both groups have mastered the language patterns at issue in speech. English-speaking children have substantial command of plurals and Yucatec-speaking children have substantial command of numeral classifiers. Accordingly, we then administered the full set of triads described earlier to samples of American English and Yucatec Maya children at ages 7 and 9 years. Looking first at stable objects, we obtained the results shown in Figure 13.3. As can be seen, English-speaking and Yucatec-speaking 7-year-old children showed an identical early bias toward shape—choosing material alternates only 12% of the time. But by age 9 years the adult pattern was visible: English-speaking children continued to favor shape, choosing material alternates only 18% of the time, whereas Yucatec-speaking children were now choosing material alternates 42% of the time. Finally, on the far right, for comparative purposes, are the adult results reported earlier. Thus, the same kind of language-group difference found among adult speakers is also found in children by age 9 years—and this result is statistically reliable.

Turning next to the results for malleable objects, where we expect the two groups to look alike, we find that English-speaking and Yucatec-speaking 7-year-old children both showed a substantial number of material choices as shown in Figure 13.4. English-speaking children choose the material alternate 42% of the time and Yucatec-speaking children choose the material alternate 46% of the time. At age nine there is essentially no change: English children choose material alternates 43% of the time and Yucatec children choose them 50% of the time. And again, the adult responses are given to the right of the chart. Overall, the similarity of response found among adult speakers for referents of this type also appears in children. However, now viewed in contrast to the developmental data, we can see that the adult results appear more strongly differentiated in a manner reminiscent of the stable-object results—which perhaps suggests some general transfer of effect from the stable object category to these malleable object stimuli.

We can bring both of these results together to display the interaction of referent type and language type across age, as shown in Figure 13.5. These composite findings show that language and thought engage in new ways in middle childhood and subsequently give both specific and global shape to adult

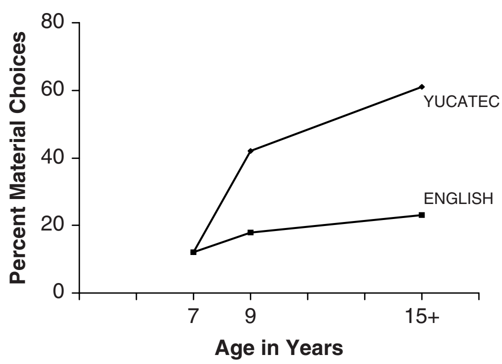


FIGURE 13.3. Developmental pattern for English and Yucatec classification preferences with stable objects: material versus shape.

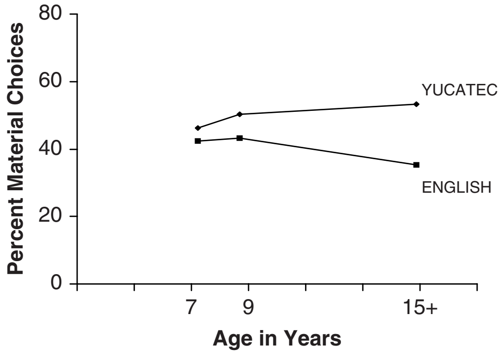


FIGURE 13.4. Developmental pattern for English and Yucatec classification preferences with malleable objects: material versus shape.

FIGURE 13.5. Developmental pattern for English and Yucatec classification preferences with both stable and malleable objects: material versus shape.

thought. Seven-year-old children show clear sensitivity to referent type independently of language group membership. They show a relative preference for material as a basis of classification with malleable objects and a relative preference for shape as a basis of classification with stable objects. Both bases of classification respond to stimulus properties and are fully available to and used by both groups. Apparently, referent type but not language type is the dominant factor in these nonverbal cognitive tasks at this age. Simply having a linguistic form in the language is not enough in itself to shape cognition. In contrast, 9-year-old children show differential sensitivity to referent type along adult lines: Their classification preferences differ where the languages differ and correspond where the languages correspond. This suggests that language categories increase in their importance for cognition between ages 7 and 9 years, and that category patterns in the linguistic structure become important in a new way. Adult responses continue to show these language-specific patterns but also trend toward consolidation into a dominant pattern for each group. The Yucatec responses converge toward material choices and the English responses toward shape choices. The split-marking pattern in English obviously militates against the complete erasure of the distinctions among referent types in that language, hence the overall trend necessarily remains subordinate to the main effect of cognition aligning with the specific linguistic treatment of a referent type. We can summarize the overall pattern of these results by saying that *young children begin by grouping different referent types in the same way and shift during middle childhood to grouping the same referent types in quite different ways as a function of the structure of their language*.[17]

Finally, the late emergence of these language effects confirms that the acquisition of individual "words" referring to the various "things" is not sufficient to create these particular cognitive effects. Children in both groups have had the vocabulary necessary to refer to these stimuli for many years, and yet at age 7 years there are no cognitive differences. But these results also make clear that simple command of the grammar in the usual sense is also not sufficient. Most of the central elements of the grammar, including number marking, have been in place for many years. Just as English-speaking children have substantial command of plurals by age 7 years, so too do Yucatec-speaking children have substantial command of numeral classifiers by this age. Seven-year-old Yucatec-speaking children reliably use classifiers when counting, draw appropriate semantic distinctions among them in comprehension tasks, and will judge a number construction lacking them as faulty. However, they do still fall short of having the full adult range of classifiers in comprehension and production. Insofar as the cognitive results derive from basic structural characteristics of the language rather than mastery of the full range of lexemes, there is no reason the effects should not appear at age 7 years rather than several years later. Something new must be happening during this middle childhood period.

THE DEVELOPMENTAL ASCENDANCE OF STRUCTURE

We have seen that the widespread tendency to think of the relation of language and thought in terms of "words" and "things" cannot be sustained. Instead, we must understand lexemes and their meaning values in terms of their place in the overall system in a language. In this sense, lexemes form an integral part of the overall meaning system of the grammar. And it is this systemic meaning that both Whorf and Vygotsky identified as the crucial aspect of language in terms of its importance for thought. The empirical case study of number marking in English and Yucatec supports their views. The contrasting systems of number marking are associated with contrasting performance on cognitive measures, in accordance with Whorf's views. These effects of language on thought are mediated not only through lexical categories: The associated inflectional pattern of plural marking has other direct effects on cognition (see Lucy 1992b, Ch. 3). Likewise, effects are not limited to simple classification tasks: Similar patterns appear in complex classification tasks (Lucy & Gaskins, 2001), in memory tasks (Lucy & Gaskins, 2003b), and in everyday behavior (Lucy 2004). And these structural effects emerge during middle childhood, in accordance with Vygotsky's views. Together, these findings suggest that the specific structure of the language that is spoken takes on new significance for cognition during this age period.

Understanding how language and thought come to relate in this new way will require taking a closer look at language development during this period. This is not a period of child language that has been heavily studied, but the available research shows that children develop many new verbal skills during this period, and most of these changes suggest that the structural element of language attains new significance as the child engages in more demanding discursive tasks. In terms of language structure, children continue their lexical development, adding new forms and reorganizing old ones so as to converge on the meanings held by adults (Ameel, Malt, & Stroms, 2008). In terms of grammatical structure, they master constructions such as passives (Chomsky, 1969) and the anaphoric use of demonstratives (Karmiloff-Smith, 1979) that enhance discourse cohesion. They also rework existing structural resources to create more coherent narratives through the sophisticated handling of temporal ordering and reported speech (Berman & Slobin, 1994; Hickmann, 1993, 2003). All of these structural developments involve taking existing structural alignments of form and meaning and either overriding or manipulating them in the service of various discursive ends. In terms of language function, children during this period also begin to use language for new forms of verbal humor and insult, as well as specialized stances such as sarcasm and flirting (e.g., Hoyle & Adger, 1998; Romaine, 1984). These skills all involve deploying one line of referential meaning while a second, sometimes diametrically opposed meaning, is also evoked in order to express a stance the speaker is taking toward the material. In formal terms, the child has learned to exploit the reflexive poetic potential of language such that one level of the message effectively "comments" on another and a new message emerges from the conjuncture. At the same time, new metalinguistic skills emerge as children become able to explain the meanings of words more effectively, setting one construction into equation with another, and as their self-corrections grow beyond a concern with referential accuracy to a concern with communicative appropriateness and rhetorical effect. In particular, the ability to recognize and appeal to a listener's presuppositions and then to manipulate their expectations and reactions suggests a growing enmeshment of language with the surrounding socially shared reality.

Collectively, these new skills reflect a growing sensitivity to and mastery of the full structural implications of language forms. Part of this mastery involves the realization that these implications are both recognized and used by others, and therefore that they can be deployed as strategic resources for achieving a variety of effects in communicative interaction. Eventually the child learns to draw on the full latent power of the shared structural means at his or her disposal. Crucially, the "problem-space" of language structure must itself be thoroughly mastered before these new functions can be erected on it (Karmiloff-Smith, 1979). The child's new capabilities are of three abstract types. First, there is deeper, more flexible mastery of the fundamentals of the meaning structure, mastery sufficient to permit the use of a single form for multiple meanings and to signal a given meaning through multiple forms. This flexibility permits greater referential precision and allows users to coordinate several messages in a single utterance, whether as speakers or hearers. Second, there is a deeper, subtler mastery of the fundamentals of the discursive space. These new discursive capacities necessarily involve shaping a message for the participants in a particular speech event. This implies an ability to understand the likely response a given utterance will elicit from a listener in a given situation and what, in turn, their interlocutors own response entails for them. In formal terms, what is emerging is the ability to coconstruct and sustain a shared reality, a common ground for the purposes of conversation. Third, both of these shifts depend on reanalyzing the deictic forms that anchor linguistic structures in on-going discourse (person, tense, modality, and evidentials). Hereafter, such deictic forms not only have reference to the default, taken-for-granted, immediate speech situation that dominates young children's speech, they also have reference to the broader shared social and interactive context, including language structures themselves, that adult speech both presupposes and helps create.

Interestingly, precisely during this period of enhanced structural mastery and associated verbal competence, children also begin to lose some of their former flexibility in language learning. Children learning language later in life will typically exhibit this loss of flexibility in the form of an *accent*, that is, a structurally driven interference with the new language. And this accent will not just be phonological but also semantic and pragmatic, as the child systematically seeks to apply the structural concepts of one or another prior language to the new language (Lucy, 2006). It is as if the child, in order to implement more sophisticated forms of discourse, is forced in some way to rigidify the existing language system in a way that interferes with learning. In other words, new verbal powers seem to be purchased at the expense of structural openness. And henceforth, each new language is "heard" through the structural paradigms of the first-learned languages.

The results of the case study reported here indicate that something similar happens with cognition during this period. For it is precisely during this period of emerging verbal skill, resting on virtuoso structural mastery and commitment to local discursive realities, that linguistic relativity effects appear. We not only see other languages through the lens of our own language in the form of a semantic accent, it seems we also come to see and think about reality itself through categories shaped by that same semantic accent. Even as the use of language structures helps liberate us from living and thinking only in the immediate reality, "enriching" our vision (per Vygotsky), its structures and their shared entailments are also becoming "habitual" constraints on our vision of reality (per Whorf).

This suggests that engagement with the inner structural logic of a language and the particular discursive world it enables provides the leverage needed to transcend the immediate moment so as to reenvision reality, to rethink it, and ultimately to remake it, precisely the practices that distinguish humans from other species. From this vantage point, linguistic relativity effects are not some unfortunate side effect of language development, but are rather its intended achievement as we recruit the inner face of our particular language structure to the shared task of reimagining the reality around us. It is crucial to see that the structural patterns in language that support this development *are* the telos of language development, the end toward which it develops. And this telos is latently there from the beginning both as a presupposition of the developmental process and as its central achievement. It is not "words" but the pattern among "words"—their structural aspect—that allows us to reach beyond the immediate speaking moment to construe an historically specific, yet enduring reality, a reality that represents the stable and enduring legacy of each language to its speakers.

Or say that the end precedes the beginning,
And the end and the beginning were always there
Before the beginning and after the end.
—T. S. Eliot

ACKNOWLEDGMENTS

An earlier version of this chapter entitled "Whorf and Vygotsky revisited: towards a comparative developmental understanding of word meaning" was presented at a National Science Foundation Conference on "Words and the World: How Words Capture Human Experience," at Lehigh University, Bethlehem, Pennsylvania, 6 June 2005. This version profited greatly from detailed comments on the first draft by volume editors Phillip Wolff and Barbara Malt and from a critical reading of the final draft by Suzanne Gaskins. The final revisions were undertaken while I was an Andrew W. Mellon Foundation supported Fellow at the center for Advanced study in the Behavioral Sciences, Stanford, CA.

Notes

1. *Four Quartets* (1943, "Burnt Norton," §V:137–143).
2. In the literature, lexemes are often designated by small caps (e.g., STINK) to differentiate them from the phonological forms in which they can appear and which are italicized (e.g., *stink*, *stank*, etc.) or from the specific referential meanings these forms take, which appear in single quotations marks (e.g., 'present tense of STINK,' 'past tense of STINK,' etc.). In this informal discussion, since it creates no confusion, the italicized form will serve to designate the lexeme. Furthermore, it is worth noting that in cases in which a morphological rather than a lexical analysis is preferred, a similar distinction needs to be drawn between morphemes, as functional units warranted by structural considerations, and morphs as their substantive local manifestations. See Lyons (1968, Ch. 5) for a discussion. Finally, it is worth emphasizing that the appearance of a lexical meaning in phonological word form, especially if recognized by native speakers as such, may become psychologically or socially significant, but it is not an essential condition.
3. So, to pursue the gender example, English also contains compound forms marking gender (e.g., *man-child, male child*) and suffixes marking gender on a base (e.g., *lioness*, *priestess*). Likewise, Spanish has some alternates (e.g., *actor* 'actor,' *actriz* 'actress') that fall outside of the *-o*/*-a* alternation and, for that matter, the same alternation can signal values other than simple gender (e.g., *manzana* 'apple,' *manzano* 'apple tree'). And Yucatec does distinguish the gender of some animals with minimal forms (e.g., *tso'* 'turkey cock,' *xtùux* 'turkey hen') and with gender markers *nohochmáak* 'old man' *xnohochmáak* 'old woman.'
4. The systematic cross-linguistic assessment of differences in the outer face, that is, denotational referential meaning, remains in its infancy. For an exemplary sustained effort at such "referential typology," see Levinson (2003) and for insight into the scope of the problem for supposedly simple concepts, see Wilkins and Hill (1995) and Lucy (1997b).
5. Although most linguists recognize that lexemes have a structural aspect, the current discussion differs from some recent accounts in two ways. First, the range of structures considered relevant to lexical meaning in this chapter is broader. Included are both the external structures of referential practice such as typical focus and range of denotation and the internal structures of semantic value such as paradigmatic alternatives and morphosyntactic potential. In contrast, some linguists restrict the use of the term *structure* or *semantic structure* to a lexeme's internal morphosyntactic potential and exclude the rest of lexical meaning as *semantic*

content. Second, these various levels of meaning are seen as routinely and necessarily in dialogue, constantly influencing each other. In contrast, it is common to assert that only the morphosyntactic potential (its semantic structure) of a lexeme interacts with or is "visible to" the rest of grammar and that the rest of lexical meaning (its semantic content) is linguistically "inert." But if the various elements of lexical meaning interact, then this separation cannot be sustained. Typological, developmental, and historical data all suggest such interactions: Changes in referential use can alter semantic value and vice versa, and within the semantic domain changes in paradigmatic alternatives can affect syntagmatic selection and vice versa. For a discussion of the narrower view of semantic structure, see Grimshaw (2005, Ch. 2); for its use in characterizing language acquisition, see Pinker (1989, especially Ch. 3).

6. The five approaches described here form an orderly set in terms of their point of departure in articulating the "metalanguage" used: (1) units in the analyst's language used without recognition of their structural properties, (2) units in the analyst's reality (domain) used without structural analysis, (3) structural analysis used, but subordinated to domain constraints (within one language), (4) structural analyses used, but subordinated to domain constraints (across languages), and (5) structures of the analyst's language used without recognition of their structural origins or implications. See Lucy (1997a) for more on the contrasting logic of structure-based and domain-based approaches.
7. Lucy (1996) articulates three levels of language and thought interaction, but the third, which has to do with institutionalized patterns of usage, will not be developed here.
8. Vygotsky did write about variation in patterns of use in his discussions of schooling. See Lucy (1996).
9. Note here that Whorf does recognize that the form-of-appearance of a meaning as a morpheme, word, or construction may also differ across languages and influence thought by virtue of its particular status. But it is just one aspect of patternment among many others.
10. For discussions of Vygotsky's use of the term *word meaning* see Wertsch (1985, Ch. 4, especially pp. 99–108) and Van der Veer and Valsiner (1991, Ch. 11). The latter clarifies that the Russian term *znachemie* would be equivalent to our notion of intensional rather than extensional meaning (1991, p. 265, n. 2).
11. This might be better worded as follows: There is a local impoverishment (the concept extracts only certain features of the object for representation) in the service of a global enrichment (the concept is inserted into a network of relations that adds features of meaning).
12. In the following line, Vygotsky reiterates the primacy of meaning (i.e., internal structural relations) over mere external form by quoting a phrase from Tolstoy. Unfortunately, the wording in this context appears to suggest that concepts precede words. For Vygotsky's critique of this view and of Tolstoy's arguments to this effect, see pp. 170–172.
13. So, we can say *k'éek'en-ó'ob'* 'pigs,' *ká'a-p'éel k'éek'en* 'two-unit pig,' *ká'a-túul k'éek'en* 'two-living pig,''*un-xéet k'éek'en* 'one-piece (of) pig,' etc., and *ká'a-ts'íit há'as* 'two-stick banana [= fruit],' *ká'a wáal há'as* 'two-flat banana [= leaf],' ká'a *kúul há'as* 'two seated banana [=plant]," etc. Note in the case of *ká'a-túul k'éek'en* 'two-living pig' that *túul* does not directly signal 'living' in a biological sense, but something broader on the lines of 'self completing or self moving'; in conjunction with *k'éek'en* 'pig/pork' or another species name it routinely yields a constructional entailment of 'animate' (Lucy 1992a, pp. 79–82, 2000). Many Yucatec unitizers generate constructional meanings of this type—a point that cannot be further developed here but that highlights in another way the structure dependence of lexical meaning.
14. English has some lexemes that can appear quantificationally neutral as to number (e.g., *sheep*), whole animal status (e.g., *lamb*), or both (e.g., *duck*). Unlike the Yucatec lexemes, however, none of these examples requires a unitizer for indefinite reference under the whole animal interpretation. See a full discussion in Lucy (1992a, Ch. 2).
15. Here, for example, is a contemporary definition of *candle* from *Webster's Seventh New Collegiate Dictionary* (1965, p. 121): "1: a long slender cylindrical mass of tallow or wax containing a loosely twisted linen or cotton wick that is burned to give light. 2: something resembling a candle in shape or use. . . ." Although the shape and structure now predominate, historically the light-giving function apparently predominated (compare the related *(in)candescent* and Latin *candere* 'shine, glow, gleam (white), etc.'

16. As should be clear from the general argument, this structure dependence is not confined to nouns, but will also be true for adjectives, verbs, and other forms (Lucy, 1994, 1997b).
17. Research on other classifier languages such as Japanese produces results similar to those found in Yucatec (see discussion in Lucy & Gaskins, 2003a).

References

Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language, 58*, 262–285.

Berman, R., & Slobin, D. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.

Bialystok, E. (1986). Children's concept of word. *Journal of Psycholinguistics Research, 15*, 13–32.

Brown, R. (1958). *Words and things*. New York: The Free Press (Macmillan).

Brown, R., & Lenneberg, E. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology*, *49*, 454–462.

Chomsky, C. (1969). *The acquisition of syntax in children from 5 to 10*. Cambridge, MA: MIT Press.

Eliot, T. S. (1943). Burnt Norton. In *Four Quartets* (pp. 13–20). New York: Harcourt, Brace, and World (Harvest).

Goldberg, A. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, *7*, 219–224.

Grimshaw, J. (2005). *Words and structure*. Stanford, CA: CSLI Publications.

Hickmann, M. (1993). The boundaries of reported speech in narrative discourse: Some developmental aspects. In J. Lucy (Ed.), *Reflexive language: Reported speech and metapragmatics* (pp. 63–90). Cambridge, UK: Cambridge University Press.

Hickmann, M. (2003). *Children's discourse: Person, space, and time across languages*. Cambridge, UK: Cambridge University Press.

Hoyle, S., & Adger, C. (Eds.). (1998). *Kids talk: Strategic language use in later childhood*. Oxford: Oxford University Press.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge: New York.

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge, UK: Cambridge University Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Levinson, S. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, UK: Cambridge University Press.

Lucy, J. (1985). Whorf's view of the Linguistic Mediation of Thought. In E. Mertz & R. Parmentier (Eds.), *Semiotic mediation: Sociocultural and psychological perspectives* (pp. 73–97). Orlando, FL: Academic Press.

Lucy, J. (1992a). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. (1992b). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. (1994). The role of semantic value in lexical comparison: Motion and position roots in Yucatec Maya. *Linguistics*, *32*, 623–656.

Lucy, J. (1996). The scope of linguistic relativity: An analysis and review of empirical research. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 37–69). Cambridge, UK: Cambridge University Press.

Lucy, J. (1997a). Linguistic relativity. *Annual Review of Anthropology*, *26*, 291–312.

Lucy, J. (1997b). The linguistics of 'color'. In C. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 320–346). Cambridge, UK: Cambridge University Press.

Lucy, J (2000). Systems of nominal classification: A concluding discussion. In G. Senft (Ed.), *Systems of nominal classification* (pp. 326–341). Cambridge, UK: Cambridge University Press.

Lucy, J. (2004). Language, culture, and mind in comparative perspective. In M. Achard & S. Kemmer (Eds.), *Language, culture, and mind* (pp. 1–21). Stanford, CA: Center for the Study of Language and Information Publications (distributed by the University of Chicago Press).

Lucy, J. (2006). *Linguistic relativity and the problem of semantic accent*. Plenary Address, 2nd Language, Culture, and Mind Conference, Paris, France.

Lucy, J., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In S. Levinson & M. Bowerman (Eds.), *Language acquisition and conceptual development* (pp. 257–283). Cambridge, UK: Cambridge University Press.

Lucy, J., & Gaskins, S. (2003a). Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner & S. Goldin-Meadow, *Language in mind: Advances in the study of language and thought* (pp. 465–492). Cambridge, MA: MIT Press.

Lucy, J., & Gaskins, S. (2003b). *Comparative developmental methods in the study of language and thought*. Formal Colloquium, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Lucy, J., & Wertsch, J. (1987). Vygotsky and Whorf: A comparative analysis. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 67–86). Cambridge, UK: Cambridge University Press.

Lyons, J. (1968). *Introduction to theoretical linguistics.* Cambridge, UK: Cambridge University Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Quine, W. (1960). *Word and object*. Cambridge, MA: Harvard University Press.

Romaine, S. (1984). *The language of children and adolescents: The acquisition of communicative competence*. New York: Blackwell.

Sapir, E. (1949). The unconscious patterning of behavior in society. In D. G. Mandelbaum (Ed.), *The selected writings of Edward Sapir in language, culture, and personality* (pp. 544–549). Berkeley: University of California Press. (Original work published in 1927.)

Silverstein, M. (1981). The limits of awareness. *Working Papers in Sociolinguistics*, no. 84. Austin: Southwestern Educational Laboratory.

Trask, L. (2004). What is a word? *University of Sussex Working Papers in Linguistics and English Language 11*.

Van der Veer, R., & Valsiner, J. (1991). *Understanding Vygotsky: A quest for synthesis*. Oxford: Blackwell.

Vygotsky, L. (1987). Thinking and speech. In R. W. Rieber & A. S. Carton (Eds.), *The collected works of L.S. Vygotsky. Vol. I: Problems of general psychology* (trans. N. Minnick) (pp. 39–285). New York: Plenum. (Original work published in 1934.)

Webster's seventh new collegiate dictionary. (1965). Springfield, MA: G. & C. Merriam.

Wertsch, J. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Cambridge University Press.

Whorf, B. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 134–159). Cambridge, MA: MIT Press. (Original work published in 1941.)

Wilkins, D., & Hill, D. (1995). When GO means COME: Questioning the basicness of basic motion verbs. *Cognitive Linguistics*, *6*, 209–259.

14 HOW WORDS CAPTURE VISUAL EXPERIENCE

The Perspective from Cognitive Neuroscience

David Kemmerer

Would it be better to sit in silence?
To think everything, to feel everything, to say nothing?
This is the way of the orange gourd.
This is the habit of the rock in the river, over which the water pours all night and all day.
But the nature of man is not the nature of silence.
Words are the thunders of the mind.
Words are the refinements of the flesh.
Words are the responses to the thousand curvaceous moments—we just manage it —sweet and electric, words flow from the brain and out the gate of the mouth.
We make books of them, out of hesitations and grammar.
We are slow, and choosy.
This is the world.

Mary Oliver (2001)

The other chapters in this volume adopt the perspectives of linguistics, psychology, and anthropology to approach the central theme of how words capture experience. Here I adopt yet another perspective—that of cognitive neuroscience—and attempt to demonstrate how recent discoveries in this rapidly growing field can help illuminate the nature of word meaning. So far the greatest progress has been made by investigations of the brain structures that underlie the linguistic encoding of the visual world, and for this reason I concentrate on this complex and intriguing set of findings. This area of inquiry is not without controversy (for reviews of several competing theoretical positions, see Caramazza & Mahon, 2006; Gainotti, 2006; Hart & Kraut, 2007b; Patterson, Nestor, & Rogers, 2007; de Vega, Glenberg, & Graesser, 2008; Mahon & Caramazza, 2008, 2009; Vigliocco, Meteyard, Andrews, & Kousta, 2009; see also the meta-analytic brain mapping study by Binder, Desai, Graves, & Conant, in press). However, there is widespread agreement about the importance of a number of studies which suggest that different visual components of word meaning—for instance, the range of colors encoded by *brown,* the type of shape encoded by *horse*, the motion pattern encoded by *gallop*, and the network of spatial relations encoded by *around*—depend on cortical regions that either overlap with or lie adjacent to some of the same regions that are engaged during the visual perception of those properties. It is this body of data that is the main focus of this chapter. I argue that these new findings, which come from both functional neuroimaging studies with normal subjects and neuropsychological studies with brain-damaged patients, support an approach to linguistic meaning, and to conceptual knowledge more generally, that is often referred to as the Embodied Cognition Framework or, as I will call it, the Simulation Framework. This approach treats semantic structures as being grounded in modality-specific sensorimotor systems, as opposed to being completely amodal in character.

The chapter is organized as follows. The first section presents the basic tenets of the Simulation Framework and summarizes three recent behavioral studies that provide evidence for it. The second section then reviews research in cognitive neuroscience that further elaborates and substantiates the Simulation Framework. It begins by describing a neurobiologically based instantiation of the Simulation Framework called Convergence Zone Theory and then surveys empirical findings from a variety of studies that address the neural correlates of the visual semantic components of words for objects, events, and spatial relations. Next, the third section covers some recent typological work on crosslinguistic variation in the lexicalization of the visual world, and points out several questions that this work poses for cognitive neuroscience. The major topics are as follows: shape classifiers and the count/mass distinction; language-specific semantic maps across the universal conceptual domain of manner-of-motion; and topological vs. projective spatial relations, the latter category being defined in terms of three different frames of reference—intrinsic, relative, and absolute. Finally, the fourth section concludes the chapter by highlighting some important general aspects of how the Simulation Framework illuminates the nature of word meaning, while at the same time acknowledging some equally important limitations of this approach.

THE SIMULATION FRAMEWORK

Many traditional approaches to the human conceptual system assume that semantic knowledge is represented separately from modality-specific systems for perception, action, and emotion (e.g., Fodor, 1975; Pylyshyn, 1984; Barsalou & Hale, 1993; Landauer & Dumais, 1997; Kintsch, 1998). According to this classic view, sensorimotor representations are transduced into amodal structures such as feature lists, semantic networks, and frames, and cognitive processes operate on these structures, not on memories of the original sensorimotor states. Moreover, the content of all types of concepts, including those encoded by words, is believed to consist entirely of combinations of these abstract symbols.

A very different line of thinking is currently being pursued by a growing number of researchers in linguistics (e.g., Hampe, 2005; Evans & Green, 2006; Bergen, 2007), philosophy (e.g., Prinz, 2005; Gallagher, 2005; Johnson, 2007), psychology (e.g., Barsalou, Simmons, Barbey, & Wilson, 2003; Gibbs, 2006; Barsalou, 2008; Semin & Smith, 2008), and neuroscience (e.g., Gallese & Lakoff, 2005; Keysers & Gazzola, 2006; Martin, 2007; Grafton, 2009), all of whom endorse one form or another of what is often called the Simulation Framework. The central claim of this approach is that semantic knowledge is not purely amodal but is instead grounded in modality-specific input/output systems. Rather than being transduced into amodal symbols, complex unimodal (e.g., visual) feature patterns that recur across different presentations of the same category of stimuli are captured by conjunctive units in correspondingly unimodal memory systems, and correlations between feature patterns across different modalities (e.g., visual and auditory) are captured by higher-order conjunctive units in more integrative cross-modal memory systems. Conceptual tasks, such as processing word meanings, are assumed to involve partial reenactments of the sensorimotor states that occurred when the referents were directly experienced. According to the Simulation Framework, these recapitulations are modality-specific in format; however, because they are driven in top-down rather than bottom-up fashion, they are prone to errors and biases, are rarely if ever represented as complete images, and are not necessarily conscious.[1] Evidence for this theory comes from an increasing number of behavioral and neuroscientific studies. Before delving into the neuroscientific data, I will attempt to convey a sense of what the behavioral research is like by summarizing three studies—all conducted by Rolf Zwaan and his colleagues—that provide evidence for linguistically induced, egocentrically anchored, perceptual simulations of shape, motion, and part-whole spatial relations (for reviews of other behavioral work see Barsalou, 2003, 2005, 2008; Pecher & Zwaan, 2005; Zwaan & Kaschak, 2009).

First, Zwaan, Stanfield, and Yaxley (2002) tested the prediction that during language comprehension people mentally represent the shapes of concrete objects denoted by count nouns. Subjects read simple sentences describing objects in certain locations, and after each item they indicated whether an object shown in a line drawing had been mentioned in the sentence. Reaction times were faster when the shape of the object in the line drawing matched the implied shape of the same object in the sentence. For example, after reading *The ranger saw the eagle in the sky*, subjects made faster responses to a picture of an eagle with outstretched wings than to a picture of an eagle with folded wings, but after reading *The ranger saw the eagle in the nest*, subjects' reaction times to the two pictures were reversed. A follow-up experiment obtained similar results when subjects simply named the depicted objects after reading the sentences. These findings are consistent with the claim that when people activate the meanings of count nouns during online sentence processing, they use contextual information to guide a mental reconstruction of the most likely shape that the designated object would have in that situation—a reconstruction that may facilitate visual processing of a subsequently presented picture (for an extension of this research see Yaxley & Zwaan, 2007).

Second, Zwaan, Madden, Yaxley, and Aveyard (2004; see also Kaschak, Madden, Therriault, Yaxley, Aveyard, Blanchard, & Zwaan, 2005) investigated whether dynamic perceptual simulations of motion are triggered during language comprehension. Subjects listened to sentences describing movement either toward or away from the observer—e.g., *The shortstop hurled the softball at you* or *You hurled the softball at the shortstop*. After each sentence, subjects were shown two pictures of objects, each presented for 500 msec and separated by 175 msec. On the critical trials, both pictures showed the kind of ball mentioned in the sentence, but the two images differed slightly in size, so that a smaller ball followed by a larger one suggested movement toward the participant, and a larger ball followed by a smaller one suggested movement away from the participant. The task was to indicate whether the two pictures showed the same object, and as predicted by the Simulation Framework, subjects were quicker when the motion trajectory suggested by the visual stimuli corresponded to the one implied by the preceding sentence. Given that the subjects were not told to relate the pictures to the sentences, and that the picture comparison task could, strictly speaking, be performed independently of the sentences, it is remarkable that the sentences nevertheless significantly influenced the subjects' responses. The authors interpret their results as supporting the view that linguistic descriptions of motion rapidly and perhaps automatically engender dynamic perceptual representations. They also point out that their study dovetails nicely with a famous experiment by Loftus and Palmer (1974) in which subjects watched short traffic safety films of car accidents and then answered questions about how fast the cars were going when they "contacted," "bumped into," "collided with," "smashed into," or "hit" each other. Subjects' judgments were affected by the speed expressed by the verb in question, and this finding has been used to argue that eyewitness testimony is highly susceptible to distortion by post-event questions. Zwaan et al. (2004, p. 618) offer the following account: "The verbs used to probe the participants' memory for the target event were cues to start dynamic perceptual simulations. Verbs associated with greater speed will produce faster perceptual simulations (i.e., more perceptual change per time unit) than verbs associated with lower speeds. These simulations, rather than the initial memories, were then used to estimate the speed of the vehicle, suggesting that words can, indeed, move mental representations."

Third, Zwaan and Yaxley (2003) examined whether spatial iconicity affects semantic relatedness judgments. Subjects read pairs of nouns that were presented vertically, one above the other, and indicated whether the members of each pair were semantically related. Some of the word pairs consisted of nouns denoting objects, or parts of objects, that typically have a vertical spatial arrangement—e.g., *branch–root, attic–basement, nose–mouth, flame–candle*. As expected, subjects' reaction times were faster when the spatial layout of the words iconically reflected the spatial layout of the designated objects (e.g., *branch* above *root*) than when the lexical and referential spatial layouts did not match (e.g., *root* above *branch*). The authors suggest that performance was influenced by spatial iconicity in the vertical condition because the two words in each pair activated their corresponding perceptual representations, which included the larger referential unit together with attentional focus on the relative locations of the parts in question—e.g., a canonically oriented tree with focus on the branches in the upper region and the roots in the lower region. When the spatial layout of the words was consistent with the spatial layout of the perceptual simulation, semantic relatedness was easier to judge.

These three studies exemplify the kinds of purely behavioral data that have been used to support the Simulation Framework. In the next section I turn to theoretical and empirical work in cognitive neuroscience that provides further elaboration and substantiation of this approach to semantic knowledge. As I will demonstrate, it is now possible to estimate with a relatively high degree of confidence some of the specific brain structures that are engaged when people perform the different types of linguistically triggered visual simulations that Zwaan and his colleagues have investigated.

CONTRIBUTIONS FROM COGNITIVE NEUROSCIENCE

Convergence Zone Theory and the Similarity in Topography Principle

As already mentioned, several competing theories are currently available regarding the organization, representation, and processing of semantic knowledge in the brain. One model that bears many important similarities to the Simulation Framework is Convergence Zone Theory (Damasio, 1989a-c; Damasio & Damasio, 1994; Tranel, Damasio, & Damasio, 1997; Tranel, Logan, Frank, & Damasio, 1997; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Meyer & Damasio, 2009), an extended version of which is Conceptual Topography Theory (Simmons & Barsalou, 2003). This approach assumes that when the exemplars of a conceptual category are experienced, they are initially represented at the cortical level as fluctuating patterns of activity across multiple modality-specific feature maps in early sensory areas. For example, watching a dog run across a field generates time-locked activation patterns in anatomically distributed regions of the visual system that contain feature maps dedicated to coding properties such as color, shape, and motion (Wandell, Dumoulin, & Brewer, 2007). Likewise, hearing a dog bark elicits activity in auditory feature maps, and petting a dog triggers activity in somatosensory feature maps. When a pattern of activity arises in a feature map, it is captured by ensembles of "conjunctive neurons" that reside in higher-order microscopic cortical areas referred to as "convergence zones" (CZs). The nature of CZs is discussed in depth in the references cited above, and only a few key points need to be elaborated here to provide a theoretical foundation for the review of literature in the section on Empirical Findings.

First, CZs exist at many hierarchical levels (Simmons & Barsalou, 2003). Within the visual system, the initial level consists of *property-specific CZs* that receive input from certain low-level feature maps (e.g., those for shape) and that contain conjunctive neurons for detecting complex patterns of activation within those maps (e.g., the distinctive shape properties of dogs). This is the main level of interest in the section on Empirical Findings, in which I describe in detail various visual semantic components of words for objects, events, and spatial relations. At the next level of coding, but still within the visual system, there are *modality-specific CZs* housing conjunctive neurons that register correlations among different visual properties of objects (e.g., the tendency for dogs to have both certain shapes and certain movements). Mounting evidence suggests that these CZs reside in the anterior temporal lobes (e.g., Quiroga Reddy, Kreiman, Koch, & Fried, 2005; Bright, Moss, Longe, Stamatakis, & Tyler, 2007; Rogers, Hocking, Noppeney, Mechelli, Gorno-Tempini, Patterson, & Price, 2007; for a review see Patterson et al., 2007; see also Lin, Chen, Kuang, Wang, & Tsien, 2007). Finally, at the apex of the processing hierarchy there are *crossmodal CZs* comprised of conjunctive neurons that create complete category "files" by detecting correspondences across different modalities (e.g., visual as well as auditory, tactile, and olfactory information about dogs). It is important to note, however, that, as pointed out by Simmons and Barsalou (2003, p. 465), "crossmodal CZs may not literally represent this conceptual content, as is typically the case in current theories that employ knowledge stores. Instead conjunctive neurons in these CZs may point to a hierarchy of lower-order conjunctive neurons that eventually activate features in feature maps." Anatomically, crossmodal CZs may depend on the anteromedial temporal lobes, especially the perirhinal cortex (e.g., Murray & Richmond, 2001; Taylor, Moss, Stamatakis, & Tyler, 2006; Lambon Ralph, Pobric, & Jeffries, 2009).[2]

Second, as alluded to in the passage quoted from Simmons and Barsalou (2003), CZs are reciprocally connected to each other and to feature maps via extensive feedforward and feedback pathways, thereby enabling both bottom-up recognition and top-down recall. This constitutes one of the key mechanisms that relates CZ theory to the Simulation Framework, as indicated by Simmons and Barsalou (2003, p. 461) in their detailed discussion of the concept of a wheel:

> On perceiving a car, the edges of a wheel are represented in visual feature maps. If selective attention focuses on this region of the perceived car, conjunctive neurons in [a property-specific] CZ capture the features in this region. Later, on reactivating these conjunctive neurons [in a top-down manner], the visual representation of this particular wheel is partially re-enacted. As the perceived wheels of subsequent cars similarly receive attentional processing, they activate overlapping conjunctive neurons in the CZ, thereby linking the visual features across different wheels to each other. The result is ... a *property simulator*, namely, a system that can simulate the various forms a property takes in different categories. As the simulator for wheel develops, it produces simulations of different wheels in different objects, such as cars, bicycles, and skates.[3]

Third, for all types of CZs, conjunctive neurons are anatomically distributed according to the Similarity in Topography principle: "The spatial proximity of two neurons in a CZ reflects the similarity of the features they conjoin. As two sets of conjoined features become more similar, the conjunctive neurons that link them lie closer together in the CZ's spatial topography" (Simmons & Barsalou, 2003, p. 457). Support for this principle comes not only from single-cell recording studies with nonhuman species (e.g., Tanaka, 1996; Graziano & Aflalo, 2007b; Kiani, Esteky, Mirpour, & Tanaka, 2007), but also from functional neuroimaging studies with humans, as discussed below.

Before reviewing a number of experimental studies that address the neural bases of different visual semantic components of words, it is worthwhile to quote at some length an example used by Damasio (1989c, p. 27) to illustrate his seminal views about how CZ theory accounts for conceptual knowledge, including word meaning:

> The presentation of a line drawing of a violin, or presentation of the word "violin" (aurally or orthographically), generates a set of time-locked activations of sensory and motor representations. The activations are generally pertinent to manipulable man-made objects, more specifically pertinent to musical instruments of the string variety, and even more narrowly so to the class of violins. In the visual realm the perceiver is likely to evoke representations of shape, motion, color, and texture which will vary from individual to individual according to the personal experience with violins that each has enjoyed. For those who have

held violins in their own hands, or even played a violin, numerous somatosensory representations will also be evoked related to tactile impressions of wood and strings, or relative to the pressure the instrument will have exerted in the perceiver's body. But that is hardly all. Auditory representations of the range of sounds produced by the instrument may also be generated; motor programs according to which the appropriate posture and motions applicable to a violin can be organized may also be evoked and readied for appropriate display; finally, a range of somatic states appropriate to one's experience with violins, e.g., like or dislike, pleasurable or painful sensation, and so on, will also be activated. In short, a wide array of representations will be generated that together define the meaning of the entity, momentarily . . . . The mechanism that permits co-activation of representations depends on devices I have called *convergence zones*, which are ensembles of neurons that "know about" the simultaneous occurrence of patterns of activity during the perceived or recalled experience of entities and events. The probability of simultaneous activation of representations prompted by a stimulus thus depends on the operation of convergence zones which, so to speak, embody a binding code for those representations . . . .

Empirical Findings

Having outlined the basic architecture of CZ theory, I now show how this model, which can be regarded as a variant of the Simulation Framework, provides a useful context for understanding a number of recent studies that have begun to address, from the perspective of cognitive neuroscience, the intriguing question of how words capture visual experience. The set of studies that will be reviewed focus on the neural underpinnings of several different visual semantic components of words for objects, events, and spatial relations. A major finding that is supported by these studies, and that I highlight throughout the review, is that, as predicted by the Simulation Framework, specific types of lexically encoded visual properties appear to depend on cortical areas that are very closely related to those that subserve the perception of the same properties.

Words for Objects Count nouns that designate different types of bounded concrete entities have complex meanings that include information about the appearance of objects in the given category. Here I focus on the neural correlates of three visual semantic components of such words: color properties, shape properties, and motion properties. During our ordinary observation of the world, these three attributes of objects are tightly bound together in unified conscious images (e.g., Edelman & Tononi, 2000; Koch, 2004). Thus, if you watch a black crow fly over a corn field, the color, shape, and motion of the bird are experienced as being perfectly fused so that, for instance, the black color does not "spill out" beyond the edges of the flapping wings. In the brain, however, these different perceptual properties are known to be represented in anatomically separate processing streams of the visual system, and, crucially, recent research suggests that the semantic representations of the corresponding properties parallel this organization. The following brief review of this research gives special emphasis to the influential work of Alex Martin and his colleagues (for a more detailed summary see Martin, 2007).

Color Properties. Many kinds of objects have typical or "canonical" colors. This applies to numerous categories of artifacts whose colors are determined by social convention (e.g., yellow bulldozers) and even more strongly to various categories of animals (e.g., white swans) and plant life (e.g., green lettuce) whose colors are genetically programmed. Such object-color associations constitute an important part of a person's semantic knowledge of the relevant nouns.

Turning to the brain, substantial evidence suggests that color constancy—i.e., the ability to see an object as having a stable color under different lighting conditions—is mediated by area V4, which occupies a portion of the lingual gyrus on the ventral surface of the occipital lobe (Zeki & Bartels, 1999). This area is activated during tasks requiring color constancy (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Zeki, Watson, Lueck, Friston, Kennard, & Frackowiak, 1991; Sakai, Watanabe, Onodera, Uchida, Kato, Yamamoto, Koizumi, & Miyashita, 1995), and damage to it causes achromatopsia, that is, acquired color blindness (Zeki, 1990; Bouvier & Engel, 2006; see also Sacks, 1995). Thus, it is generally assumed that V4 plays a crucial role in conscious, passive color sensation (for a splendid, albeit now somewhat dated, discussion, see Zeki, 1993; for recent advances see Conway, Moeller, & Tsao, 2007).

To identify the cortical regions that underlie semantic knowledge about the canonical color properties of objects, Martin, Haxby, Lalonde, Wiggs, and Ungerleider (1995) conducted two experiments using positron emission tomography (PET). In both experiments the subjects' task was to produce words denoting the typical colors of objects, but in the first experiment the stimuli were black-and-white line drawings of objects (e.g., a drawing of a child's wagon, to elicit *red*), and in the second they were printed words for objects (e.g., *wagon*, again to elicit *red*). Relative to a baseline task in which the subjects simply named each object/word that was presented, both of the experimental tasks elicited bilateral activation in an area of the fusiform gyrus just anterior to V4. This area may function as a property-specific CZ that is reciprocally connected with V4 and that houses distinct populations of conjunctive neurons for distinct color categories. Further evidence for this view comes from several additional studies. For example, the fusiform color area is engaged when subjects name the colors of chromatic line drawings of objects (Chao & Martin, 1999), when subjects imagine named colors (Howard, Ffytche, McKeefry, Ha, Wodruff, Bullmore, Simmons, Williams, David, & Brammer, 1998), and when subjects with synesthesia report automatically experiencing certain colors on hearing certain words (Paulesu, Harrison, Baron-Cohen, Watson, Goldstein, Heather, & Frackowiak, 1995) (see Fig. 14.1). Moreover, direct electrical stimulation of this

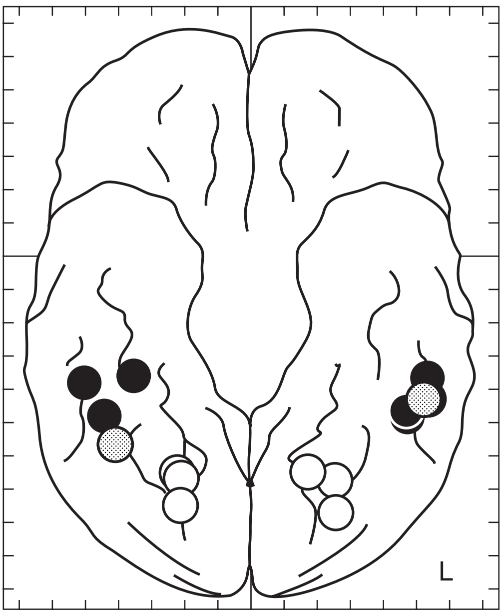


FIGURE 14.1. Summary of findings suggesting that color information is stored in the ventral temporal lobes, anterior to the regions that mediate passive color perception. White circles indicate the location of regions active during passive color perception (Corbetta et al., 1990; Sakai et al., 1995; Zeki et al., 1991); black circles indicate the location of regions in the ventral temporal lobes active when subjects generated color words (Martin et al., 1995, two studies; Wiggs et al., 1999; Chao & Martin, 1999); the gray circle on the left hemisphere shows the location of the region active when color-word synesthetes experienced color imagery (Paulesu et al., 1995); the gray circle on the right hemisphere shows the location of the region active in normal subjects during a color imagery task (Howard et al., 1998).

(Adapted from Martin et al., 2000.)

area produces artificial color percepts (Murphey, Yosher, & Beauchamp, 2008; see also Wandell, 2008). Most impressive of all, however, is a recent functional magnetic resonance imaging (fMRI) study by Simmons, Ramjee, Beauchamp, McRae, Martin, and Barsalou (2007) that predicted and confirmed partially overlapping patterns of activation in the fusiform color area during the following two conditions: first, a conceptual property verification task in which subjects judged whether the type of color denoted by an adjective (e.g., *green*) generally applies to the type of object denoted by a noun (e.g., *lettuce*); and second, the Farnsworth-Munsell 100 Hue Test, which is a perceptual task that requires not just color constancy, but also close attention to the distinctions between color categories (see also Beauchamp, Haxby, Jennings, & DeYoe, 1999). According to Simmons et al. (2007, p. 2808), these findings indicate that, in keeping with the Simulation Framework, "retrieving property knowledge shares [part of] the neural substrate underlying property perception."

On the other hand, an opponent of the Simulation Framework—that is, someone who is skeptical of the view that conceptual knowledge is grounded in the brain's modality-specific systems—might argue that the fusiform activations that Simmons et al. (2007) observed during the property verification task could reflect nonsemantic visual imagery rather than intrinsically semantic content. Simmons et al. (2007, pp. 2807-2808) address this criticism as follows:

> The neural bases of color perception and explicit color imagery share many commonalities, and were we to ask subjects to explicitly imagine colors, we would expect to see areas of activation similar to those reported here. This expectation follows naturally from a theoretical stance that perceptual information is used in many areas of cognition, including imagery and memory.... Importantly, at no point were subjects instructed to use imagery to perform the property verification task, nor from the amodal perspective should imagery even be necessary. In fact, most amodal accounts would posit that property information is stored propositionally with the relevant concept information. This being the case, it would seem extremely odd for a proponent of amodal accounts to argue that the task cannot be performed using the amodal representations central to amodal theories, but instead must be performed using additional, ancillary, effortful processes.... If property information is amodal, then why would one need to "imagine" the property? Yet subjects do activate modality-specific cortex when accessing property knowledge.

Moreover, returning to the topic of object-color associations, further evidence for the Simulation Framework comes from a recent behavioral study in which subjects were asked to adjust, via interactive manipulation of a graphic display, the color of natural fruits until they appeared gray. Remarkably, the stimuli were not perceived to be gray until their color was shifted away from the observers' independently identified gray point in the direction opposite the canonical color of the fruit (e.g., slightly toward blue for a banana). This discovery suggests that "color sensations are not determined by the incoming sensory data alone, but are significantly modulated by high-level visual memory" (Hansen, Olkkonen, Walter, & Gegenfurtner, 2006, p. 1367). Indeed, this phenomenon might be generated in part by top-down projections from the fusiform color area to V4.

Finally, it is noteworthy that several recent studies point to an intriguing left-hemisphere-mediated interaction between linguistic and perceptual representations of color. Gilbert, Regier, Kay, and Ivry (2006) report an experiment in which subjects focused their visual attention on a central fixation marker and were then shown a ring of 12 colored squares. One of the squares (the target) was a different hue than the others, and the task was to indicate as quickly as possible whether it was in the left or right visual field. Responses were significantly faster when the target and distractor colors had different names (e.g., a blue target among green distractors) compared to when they had the same name (e.g., a light green target among somewhat darker green distractors). However, this category effect occurred only when the targets were presented in the right visual field. Furthermore, a follow-up experiment showed that the effect disappeared when subjects performed a secondary task requiring verbal working memory, but not when they performed a secondary task requiring spatial working memory. Because information in the right visual field is routed first to the language-dominant left hemisphere, whereas information in the left visual field is routed first to the right hemisphere and must then cross the corpus collosum to access lexical codes, the results can be interpreted as providing evidence for a lexical influence on rapid color discrimination. Gilbert et al. (2006, p. 493) summarize the implications of their study as follows:

> Previous studies addressing the possible influence of language on perception have tended to look for a simple yes or no answer to the question. Our findings suggest a more complex picture, based on the functional organization of the brain. The [left hemisphere] appears to sharpen visual distinctions between lexically defined categories and to blur visual distinctions within these categories, whereas the right hemisphere does so much less, if at all. To the degree that these results can be generalized to everyday perception, our representation of the visual world may be, at one and the same time, filtered and not filtered through the categories of language.

Similar findings have been obtained in several subsequent studies (Drivonikou, Kay, Regier, Ivry, Gilbert, Franklin, & Davies, 2007; Roberson, Pak, & Hanley, 2007; Franklin, Drivonikou, Bevis, Davies, Kay, & Regier, 2008), and future research will undoubtedly delineate the nature of this fascinating phenomenon in even greater detail, both behaviorally and neurobiologically.

Shape Properties. Perhaps the most critical visual semantic component of count nouns is shape. This point is expressed quite forcefully in the following passage from Landau and Jackendoff (1993, p. 218):

> In the average [English-speaking] adult vocabulary, there are roughly 10,000 names for things—count nouns that label different kinds of objects. For a large proportion of object categories, shape is among the most important criteria for identification, and in particular for judgments of what a thing should be called: Categories of things with the same shape, including natural kind objects and artifacts, often share the same name.

Based on many neuropsychological and functional neuroimaging studies, it is now well established that within the visual system the shape properties of objects are represented in the ventral occipitotemporal processing stream (e.g., Farah, 2004; Haxby, Gobbini, & Montgomery, 2004; Milner & Goodale, 2006; Vandenbulcke, Peeters, Fannes, & Vandenberghe, 2006). Currently, one of the most intensively investigated questions is the following: Across the large expanse of ventral temporal cortex, are CZs for the shape properties of different categories of objects uniformly distributed, or are they clustered together in patches (Reddy & Kanwisher, 2006; Op de Beeck, Haushofer, & Kanwisher, 2008)? Although there is some evidence for distributed coding, there is even greater evidence—consistent with the Similarity in Topography principle—that certain areas are preferentially responsive to certain categories of objects, particularly faces (e.g., Kanwisher & Yovel, 2006; Tsao, Freiwald, Tootell, & Livingstone, 2006), nonfacial body parts (e.g., Peelen & Downing, 2007; see also Kemmerer & Tranel, 2008), animals (e.g., Chao, Haxby, & Martin, 1999; Chao, Weisberg, & Martin, 2002), tools (e.g., Chao et al., 1999; Chao et al., 2002), scenes (e.g., Epstein & Kanwisher, 1998; Epstein, DeYoe, Press, Rosen, & Kanwisher, 2001), and printed words (e.g., McCandliss, Cohen, & Dehaene, 2003; Gaillard, Naccache, Pinel, Clemenceau, Volle, Hasboun, Dupont, Maulac, Adam, & Cohen, 2006).

Here I focus on several functional neuroimaging studies that point to segregated cortical representations of the shapes of animals and tools.[4] In one set of fMRI experiments, Chao, Haxby, and Martin (1999) evaluated perceptual processing of animals and tools by using passive viewing tasks and delayed match-to-sample tasks; in addition, they evaluated conceptual processing of animals and tools by using picture naming tasks and property verification tasks (i.e., answering yes/no questions such as "Forest animal?" and "Kitchen tool?" in response to printed words for animals and tools). Across all of the tasks, significantly greater bilateral activation for animals was consistently found in a lateral portion of the fusiform gyrus, whereas significantly greater bilateral activation for tools was consistently found in a medial portion of the fusiform gyrus (see Fig. 14.2, left panel). As the authors point out, it is especially interesting that these adjacent but nevertheless distinct regions of the fusiform gyrus were activated not only by pictures but also by words.

The same question arises here, however, as arose for the investigation by Simmons et al. (2007) of object-color associations. In particular, we might assume that the activations evoked by words were merely a reflection of the subjects' deliberate efforts to conjure up explicit visual images of the shapes of the lexically encoded animals and tools. Evidence against this interpretation, and in favor of the hypothesis that the lexically driven category-related fusiform activations index automatic semantic processing, comes from a recent fMRI study by Wheatley, Weisberg, Beauchamp, and Martin (2005) that took advantage of the neurophysiological phenomenon known as "repetition suppression." Basically, repetition suppression is a decrease in neural response associated with the repeated presentation of identical, or conceptually related, stimuli. It reflects the

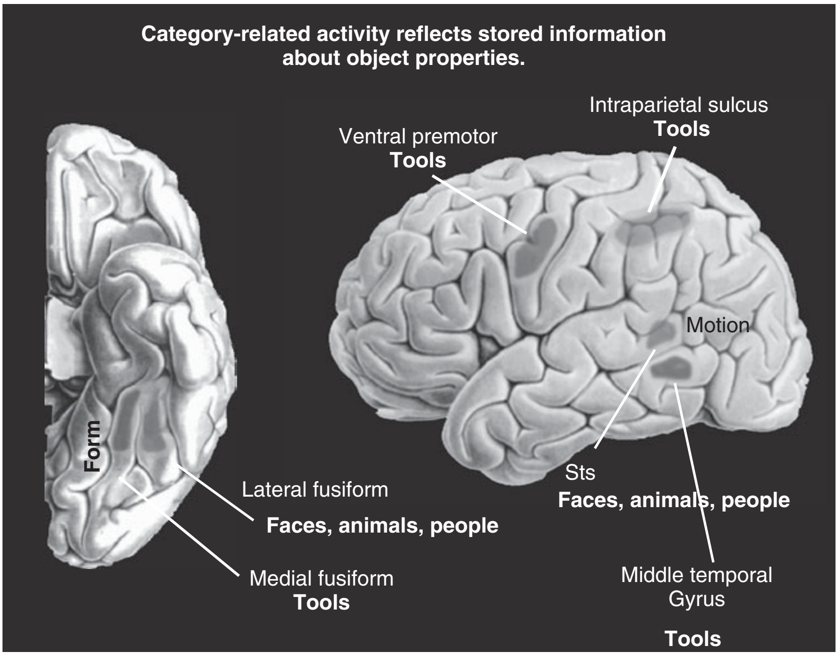


FIGURE 14.2. Grayscale rendition of Color Plate 8. See Color Plate 8 for interpretation. Distinct activation patterns for animal (including human) and tool concepts as revealed by functional neuroimaging studies employing both pictures and printed words as stimuli. The left image shows parallel but segregated lateral (red) and medial (blue) sectors of the fusiform gyrus that may contain assemblies of conjunctive neurons for capturing the complex shape patterns of the two categories of objects. The right image shows parallel but segregated superior temporal sulcus (STS, red) and middle temporal gyrus (blue) regions that may contain assemblies of conjunctive neurons for capturing the complex motion patterns of the two categories of objects. The right image also shows ventral premotor and intraparietal sulcus regions that may contribute to visuomotor knowledge of tools.

(This figure was kindly provided by Alex Martin.)

well-established fact that prior exposure to a stimulus leads to greater processing efficiency (for a review see Grill-Spector, Henson, & Martin, 2006). In the study by Wheatley et al. (2005), subjects read rapidly presented word pairs (each word presented for 150 msec with a 100-msec interstimulus interval) that were either unrelated (e.g., *celery giraffe*), related (e.g., *horse goat*), or identical (e.g., *camel camel*). The investigators found that as the degree of semantic relatedness between the two words progressively increased for a particular category—in this case, animals—the neural activation evoked by the second word progressively decreased in the lateral fusiform gyrus, this being the same area that Chao et al. (1999) linked with the animal category. This repetition suppression effect may be the neural manifestation of automatic priming of semantically related words (see also Gold, Balota, Jones, Powell, Smith, & Anderson, 2006; Wible, Han, Spencer, Kubiciki, Niznikiewicz, Jolesz, McCarley, & Nestor, 2006). Moreover, given the processing time constraints imposed by the task, it is highly unlikely that the fMRI results reflect explicit visual images that the subjects deliberately generated *after* understanding the words. Rather, as Wheatley et al. (2005, p. 1877) put it:

> It may be that the visual image of an object is automatically retrieved as an unconscious and obligatory by-product of normal word reading. In this case, automatic, implicit generation of an object image would be the mechanism by which we access an important property underlying the meanings of words denoting concrete entities. In this sense, implicit visual imagery would be an obligatory component of reading for meaning.

For related data and arguments see Hauk, Davis, Kherif, and Pulvermüller (2008a).

The convergent results of the studies by Chao et al. (1999) and Wheatley et al. (2005) suggest that the shape features of the meanings of concrete count nouns are captured by conjunctive neurons in ventral temporal CZs that not only overlap partially with the CZs that subserve visual perception of the same properties (consistent with the Simulation Framework), but are also clustered according to semantic category (consistent with the Similarity in Topography principle). Further evidence for these ideas comes from a number of other studies (Martin, Wiggs, Ungerleider, & Haxby, 1996; Thompson-Schill, Aguirre, D'Esposito, & Farah, 1999; Okada, Tanaka, Nakai, Nishizawa, Inui, Sadato, Yonekura, & Konishi, 2000; Chao et al., 2002; Whatmough, Chertkow, Murtha, & Hanratty, 2002; Kan, Barsalou, Solomon, Minor, & Thompson-Schill, 2003; Martin & Weisberg, 2003; Devlin, Rushworth, & Matthews, 2005; Mechelli, Sartori, Orlandi, & Price, 2006; Noppeney, Price, Penny, & Friston, 2006; Mahon, Anzellotti, Schwarzbach, Zampini, & Caramazza, 2009; for studies focusing specifically on hemispheric asymmetries see Lincoln, Long, & Baynes, 2007; Gilbert, Regier, Kay, & Ivry, 2008; for pertinent interelations between human fMRI and monkey single-cell data see Kriegeskorte, Mur, Ruff, Kiani, Bodurka, Esteky, Tanaka, & Bandettini, 2008). Thus, returning to the purely behavioral study by Zwaan et al. (2002) mentioned in the section on The Simulation Framework, it may be that when people read sentences such as *The ranger saw the eagle in the sky* and *The ranger saw the eagle in the nest*, the contextually appropriate eagle images that they mentally simulate depend on the lateral fusiform gyrus.

Motion Properties. Yet another visual semantic component of many count nouns involves the characteristic motion patterns of the designated objects. To give a few simple examples, part of our knowledge of the word *rabbit* is the typical hopping movement of this kind of animal, and part of the meaning of *scissors* is the distinctive cutting movement of this kind of tool. As with the attributes of color and shape, the motion information encoded by count nouns appears to be neurally represented in cortical areas that are closely related to those that subserve perception of the same properties.

It is well established that area MT (also known as MT+, hMT+, V5, and hOc5), which resides in the vicinity of the anterior occipital and lateral occipital sulci, is critically involved in the passive perception of moving visual stimuli (e.g., Malikovic, Amunts, Schleicher, Mohlberg, Eickhoff, Wilms, Palomero-Gallagher, Armstrong, & Zilles, 2007). Damage to this area causes akinetopsia, that is, acquired motion blindness, a rare neurological syndrome that is just as striking as achromatopsia for its specificity (Zeki, 1991; Zihl, Von Cramon, Mai, & Schmid, 1991). Although MT does not itself appear to distinguish systematically between different types of object-associated motion, it projects to higher-order posterolateral temporal areas that clearly do. These areas consist of property-specific CZs that are organized in the following way according to the Similarity in Topography principle (see Fig. 14.2, right panel). One processing stream extends from MT into regions of the posterior superior temporal sulcus that respond preferentially to the sight of biological (e.g., human) motion patterns, perhaps with further segregation according to body parts (for reviews of a large literature see Grossman, 2006, and Blake & Shiffrar, 2007; for a neurophysiologically plausible computational model see Giese & Poggio, 2003, and Giese, 2006). Another processing stream extends from MT into areas of the posterior middle temporal gyrus that respond preferentially to the sight of nonbiological (e.g., tool) motion patterns (Beauchamp, Lee, Haxby, & Martin, 2002, 2003; Martin & Weisberg, 2003; Beauchamp & Martin, 2007; see also Martin et al., 1996; Tranel et al., 1997a; Kellenbach, Brett, & Patterson, 2003; Boronat, Buxbaum, Coslett, Tang, Saffran, Kimberg, & Detre, 2005; Noppeney, Josephs, Kiebel, Friston, & Price, 2005; Weisberg, van Turennout, & Martin, 2007).

During the past 20 years, many functional neuroimaging studies have provided evidence that these posterolateral temporal areas—especially those in the left hemisphere—contribute not only to the transient perception, but also to the long-term semantic representation, of object-associated motion properties encoded by concrete count nouns. Some of the original studies employed the so-called "verb generate" paradigm in which subjects are presented with the name of an object (e.g., *knife*) and must produce a semantically related motion verb (e.g., *cut*) (for a review of the early literature see Martin, Ungerleider, & Haxby, 2000; for further discussion of motion verbs see the section on Words for Events). More recently, it has been found that semantic representations of biological (e.g., human) and nonbiological (e.g., tool) motion properties are supported by the two segregated perceptual processing streams described in the previous paragraph. For instance, in the study by Chao et al. (1999) that I discussed in the context of shape properties, another important finding was that the pathway for biological motion patterns was recruited more than the one for nonbiological motion patterns during tasks involving either animal pictures or animal words, whereas the opposite asymmetry occurred during tasks involving either tool pictures or tool words (see Fig. 14.2, right panel). As shown in Figure 14.2, tool pictures as well as tool words also elicited activation in two regions—the anterior intraparietal sulcus and the ventral premotor cortex—that subserve visuomotor knowledge about how such objects are grasped and manipulated; however, I will not discuss these results further, as my focus here is restricted to visual information (for an in-depth review see Lewis, 2006; for another perspective see Mahon & Caramazza, 2005, 2008, 2009). The available data, therefore, suggest that the motion patterns associated with the types of objects encoded by many count nouns are captured by conjunctive neurons in category-related CZs in the posterolateral temporal cortex that reside very close to those that are engaged during perception of the same motion patterns.

Summary. The research reviewed above supports the idea that, as proposed by CZ theory and the broader Simulation Framework, the typical color, shape, and motion properties of various categories of objects encoded by count nouns depend on anatomically segregated regions of the temporal lobe that are either overlapping with or anterior to some of the regions that underlie perception of those properties. Consider, for example, the visual components of the meaning of the word *elephant*.

The canonical gray color of elephants seems to be captured by conjunctive neurons located close to those for color perception; the idiosyncratic shape of elephants seems to be captured by conjunctive neurons located close to those for shape perception; and the characteristic movements of elephants seem to be captured by conjunctive neurons located close to those for motion perception. When the word *elephant* is heard or read, these cortically distinct representations of color, shape, and motion properties are automatically activated, perhaps via certain conjunctive neurons in an anterior temporal region that houses higher-order modality CZs for binding in long-term semantic memory the diverse visual attributes of objects (Bright et al., 2007). This spatially distributed but temporally synchronous[5] pattern of activation may, but need not, co-occur with an explicit conscious image of an elephant. Either way, the two most important points to bear in mind are, first, that the relevant fragments of semantic knowledge appear to be represented in modality-specific and, even more narrowly, property-specific format, and second, their transient coactivation on presentation of the word *elephant* can be interpreted as constituting a rough reenactment or simulation of how the properties would normally be represented during perception.

Words for Events Verbs are the preferred class of words for describing events, and many verbs encode types of events that involve particular visual motion patterns. To take an especially well-studied case, Slobin (2000, 2003, 2004) notes that in English the multidimensional psychological space of "manner of locomotion" is intricately partitioned into discrete categories by well over 100 verbs that fall into specialized subclasses such as the following: rapid motion (e.g., *dash, hurry, rush, scramble, sprint*), leisurely motion (e.g., *amble, drift, mosey, saunter, stroll*), smooth motion (e.g., *glide, slide, slink, slip, slither*), awkward motion (e.g., *limp, lurch, stagger, stumble, trip*), furtive motion (e.g., *creep, sidle, skulk, sneak, tiptoe*), manners of walking (e.g., *march, plod, sashay, strut, trudge*), and manners of jumping (e.g., *bound, hop, jump, leap, spring*). Such highly specialized verbs are not just dictionary entries, but are actively employed by speakers in a variety of contexts, including oral narrative, spontaneous conversation, creative writing, naming videoclips of motion events, and speeded fluency, i.e., listing as many motion verbs as possible in 1 minute (Slobin, 2003).

As I indicated in the discussion of the motion properties of objects, several functional neuroimaging studies have shown that the "verb generate" task elicits activation in the posterolateral temporal cortex anterior to MT, predominantly in the left hemisphere. If this reflects retrieval of stored visual semantic information about motion, we would expect to find similar patterns of activation during other tasks that require access to the motion component of verb meanings. This prediction has been confirmed in two studies by Joseph Kable and colleagues. In the first study, Kable, Lease-Spellmeyer, and Chatterjee (2002) asked subjects to perform the same kind of semantic similarity judgment task in two conditions, one with pictures of events as stimuli and the other with the corresponding verbs as stimuli. In both conditions, subjects saw three items in a triangular array—one at the top of the screen and two at the bottom—and had to determine whether the top item was more similar to the left-hand bottom item or to the right-hand bottom item. For example, they had to recognize that "digging" is more similar to "shoveling" than to "listening." The investigators also ran control experiments using pictures of objects and corresponding nouns as stimuli. When subjects performed the judgment task with pictures, there was greater activation for events, compared to objects, in MT and in nearby portions of the posterolateral temporal cortex, with stronger signals in the right than the left hemisphere. The engagement of MT is consistent with previous evidence that this region responds not only to veridical motion but also to static pictures that imply motion, such as a dolphin photographed in mid-leap out of the ocean (Kourtzi & Kanwisher, 2000; Senior, Barnes, Giampietro, Simmons, Bullmore, Brammer, & Davis, 2000; Senior, Ward, & David, 2002; Urgesi, Moro, Candidi, & Aglioti, 2006; Assmus, Giessing, Weiss, & Fink, 2007; but see also Alford, van Donkelaar, Dassonville, & Marrocco, 2007; Fawcett, Hillebrand, & Singh, 2007). When subjects performed the judgment task with words, there was greater activation for verbs, compared to nouns, in portions of the posterolateral temporal cortex anterior to MT, but not in MT itself; moreover, the signals were stronger in the left than the right hemisphere. In a follow-up study, Kable, Kan, Wilson, Thompson-Schill, and Chatterjee (2005) replicated these findings under conditions that employed the same kind of judgment task but with stimuli that required greater discrimination of subtle aspects of the visual motion patterns encoded by verbs. For instance, subjects had to determine that "skipping" is more similar to "bouncing" than to "rolling."

Kable et al. (2005) offer the following interpretation of their discovery that words for events activate regions anterior to MT, whereas pictures of events activate MT itself: "These anterior-posterior differences between words and pictures could be evidence of a gradient of motion information represented in the occipito-temporal cortex, with areas closer to [MT] representing more concrete visual information and areas closer to the middle temporal gyrus representing more abstract propositional information" (p. 1863). The notion of a concrete-to-abstract gradient is plausible, but it is not clear if motion information really becomes more "propositional" as it becomes more "abstract." From the perspective of CZ theory and the Simulation Framework, the information is still modality-specific in format, and it may consist of visual motion properties at varying levels of detail, captured by hierarchies of conjunctive neurons in CZs extended along the posterolateral temporal cortex. Although it is indisputable that the types of motion patterns implied by static pictures are more concrete than those encoded by verbs, it is also the case that, as emphasized by both Slobin (2000, 2003, 2004) and Kable et al. (2005), the types of motion patterns encoded by verbs can be remarkably detailed. In fact, it is precisely for this reason that some linguists, such as Jackendoff (2002, p. 350), have argued that the semantic nuances distinguishing verbs such as *walk, jog, limp, strut,* and *shuffle* are best left to modality-specific sensorimotor systems.

In a recent fMRI study, Kemmerer, Gonzalez Castillo, Talavage, Patterson, and Wiley (2008) employed a "triads" paradigm exactly like the one used by Kable et al. (2002, 2005) to investigate the neural substrates of the following five classes of verbs, based on Levin (1993): Running verbs (e.g., *run, jog, walk*), Speaking verbs (e.g., *shout, mumble, whisper*), Hitting verbs (e.g., *hit, poke, jab*), Cutting verbs (e.g., *cut, slice, hack*), and Change of State verbs (e.g., *shatter, smash, crack*). Relative to a baseline condition, the five verb classes evoked complex and widely distributed patterns of activation that differed from each other in many theoretically important ways. For present purposes, however, what is most relevant is that all five classes significantly engaged the left posterolateral temporal cortex anterior to MT, just as in the studies of Kable et al. (2002, 2005). Moreover, there was some evidence for a rough body-part-based organization, as unique portions of the posterolateral temporal cortex were found to be activated by Running verbs (encoding distinctive leg movements), Cutting verbs (encoding distinctive hand movements), and Speaking verbs (encoding distinctive mouth movements)[6] (see Fig. 14.3).

Interestingly, in four other studies Daniel Tranel and colleagues claim to have identified links between verb processing and area MT itself, suggesting that, contrary to the findings of Kable et al. (2002, 2005) and Kemmerer et al. (2008), this region may actually play a role in representing the motion component of verb meanings. First, in a PET study Damasio, Grabowski, Tranel, Ponto, Hichwa, and Damasio (2001) presented subjects with pictures of actions being performed with tools (e.g., stirring a cup of coffee with a spoon) and asked them to name the actions in one condition (e.g., *stirring*) and to name the tools in another condition (e.g., *spoon*). The subtraction of naming tools from naming actions revealed robust activation in an area that the authors refer to as MT, suggesting that the engagement of this region was driven

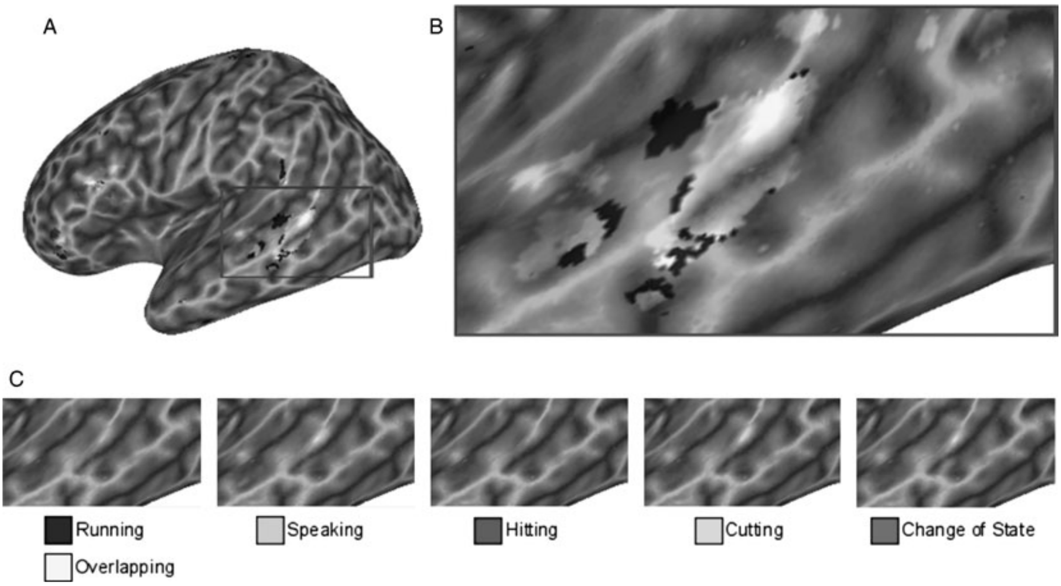


FIGURE 14.3. Grayscale rendition of Color Plate 9. See Color Plate 9 for interpretation. Activation patterns for five classes of verbs—Running, Speaking, Hitting, Cutting, and Change of State—in the posterolateral temporal cortex (PLTC). (A) Activations rendered on the left hemisphere of an inflated brain, with the PLTC enclosed in a red box. Yellow patches signify areas of overlapping activation for two or more verb classes. Other colored patches signify areas of activation unique to particular verb classes, according to the color key in (C). (B) Enlargement of the red box in (A). (C) Activations for each separate verb class in the territory of the PLTC indicated by the red box in (A) and (B).

(Adapted from Kemmerer et al., 2008.)

more by the verb retrieval task than by the pictorial stimuli. Second, in another PET study Tranel, Martin, Damasio, Grabowski, and Hichwa (2005) used a similar kind of experimental design as Damasio et al. (2001), with subjects performing both action-naming and tool-naming tasks for the same stimuli, except in this study the target verbs and nouns were homophones (e.g., *comb* to name both the action and the tool). Once again, the subtraction of naming tools from naming actions revealed robust activation in an area that the authors refer to as MT. Third, in a large-scale lesion study with 90 brain-damaged patients, Tranel, Kemmerer, Adolphs, Damasio, and Damasio (2003) found that many patients who failed two standardized tests that probe conceptual knowledge of actions had lesions that included either MT or the white matter underneath it. Fourth, in another large-scale lesion study with 78 brain-damaged patients, Tranel, Manzel, Asp, and Kemmerer (2008) found that several patients who failed two action naming tests—one with static pictures as stimuli and the other with dynamic videoclips as stimuli—had lesions encompassing MT (see Fig. 14.4).

These four studies seem to provide some leverage for the idea that, contrary to the findings of Kable et al. (2002, 2005) and Kemmerer et al. (2008), area MT may in fact comprise part of the neural basis for the visual motion component of verb meanings (for convergent fMRI evidence see Pirog Revill, Aslin, Tanenhaus, & Bavelier, 2008; and for relevant behavioral studies see Kaschak et al., 2005, Meteyard, Bahrami, & Vigliocco, 2008, and Meteyard, Zokaei, Bahrami, & Vigliocco, 2008). The studies are not without limitations, however. For instance, neither PET study precisely localized MT, and the relevant patients in the two lesion

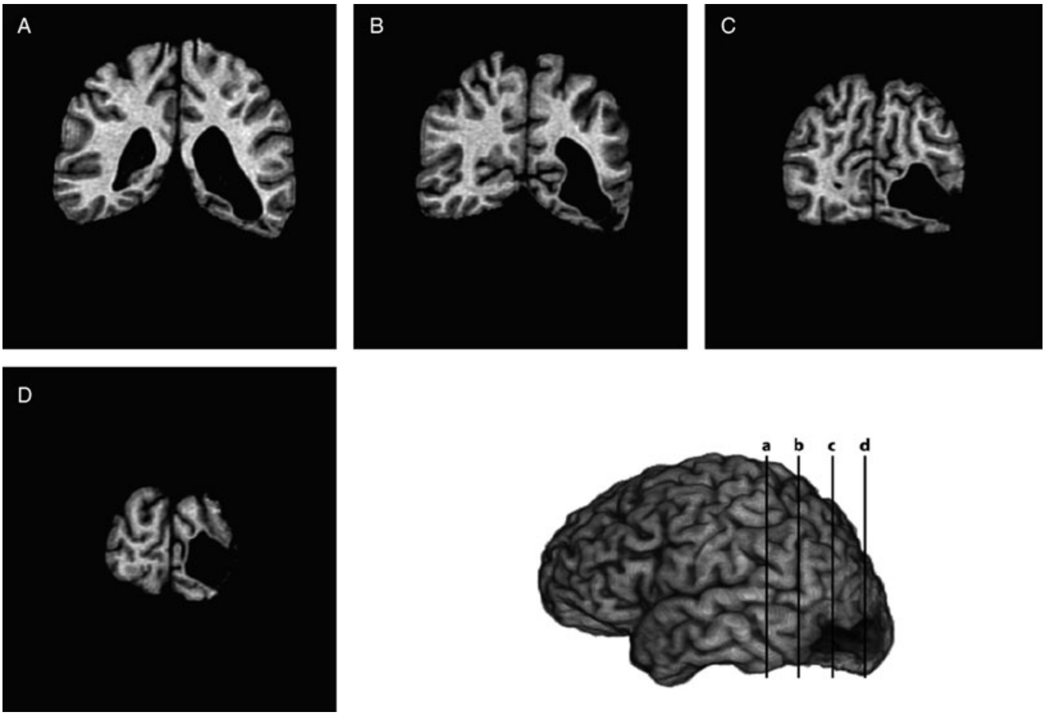


FIGURE 14.4. Example of a patient with a left ventrolateral occipital lesion (which includes the MT sector) and impaired naming of both static pictures of actions and dynamic videoclips of actions. The lesion is shown in a magnetic resonance scan obtained in the chronic epoch, reconstructed in standard brain space. The left hemisphere lateral perspective is shown, along with coronal cuts (black lines, a–d) that depict the cortical and white matter extent of the lesion in the ventrolateral occipital region. The left hemisphere is on the right in the coronal sections.

(Adapted from Tranel et al., 2008.)

studies had damage extending beyond MT. Further research is clearly needed to investigate these issues more carefully (e.g., Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008), but transcending the empirical minutiae is the overarching discovery that the visual motion patterns encoded by verbs do appear to be captured by conjunctive neurons in posterolateral temporal CZs that are closely related to those that subserve perception of the same properties. So, to return briefly to the purely behavioral study by Zwaan et al. (2004) described in the section on The Simulation Framework, it seems likely that when dynamic, egocentrically anchored simulations of motion are triggered by sentences such as *The shortstop hurled the softball at you* or *You hurled the softball at the shortstop*, those simulations are neurally instantiated in a sector of the posterolateral temporal cortex that resides near the region that underlies actual motion perception (see also Kaschak et al., 2005).

Although this chapter is concerned primarily with how words capture visual experience, at this juncture it would be remiss not to mention that one of the most fascinating recent developments in our understanding of the neural substrates of verb meaning involves motoric rather than visual information. Within the frontal lobes, both the primary motor cortex and the adjacent premotor cortex are somatotopically organized, which is to say that they contain maps of the muscular layout of the body—for instance, in the primary motor cortex the lips/tongue are represented in the ventrolateral sector, the arms/hands are represented in the dorsolateral sector, and the legs/feet are represented in the dorsomedial sector.[7] Single-cell recording studies with macaque monkeys, as well as studies using a variety of methods with humans, have revealed that body-part-specific motor areas are activated not only when certain types of actions are executed, but also when they are seen or heard (e.g., Buccino, Binkofski, Fink, Fadiga, Fogassi, Gallese, Seitz, Zilles, Rizzolatti, & Freund, 2001; Calvo-Merino, Grezes, Glaser, Passingham, & Haggard, 2006; Gazzola, Aziz Zadeh, & Keysers, 2006; Caetano, Jousmäki, & Hari, 2007; Filimon, Nelson, Hagler, & Sereno, 2007; Saygin, 2007; Van Schie, Koelewijn, Jensen, Oostenveld, Maris, & Bekkering, 2008; Serino, De Filippo, Casavecchia, Coccia, Shiffrar, & Ladavas, in press). Most importantly in the present context, there is increasing evidence that the motoric information encoded by action verbs is also represented by motor-related structures in the frontal lobes. In particular, several studies employing a wide range of brain mapping techniques have shown that verbs encoding face actions (e.g., *lick*), arm/hand actions (e.g., *pick*), and leg/foot actions (e.g., *kick*) differentially engage the corresponding somatotopically mapped primary motor and premotor regions (for reviews of a rapidly growing literature see Pulvermüller, 2008; Hauk, Shtyrov, & Pulvermüller, 2008b; Kemmerer & Gonzalez Castillo, 2009; Kemmerer, forthcoming; for a review of pertinent behavioral data see Fischer & Zwaan, 2008). These findings support the provocative notion that the motoric aspects of the meanings of action verbs—that is, what we might call the semantics of kinematics—are not part of an amodal symbolic representation in the brain, but are instead subserved by frontal cortical structures that overlap at least partly with those underlying the execution, imagination, observation, and audition of actions. Further evidence for this view comes from studies indicating that damage in the left primary and premotor areas often disrupts knowledge of the meanings of action verbs (e.g., Kemmerer & Tranel, 2003; Bak, Yancopoulou, Nestor, Xuereb, Spillantini, Pulvermüller, F., & Hodges, 2006; Grossman, Anderson, Khan, Avants, Elman, & McCluskey, 2008; Kemmerer, Rudrauf, Manzel, & Tranel, submitted; see also Tranel et al., 2003). However, as yet no neuropsychological studies have directly and systematically tested the prediction that the meanings of verbs for lip/tongue, arm/hand, and leg/foot actions should be differentially impaired by lesions affecting the relevant somatotopically mapped motor regions. This is a topic that clearly warrants further research (for a preliminary inquiry see Kemmerer & Tranel, 2000b).

Words for Spatial Relations Finally, speakers often use locative prepositions to describe the static spatial arrangements of objects in the visual field. During the past few decades, the literature in linguistics on the meanings of locative prepositions has expanded dramatically, leading from the classic work of Talmy (1983) and Herskovits (1986) to a variety of new approaches, such as those of Evans (2003) and Conventry and Garrod (2004) (for a broad survey of recent proposals, see Evans & Chilton, 2009). In striking contrast, very little research in cognitive neuroscience has sought to identify the brain structures that mediate this complex and intriguing conceptual domain. In what follows, I begin by summarizing several aspects of the meanings of locative prepositions, and then I briefly review a few neuropsychological and functional neuroimaging studies that have focused on these types of words. As I will show, these studies suggest that the cortical areas underlying the linguistic encoding of spatial relations are close to, but nevertheless distinct from, those underlying the representation of spatial relations for purely perceptual purposes.

The types of spatial scenarios designated by locative prepositions generally involve two objects—the "figure" (F), which is the thing to be located, and the "ground" (G), which is an object that serves as a point of reference. For example, in the sentence *The beer is in the refrigerator*, the noun-phrase *the beer* specifies F, the noun-phrase *the refrigerator* specifies G, and the preposition *in* specifies the nature of the spatial relationship between them. For the most part, locative prepositions express spatial relations in terms of very sketchy or schematic structural properties of the objects involved; metrical details are usually ignored, such as the exact sizes, shapes, and orientations of the objects, or the precise distances between them. For instance, in the example given above, the real-world situation that the sentence refers to might consist of a geometrically rich, idiosyncratic, three-dimensional spatial layout, perhaps involving a Corona longneck standing upright on the top shelf of a big refrigerator with many levels and compartments; yet the semantic structure of *in* is very austere and skeletal, as it abstracts away from these spatial particularities and instead treats the beer bottle as just a dimensionless point and the refrigerator as simply an idealized container. Hence, locative prepositions designate "categorical" as opposed to "coordinate" spatial relationships (Postma & Laeng, 2006).

Two major subclasses of locative prepositions are often distinguished—topological and projective. Topological prepositions refer to spatial relations that involve various types of "coincidence" of F and G. In addition to *in,* which expresses containment, other topological prepositions include *on* for contact, *around* for encirclement, and *through* for penetration. Some topological prepositions actually encode a combination of geometric and functional features—e.g., *on* expresses not only contiguity but also the force-dynamic notion of support (Coventry & Garrod, 2004). Projective prepositions, on the other hand, employ a different strategy of locational reference for other sorts of spatial arrays. They specify the location of F as being within a search-domain that is projected from one of the major facets or dimensional axes of G. Thus, *in front of* and *in back of* (or *behind*) designate relations of anteriority and posteriority with respect to the front/back axis of G, and *above* and *below* designate relations of superiority and inferiority with respect to the vertical axis of G. Determining the principal axes of G is not always straightforward, however, because it depends on the frame of reference that is adopted. Focusing on just the front/back distinction, many objects have what might be called an intrinsic front, which is based on factors such as the canonical direction of the sense organs (for people and animals), the canonical direction of motion (for vehicles), or the canonical direction of encounter (for TVs, computers, etc.). But some objects, such as trees, lack an intrinsic front, in which case an egocentric "orientation mirroring" frame of reference is adopted so that the front of the object is conceptualized as the side facing the observer.

Many locative prepositions appear to have a network of distinct but closely related meanings that are organized around a central or prototypical meaning. For instance, *in* ideally describes a spatial relation of containment in which G (1) is a three-dimensional object, (2) is hollow, and (3) completely encloses F. But each of these conditions can be violated, thereby yielding extended meanings as exemplified in the following situations: (1) a person standing inside a circle painted on the floor (G is a two-dimensional object), (2) a nail that has been pounded into a board (G is solid), and (3) an apple in a bowl even though it rests of top of other fruit so that it is technically above the horizontal upper edge of the bowl (F is not enclosed by G). Accounting for this remarkable flexibility of locative prepositions has proven to be a difficult challenge for semantic analysis.

Turning to the brain, Landau and Jackendoff (1993) speculated that the meanings of locative prepositions might be neurally instantiated in the left inferior parietal lobule, a brain region known to be involved in the perceptual representation of schematic, categorical spatial relations between objects (for reviews see Jager & Postma, 2003; Laeng, Chabris, & Kosslyn, 2003; Postma & Laeng, 2006). Several recent studies have not only corroborated this proposal but have made it more precise by suggesting that the critical cortical region may be the left supramarginal gyrus (SMG; for reviews see Kemmerer, 2006, 2009). Damasio, Grabowski, Tranel, Ponto, Hichwa, and Damasio (2001) report a 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:

first, naming F with an appropriate noun, and second, 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 SMG. They did not indicate which prepositions were targeted for production, but it appears that a mixture of topological and projective prepositions was included, which suggests that the SMG activation reflects semantic processing of both types (for convergent fMRI data see Noordzij, Neggers, Ramsey, & Postma, 2008, and Amorapanth, Widick, & Chatterjee, in press).

Additional evidence comes from a neuropsychological study conducted by Tranel and Kemmerer (2004; see also Kemmerer & Tranel, 2000a, 2003; Kemmerer, 2005). They administered a set of four standardized tests that collectively evaluate production, comprehension, and semantic analysis of 12 English prepositions (encoding topological relations as well as several kinds of projective relations) to 78 brain-damaged patients with lesions distributed throughout the left and right cerebral hemispheres, and then contrasted the lesion sites of the patients who were impaired on the tests with the lesion sites of those who were unimpaired. Poor performance was linked with damage in the left SMG and the left frontal operculum (see Fig. 14.5). The involvement of the left SMG strengthens the hypothesis that this region plays an essential role in representing the meanings of locative prepositions (for convergent neuropsychological data see Wu Waller, & Chatterjee, 2007, and Amorapanth et al., in press). The investigators did not, however, conduct separate analyses to determine whether the different semantic classes of prepositions dissociated from each other behaviorally and neuroanatomically, nor did they investigate whether prototypicality influenced the results. 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, Weiss, Mohlberg, Pieperhoff, Eickhoff, Gurd, Marshall, Shah, Fink, & Zilles, 2004), and semantic working memory, possibly in Brodmann area 45 (e.g., Badre & Wagner, 2007).

In a follow-up experiment with just those patients who failed all four preposition tests ($n = 6$), Tranel and Kemmerer (2004) assessed nonlinguistic visuospatial processing by administering a large battery of standardized neuropsychological tests (Benton & Tranel, 1993). Although a few of the tests emphasize sensitivity to metrically precise coordinate spatial relations (e.g., subtle variations in line orientation), the majority of them require an appreciation of more schematic categorical spatial relations (e.g., the typical arrangements of the parts of complex objects). Overall, the patients performed extremely well on the various tests, revealing a strong dissociation between impaired linguistic and preserved perceptual processing of spatial relations. Moreover, Kemmerer and Tranel (2000a) describe a patient who manifested a dissociation that was the opposite of the kind manifested by the patients in the study of Tranel and Kemmerer (2004)—namely, intact knowledge of the meanings of locative prepositions but impaired nonlinguistic visuospatial processing of both coordinate and categorical spatial relations.

This neuropsychological double dissociation constitutes evidence that the spatial image schemas expressed by locative prepositions are separate from those that are required to execute certain kinds of visuospatial tasks (for convergent behavioral data see Munnich, Landau, & Dosher, 2001; see also Kemmerer, 1999). In terms of CZ theory, these two types of spatial representations appear to be mediated by distinct populations of conjunctive neurons in anatomically close but, importantly, nonoverlapping portions of the left inferior parietal lobule, especially the SMG. Does this pose a challenge to the Simulation Framework? Not necessarily. For one thing, even though the meanings of locative prepositions do not seem to be needed to perform certain types of perceptual tasks, it is still reasonable to assume that they are perceptually grounded. After all, they derive in large part from visual experience, and it is quite possible that they employ the relatively high-level

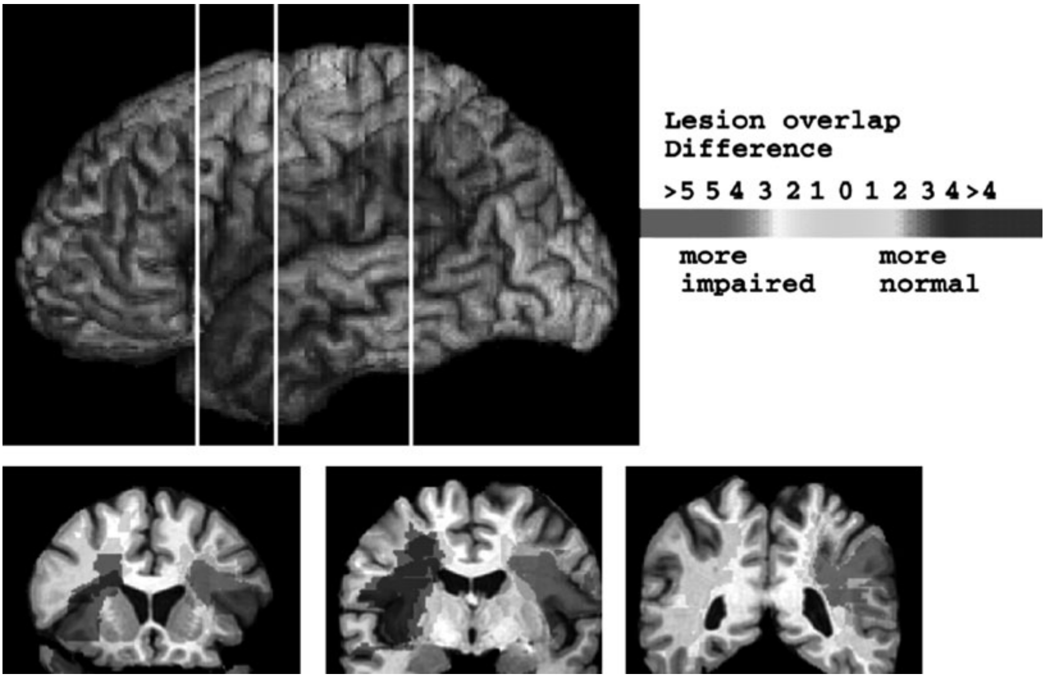


FIGURE 14.5. Grayscale rendition of Color Plate 10. See Color Plate 10 for interpretation. Results from a lesion study in which 78 brain-damaged subjects performed the Matching Test, which assesses knowledge of English locative prepositions. For each test item ($n = 50$), the subject is shown three pictures of objects in various spatial relationships and is asked which picture best represents the meaning of a particular preposition. For example, in one item the preposition is *in* and the three pictures show (1) one window above another window on the outside of a house, (2) eggs in a carton, and (3) a boy on a swing. The figure shows the subtraction of lesion overlaps for 15 unimpaired subjects from the lesion overlaps for 15 impaired subjects. The color bar indicates the number of lesions in the overlap difference (the difference reached as high as 7 in the red-coded zone for "more impaired"). The top panel shows a lateral view of the left hemisphere and reveals that, relative to unimpaired subjects, impaired subjects more often had lesions in the left frontal operculum extending posteriorly into the left inferior parietal lobule, specifically the supramarginal gyrus. The vertical white lines in the top panel denote the planes of the coronal sections depicted in the bottom three panels. The leftmost line, and corresponding leftmost bottom panel, indicates the plane of the frontal operculum, and the rightmost line, and corresponding rightmost bottom panel, indicates the plane of the supramarginal gyrus.
(Adapted from Tranel & Kemmerer, 2004.)

representational resources of this particular modality to capture various austere categories of spatial relations such as containment (*in*), contact (*on*), and penetration (*through*). Pursuing this line of thinking a step further, it may even be the case that the transient activation of prepositional meanings during online language processing involves, at least to some extent, embodied simulations of skeletal spatial arrays, most likely modulated by contextual factors. This is clearly a strong prediction of the Simulation Framework (e.g., see many of the chapters in Hampe, 2005), but I am not aware of any experimental studies that have attempted to test it. If such simulations do occur, however, their content is probably very language-specific in content; that is, it probably reflects the construal of space for communicative purposes much more than for purely perceptual purposes. Gentner and Boroditsky (2001) make essentially the same point when they state that spatial relational concepts are among the least likely to be "given by the world" and hence tend to exhibit strong language-specific influences. Support for this view comes not only from the neuropsychological double dissociation reported by Tranel and Kemmerer (2004), but also from the discovery of extensive crosslinguistic variation in the semantic domain of spatial relations, as discussed below.

Summary

In recent years, a rapidly growing body of research in cognitive neuroscience has been generating evidence for the view that word meanings are not represented in a single amodal brain region, similar to entries in a dictionary. Instead, they appear to be stored in multiple, anatomically distributed, modality-specific, and, even more narrowly, property-specific cortical areas, and their ephemeral activation during online language processing seems to involve the automatic, implicit recapitulation of sensorimotor states. I have reviewed a number of studies that focus on the neural substrates of the visual semantic components of words for objects, events, and spatial relations. Overall, these studies suggest that different visual semantic components are housed in cortical areas that are closely related to those that underlie the perception of the relevant properties. In particular, for the color, shape, and motion components of word meanings, there is mounting evidence that the online activation of the pertinent information takes the form of covert, property-specific, perceptual simulations. Such simulations may also occur for the types of spatial image schemas that are encoded by locative prepositions, but this still awaits careful investigation.

SOME QUESTIONS RAISED BY CROSS-LINGUISTIC VARIATION

The previous section demonstrated that cognitive neuroscience is starting to contribute, in its own unique way, to our understanding of how words capture visual experience. So far, however, this approach has been applied almost exclusively to English. In fact, as someone who was originally trained in linguistic typology before migrating to the field of cognitive neuroscience, I am acutely aware of the fact that crosslinguistic variation in lexical-semantic systems—variation of the sort that is discussed in many of the other chapters in this volume—has been almost entirely ignored by the majority of researchers who are currently investigating the organization, representation, and processing of conceptual knowledge in the brain (for an attempt at consciousness raising see Kemmerer, 2006). In this section I present thumbnail sketches of just a few forms of variation involving the lexical encoding of three types of visual information—shape, motion, and spatial relations—and for each one I consider some questions that arise concerning the underlying neural substrates.[8]

Shape

As noted earlier, Landau and Jackendoff (1993) maintain that the vocabulary of the average adult English speaker contains roughly 10,000 count nouns that designate types of objects that are bounded and individuated primarily according to inherent shape. At the same time, however, English speakers also make regular use of a large number of mass nouns that designate types of substances with no inherent shape. To describe a particular instance of a certain type of substance, speakers often combine the appropriate mass noun with one of several possible terms that provide rough geometric information about the form that the substance happens to take on that occasion—e.g., *a sheet of paper, a stack of paper, a pile of paper, a wad of paper*.

Some languages in the world have predominantly mass nouns together with rich systems of nominal classification that are much more elaborate than what is found in English and other Indo-European languages (e.g., Senft, 2000; Aikhenvald, 2003). For example, Lucy (1992) has argued that in Yucatec Maya (Mayan, Mexico) the noun *lo'bal* could be glossed "banana stuff" because, depending on the shape-discriminating classifier with which it co-occurs, it can be used to refer to a single banana, a bunch of bananas, a bit of the fruit, the leaf of the plant/tree, or the whole plant/tree:

'un-tz'iit	*lo'bal*	"one one-dimensional banana (i.e., the fruit)"
'un-kuuch	*lo'bal*	"one load banana (i.e., the bunch)"
'um-p'iit	*lo'bal*	"one bit banana (i.e., a bit of the fruit)"
'un-waal	*lo'bal*	"one two-dimensional banana (i.e., the leaf)"
'un-kuul	*lo'bal*	"one planted banana (i.e., the plant/tree)"

Likewise, a single classifier in Yucatec Maya can impose the same kind of individuation on a variety of different materials, "in much the way that a cookie-cutter cuts up undifferentiated dough," as Levinson (1996, p. 185) observed in a discussion of Lucy's (1992) study:

'un-tz'iit	*lo'bal*	"one one-dimensional banana (i.e., the fruit)"
'un-tz'iit	*kib'*	"one one-dimensional wax (i.e., a candle)"
'un-tz'iit	*che'*	"one one-dimensional wood (i.e., a stick)"
'un-tz'iit	*nal*	"one one-dimensional corn (i.e., an ear)"

Crosslinguistically, the semantic features of shape that are most commonly expressed by classifiers are the one-dimensional (1D) long shape, the two-dimensional (2D) flat shape, and the three-dimensional (3D) round shape. In some languages, however, finer features are encoded. A good illustration is Japanese, which, according to Sanches (1977), has the following system for classifying entities according to predominant dimension: 1D (-*hon*); 2D with length predominating (-*suji*); 2D with length and breadth equally important (-*mai*); 2D with height and breadth equally important (-*men*); 3D with length and breadth predominating (-*hen*); 3D cubic (-*cho*); 3D irregularly shaped (-*kai*); and 3D spherical (-*ko*). Based in part on an analysis of how Japanese children acquire this complex system, Inoue (2000) proposes that classifier choice is highly dependent on what she calls "visualizing ability"—a notion that accords nicely with the Simulation Framework. As she puts it, "speakers have to imagine just how round or tall an object-noun can be in order for the referent to still be associated with the numeral classifier *hon* ('cylindrical object'). If a pen is a prototypical instance of a referent for the numeral classifier *hon*, how about a 10-meter telephone pole? Or a coffee mug?" (Inoue, 2000, p. 219).

A small but growing body of literature in cognitive neuroscience has begun to address both semantic and syntactic aspects of the count/mass distinction in English and in a few other Indo-European languages, especially Italian (for a short summary, see Semenza, 2005; for studies investigating specifically the semantic aspects of the distinction, see Bisiacchi, Mondini, Angrilli, Marinelli, & Semenza, 2005; Taler, Jerema, & Saumier 2005; Crutch & Warrington, 2007). But as yet no research has explored the neural substrates of the kinds of shape classifiers mentioned above. Are the highly schematic meanings of such classifiers captured by conjunctive neurons in the ventral temporal cortex, and if so, where exactly do they reside? Could future research in cognitive neuroscience help explain why the crosslinguistically most prevalent shape classifiers are those that specify 1D, 2D, and 3D forms? How does the cortical representation of *lo'bal* in Yucatec Maya differ from that of *banana* in English, especially with regard to the regions of the fusiform gyrus that store the shape properties of objects? For speakers of Yucatec Maya, does *lo'bal* automatically trigger perceptual simulations of not only single bananas, but also bunches of bananas, the leaf of the plant/tree, or even the whole plant/tree? These are only a handful of the many intriguing questions that classifier languages pose for cognitive neuroscience in general, and for the Simulation Framework and CZ theory in particular.

Motion

As already indicated, English has a well-developed inventory of "manner of locomotion" verbs that encode subtle semantic distinctions along multiple dimensions such as visual pattern, motor pattern, rate, and social-emotional evaluation. But this inventory differs significantly from the inventories in other languages. Even English and German, which are closely related languages, differ somewhat in this domain—e.g., German has no exact equivalents to English *scamper, scurry, scuttle,* and *scramble,* and English has no exact equivalents to German *stapfen, stiefen, trampeln,* and *stampfen,* which designate different kinds of firm, heavy walking (Snell-Hornby, 1983). These differences are quite minor, however, compared to the much more substantial variation that has been documented worldwide (Levinson & Wilkins, 2006; but see also Malt, Gennari, Imai, Ameel, Tsuda, & Majid, 2008). Building on a foundation of previous work by a number of typologists, Slobin (2004) reports extensive crosslinguistic diversity in how motion events are described, and he attempts to account for this diversity by positing a "cline of manner salience." In high-manner-salient languages (e.g., Germanic and Slavic languages, Hungarian, and Mandarin), there is an easily accessible morphosyntactic slot for expressing manner (such as the main verb position in English), and this encourages speakers to attend to, and eventually lexicalize, increasingly finer manner details, which in turn causes a rich lexicon of manner morphemes to arise diachronically and to even influence co-speech gesture (Kita & Özyürek, 2003; see also Kemmerer, Chandrasekaran, & Tranel, 2007). At the other end of the continuum, in low-manner-salient languages (e.g., Romance languages, Turkish, and Hebrew), manner information is grammatically subordinated to path information, so the former is provided only when it must be foregrounded for some reason, and there is less motivation to create a large inventory of manner morphemes. The tremendous difference between the expressive resources of languages at the two extremes can be seen by contrasting English with both Spanish and French. For example, *creep, glide, slide, slip,* and *slither*[9] are all translated into Spanish as *escabullirse,* and *bound, hop, jump, leap,* and *spring* are all translated into French as *bondir.*

Shifting to the brain, it is interesting to consider that CZ theory leads to the following hypothesis regarding the neural substrates of the visual component of locomotion verbs. Perhaps the crosslinguistic diversity in the lexicalization of manner of locomotion is reflected, at least in part, in corresponding neurobiological diversity in the spatial arrangements and "tuning curves" of conjunctive neurons in CZs within the mosaic of cortical areas extending anteriorly from area MT into the posterolateral temporal cortex, including the region that responded preferentially to Running verbs in the fMRI study of Kemmerer et al. (2008) (see Fig. 14.3). According to this hypothesis—more precisely, according to the Similarity in Topography principle—the layout of the relevant conjunctive neurons is systematically different for English speakers compared to, say, Spanish speakers. For English speakers there are separate but tightly clustered conjunctive neurons for the closely related visual motion patterns encoded by *creep, glide, slide, slip,* and *slither;* however, for Spanish speakers such conjunctive neurons do not exist because (1) the Spanish manner verb lexicon does not make any of those subtle semantic distinctions (the whole spectrum is covered by just one verb, *escabullirse*), and (2) there is no independent reason to expect those particular distinctions to be "natural" in the sense of being universally employed in the nonverbal categorization of motion events (see Slobin, 2000, 2003, for relevant data and discussion from the perspective of language acquisition). As the spatial resolution of brain mapping techniques continues to improve (e.g., Mur, Bandettini, & Kriegeskorte, 2009), it may eventually become feasible to test hypotheses of this nature, thereby shedding further light on the biological bases of the meanings of motion verbs. For present purposes, the essential point is this: It may not be a coincidence that prominent theorists in both linguistic typology (e.g., Croft, 2001; Haspelmath, 2003) and cognitive neuroscience (e.g., Simmons & Barsalou, 2003; Crutch & Warrington, 2003) increasingly use the mapping metaphor (i.e., analyzing semantic domains in terms of geometric spaces) in their characterizations of the organization of conceptual knowledge. Perhaps the metaphor is more appropriate than we have hitherto realized (for reviews of pertinent neurocomputational modeling, see Kohonen & Hari, 1999; Graziano & Aflalo, 2007b).

Spatial Relations

Even greater cross-linguistic variation has been documented in the domain of spatial relations than in the domains of physical shape and manner of locomotion, perhaps because, as already noted, spatial relational concepts are less likely to be "given by the world" (Gentner & Boroditsky, 2001). Several recent books provide in-depth descriptions of the remarkable diversity, as well as the many overarching commonalities, of the categorical spatial coding systems manifested in carefully selected samples of the 6000+ languages of the world (Levinson, 2003b; Levinson & Wilkins, 2006). Here I mention just a few forms of variation involving topological and projective relations, and I briefly introduce some questions and hypotheses about how these types of variation might be implemented in the brain (for a deeper discussion of these issues, see Kemmerer, 2006, 2009).

In the topological realm, a great deal of crosslinguistic diversity has been documented, but several underlying patterns have also emerged. In a study reported by Levinson and Meira (2003), nine unrelated languages[10] were investigated by comparing native speaker responses to a standardized set of 71 pictures showing a wide range of topological relations (for more recent developments see Levinson & Wilkins, 2006). The 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 were uncovered. 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, as 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 finer 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, as *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 (2003) also report many intriguing cases of category fractionation in other languages, such as the rather 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).

The neuropsychological and functional neuroimaging studies summarized in the section on Words for Spatial Relations suggest that the left inferior parietal lobule—in particular, the SMG—is a key cortical region for representing the meanings of English locative prepositions, including those that encode topological spatial 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 & Wilkins, 2006). By conducting high-resolution brain mapping studies with such materials, it might 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 inferior parietal lobule, 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 neuronal ensembles, in accord with the Similarity in Topography Principle.

Turning to the realm of projective spatial relations, I indicated earlier that when English speakers need to say that F (e.g., a ball) is located within a region that is projected from a G without an intrinsic front or back (e.g., a tree), they use an egocentric "orientation mirroring" frame of reference, so that the front of G is construed as the side facing the observer. Thus, a statement such as *The ball is in front of the tree* means that the ball is between the tree and the observer. However, in some languages, such as Hausa (West Chadic, Niger), an egocentric "orientation-preserving" frame of reference is employed instead, so that the observer's front-back bodily axis is mapped onto G without any rotation. In this type of system, a statement such as *The ball is in front of the tree* means that the ball is on the side of the tree opposite to the observer. Yet another strategy is to avoid the egocentric frame of reference entirely by rejecting the assumption that trees lack intrinsic fronts and backs. This approach is taken by Chamus (Nilo-Saharan, Kenya), which treats the front of a tree as the side it leans toward, or, if it is vertical, the side with the biggest branch or the most branches. An even more radical departure from English can be found in languages that specify all projective relations on the horizontal plane in terms of an absolute frame of reference that provides a set of fixed bearings or cardinal directions. In systems of this sort, we might describe the spatial array involving the ball and tree by saying *The ball is north of the tree*, a statement that is completely independent of both the observer's perspective and the intrinsic facets of the tree. Absolute systems are fundamentally geocentric, and languages often base terms for cardinal directions on stable environmental features such as mountain slopes, river drainages, and prevailing wind patterns. For example, Tzeltal (Mayan, Mexico) has an absolute system that is anchored in the mountain incline of the local landscape, giving rise to three directional terms: *alan* "downhill" (roughly north), *ajk'ol* "uphill" (roughly south), and *jejch* "across" (either east or west) (Brown & Levinson, 1993; Brown, 2006). Although the terminology of absolute systems typically derives from environmental landmarks, such systems are fully abstracted (at least in some languages[11]), 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 truly remarkable ability, as demonstrated by Levinson (2003b). Here is just one of many eye-opening examples and analyses that he presents throughout his book, this particular one focusing on Guugu Yimithirr (GY) (Pama-Nyungan, Australia), whose speakers use exclusively the absolute frame of reference for characterizing horizontal projective relations (Levinson, 2003b, p. 114):

> 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).

Languages such as this constitute an excellent example of how semantic typology can inspire future research on the neural representation of categorical spatial relations. As noted, the speakers of such languages must constantly compute their 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 retinas 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, Thalu, Phillips, Wiltschko, & Wiltschko, 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, Benhamoou, Girard, Ciccione, Roos, Sudre, & Benvenuti, 2007), and locusts perceive polarization patterns in the blue sky and use them as cues for spatial orientation (Heinze & Homberg, 2007). But for people in "absolute" communities, the mental compass that generates their superb sense of direction—a sense that may be comparable in accuracy to that of homing pigeons (Levinson, 2003b, 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, 2003b, p. 278; see also Haun, Rapold, Call, Janzen, & Levinson, 2006). It is reasonable to suppose that the relevant brain areas include parietal and hippocampal structures that have been implicated in both constructing landmark-based cognitive maps of the environment and monitoring our perspective on and movement through them (e.g., Ekstrom, Kahana, Caplan, Fields, Isham, Newman, & Fried, 2003; Hartley, Maguire, Spiers, & Burgess, 2003; Janzen & van Turennout, 2004; Hafting, Fyhn, Molden, Moser, & Moser, 2005; Spiers & Maguire, 2006; Leutgeb, Leutgeb, Moser, & Moser, 2007). However, because the use of the mental compass does not necessarily require input from visually perceived landmarks, 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 (Sargolini, Fyhn, Hafting, McNaughton, Witter, Moser, & Moser, 2006). To be sure, I would hardly expect my fellow brain scientists to attempt to set up an fMRI scanner in Hopevale, North Queensland, Australia in order to gather functional neuroimaging data about the biological bases of spatial description in Guugu Yimithirr. But I do hope that in the coming years an increasing number of these researchers will develop a genuine appreciation of the neuroscientific questions raised by languages such as this, and that ultimately such questions will somehow be addressed.

Summary

Recent research in the nascent field of semantic typology has begun to reveal considerable crosslinguistic variation in numerous domains of word meaning. Studies along these lines point to what are sometimes called language-specific semantic maps across universal conceptual spaces. I have briefly surveyed some of the major forms of diversity in the visual semantic fields of shape, motion, and spatial relations. Crosslinguistic variation of this nature has already had a significant influence on developmental psychologists who investigate the acquisition of language during childhood (e.g., Bowerman & Levinson, 2001; Guo, Lieven, Budwig, Ervin-Tripp, Nakamura, &

Ozcaliksan, 2008) as well as on cognitive scientists who investigate the relation between language and thought (e.g., Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2005); however, it has largely been ignored by cognitive neuroscientists who investigate the organization, representation, and processing of conceptual knowledge in the brain. This is unfortunate, because typological data raise many interesting questions about the neural substrates of word meaning—questions that go beyond the idiosyncrasies of English to embrace much richer patterns of similarities and differences among the 6000+ languages of the world.

CONCLUSIONS

In this chapter I have discussed the issue of how words capture experience from a cognitive neuroscience perspective. I have deliberately concentrated on the lexical encoding of the visual world, because that is where the most progress has been made with regard to identifying the underlying brain structures. Although the neuropsychological and functional neuroimaging studies that I have reviewed are restricted entirely to English and are therefore quite parochial from the point of view of semantic typology, they are nevertheless broadly consistent with the Simulation Framework insofar as they suggest that different kinds of visual semantic components of words depend on cortical regions that are closely related to those that process the same properties during perception. I have not discussed any studies involving semantic representations in nonvisual sensory modalities, but it is worth noting that they provide further evidence for the Simulation Framework (e.g., Kellenbach, Brett, & Patterson, 2001; James & Gauthier, 2003; de Araujo, Rolls, Velazco, Margot, & Cayeux, 2005; Simmons, Martin, & Barsalou, 2005; Goldberg, Perfetti, & Schneider, 2006a,b; González, Barros-Loscertales, Pulvermüller, Meseguer, Sanjuán, Belloch, & Ávila, 2006; Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008; Hoenig, Sim, Bochev, Herrnberger, & Kiefer, 2008; Foroni & Semin, 2009; and Hwang, Palmer, Basho, Zadra, & Müller, 2009). For instance, Goldberg, Perfetti, and Schneider (2006a) found that making semantic decisions about visual, auditory, tactile, and gustatory aspects of word meaning activated the corresponding sensory brain regions.

In a recent summary of the Simulation Framework, Zwaan and Madden (2005, p. 242) point out that this theory will seem trivial to some people yet counterintuitive to others: "[It] will seem trivial to the lay person, or even to people with great expertise in the use of language, such as novelists and poets. Of course words can be used to conjure up images in the reader's [or listener's] mind! However, these same claims will seem counterintuitive to researchers trained in traditional cognitive science. To them, the claim that meaning can be captured by experiential representations does not make sense." I submit that this is because the core assumption of the traditional view is that semantic structures are *by definition* entirely amodal in character. It is precisely this assumption—it might even be called an axiom—that advocates of the Simulation Framework are attempting to challenge through careful theoretical argumentation and an increasingly sophisticated and compelling progression of psychological and neuroscientific investigations.

Many questions remain unresolved, however, especially regarding the true nature and function of lexically triggered simulations of sensorimotor states. In my opinion, one of the top priorities for future research should be to clarify, both theoretically and empirically, the distinctions between the following three types of modality-specific information processing. First, there are *conscious perceptual experiences*, as when we open our eyes and see, for example, a dog. These experiences are subserved by an array of neural mechanisms, beginning in the retina and proceeding all the way up to, and beyond, the highest levels of the visual system. Second, there are *explicit perceptual simulations*, as when we close our eyes and voluntarily, effortfully imagine a dog. Although the topic of mental imagery has a long and controversial history, the weight of evidence strongly favors the view that the top-down generation of explicit visual images shares a great deal, but certainly not all, of the cortical territory underlying bottom-up perception (for a comprehensive review see Kosslyn, Thompson, & Ganis, 2006). Third, there are *implicit perceptual simulations*, as when we hear or read the word *dog*, and the word automatically, reflexively triggers an unconscious visual representation of a dog. According to the Simulation Framework, these types of simulations constitute substantial portions of the meanings of words. However, relatively little attention has been devoted to delineating the precise ways in which they are similar to and different from the types of representations that are employed in veridical perception and explicit imagery. As I emphasized in the review of experimental studies in the section on Empirical Findings, cognitive neuroscientists have frequently found that when subjects perform semantic tasks involving, say, the shape properties of objects encoded by concrete count nouns, there is activation in some of the same cortical areas that are engaged during both the perception and the imagination of those properties. And yet it is patently obvious that understanding a word such as *dog* is not at all the same as passively seeing a dog in the world, nor is it the same as voluntarily evoking an explicit image of a dog in the mind's eye. A few cognitive neuroscientists, such as Alex Martin and his colleagues, have recently begun to use advanced functional imaging methods, such as repetition suppression and priming paradigms, to tease apart these three kinds of modality-specific information processing, with the ultimate goal of isolating and properly characterizing the neural substrates as well as the cognitive content of the visual semantic components of words. This line of research is still in its infancy, however, and most of the work still lies ahead.

The usual goal of communication is, of course, to set up "the same thought" in the receiver's brain as is currently taking place in the sender's brain. In this context, words can be regarded as instructions for running embodied simulations. As noted, however, these simulations are generally implicit, which is to say, covert, unconscious, and automatic. Although we can, if we wish, bring these sensorimotor reenactments into the light of awareness, our brains usually hide them from us, presumably so we can consciously attend to other things, such as how the message fits into our belief system, whether the speaker is being sincere, how we intend to respond, and so on (Frith, 2007). The implicit nature of lexically triggered simulations is actually quite consistent with other recent discoveries about the fundamental role that embodied simulations play in our mental lives, especially when it comes to reflexively understanding each other's actions, emotions, and sensations. The basic idea is simple: By virtue of having common brain circuits and common sensorimotor and affective experiences, people can, so to speak, automatically translate the sights and sounds of what other individuals do and feel into the language of their own actions and feelings (for reviews see, e.g., Gallese, Keysers, & Rizzolatti, 2004; Decety & Lamm, 2006; Keysers & Gazzola, 2006; Grafton, 2009). Among the many forms of evidence for this view are the following findings. First, as I pointed out at the end of the section on Words for Events, the observation of an action engages some of the same somatotopically mapped neural networks that are active during its execution (e.g., Buccino et al., 2001; Filimon et al., 2007). This kind of motor resonance is strongest for goal-directed actions that the observer is skilled at performing (e.g., Calvo-Merino et al., 2006; Cross, Hamilton, & Grafton, 2006), but it can also be triggered by (1) "degraded" actions that are perceived only as point-light displays (e.g., Saygin, Wilson, Hagler, Bates, & Sereno, 2004; Saygin, 2007), (2) actions that are merely implied by static pictures (e.g., Longcamp, Tanskanen, & Hari, 2006; Urgesi et al., 2006), and (3) actions that are heard but not seen (e.g., Gazzola et al., 2006; Caetano et al., 2007). Second, deciphering the intentions of observed actions—e.g., whether a person who is grasping a cup intends to drink from it or clean it—also evokes motor simulations (e.g., Iacoboni, Molnar-Szakacs, Buccino, Mazziotta, & Rizzolatti, 2005; Hamilton & Grafton, 2006; but see also Brass, Schmitt, Spengler, & Gergely, 2007). Third, recognizing and empathizing with other people's emotions involve covertly recapitulating the types of body states that generate them (e.g., Wicker, Keysers, Plailly, Royet, Gallese, & Rizzolatti, 2003; Avenanti, Bueti, Galati, & Aglioti, 2005; Harrison, Singer, Rotshtein, Dolan, & Critchley, 2006). Fourth, the observation of someone being touched on a particular part of their body induces activation in the somatosensory cortices of the viewer, as if the viewer's own body were the subject of tactile stimulation (e.g., Keysers, Wicker, Gazzola, Anton, Fogassi, & Gallese, 2005; Hasson, Nir, Levy, Fuhrmann, & Malach, 2004; Blakemore, Bristrow, Bird, Frith, & Ward, 2005). Taken together, these findings, together with a large and rapidly growing assortment of additional discoveries involving both normal and pathological cognition, provide strong support for the Simulation Framework. The point I wish to highlight, however, is that all of these types of sensorimotor reenactments are usually implicit in the same way that those underlying word comprehension are usually implicit. Overall, the available data suggest that the same neurocognitive mechanism—simulation—lies behind the entire range of phenomena.

At the same time, though, it is clear that words are special, as they reflect historically shaped, culturally shared conventions for conceptual coordination. Words are not merely coded instructions for covertly reenacting certain kinds of sensorimotor states, they are *language-specific* instructions for running such reenactments. The previous section barely scratched the surface of the tremendous diversity that linguists specializing in semantic typology have already found in the multifaceted domain of the lexicalization of visual experience. The fact that so much crosslinguistic diversity exists cannot be overemphasized because, as Levinson (2003a, p. 29) observes, "*that's the fundamentally interesting thing* about language from a comparative point of view. We are the only known species whose communication system is profoundly variable in both form and content. . . . So we can't have the same kind of theory for human communication that we have for bee or even monkey communication. . . ." Tomasello (2003, p. 1) makes essentially the same point when he states that one of the most bizarre traits of *Homo sapiens* is that "whereas the individuals of all nonhuman species can communicate effectively with all of their conspecifics, human beings can communicate effectively only with other persons who have grown up in the same linguistic community—typically, in the same geographic region." The significance of this basic fact has yet to be fully grasped by and absorbed into the branch of cognitive neuroscience that concentrates on the representation, organization, and processing of conceptual knowledge. Hence many intriguing questions await investigation. What is the relation between crosslinguistic variation in body part terms (Brown, 2005a,b; Enfield, Majid, & van Staden, 2006) and the mapping of the extrastriate body area (Peelen & Downing, 2007; Kemmerer & Tranel, 2008)? What is the relation between crosslinguistic variation in landscape terms (Burenhult & Levinson, 2008) and the mapping of the parahippocampal place area (Epstein & Kanwisher, 1998; Aziz-Zadeh, Fiedbach, Naranayan, Feldman, Dodge, & Ivry, 2008)? What is the relation between crosslinguistic variation in verbs of "cutting and breaking" (Majid, Boster, & Bowerman, 2008) and the mapping of the posterolateral temporal cortex (Blake & Shiffrar, 2007; Kemmerer et al., 2008)? My hope is that this chapter will help inspire more scientists to explore these and countless other questions about how words capture visual experience.

Notes

1. For an introduction to this line of thinking in the context of a textbook, see Smith and Kosslyn (2007, Ch. 4).
2. A qualifier: Independently of long-term memory, during passive perception, multisensory integration may occur at much earlier stages of processing, as suggested by Beauchamp (2005) and Ghazanfar and Schroeder (2006).
3. For useful figures that illustrate this theory, see Smith and Kosslyn (2007, pp. 166–168).
4. In the studies to be discussed, tools are usually restricted to small man-made objects that are manipulated in conventional ways to serve specific functions.

5. Hart and Kraut (2007a) propose that anatomically segregated feature representations are "glued" together during online semantic processing by means of neural firing rates that are synchronized with a 30-Hz gamma rhythm. This synchronization process may be mediated in part by the pulvinar nucleus in the thalamus.
6. The area uniquely associated with Speaking verbs may reflect not only the activation of lip/tongue-related visual motion representations—e.g., the distinctive appearances of *screaming* versus *whispering* (Allison et al., 2000; Pelphrey et al., 2005)—but also the activation of vocalization-related auditory representations—e.g., the distinctive sound patterns, especially loudness levels, of *screaming* versus *whispering* (Schirmer & Kotz, 2006).
7. To obtain a sense of the additional complexities, consider the recent finding that the lateral precentral gyrus of the macaque brain contains at least three separate hand representations (one in the primary motor cortex, a second in the ventral premotor cortex, and a third in the dorsal premotor cortex) whose topographic partitioning reflects multiple, conflicting mapping requirements, including the need to encode distinct repertoires of complex, ethologically relevant movements, such as climbing/leaping behaviors, reaching behaviors, hand-to-mouth behaviors, defensive behaviors, and central space/manipulation behaviors (for reviews see Graziano, 2006; Granziano & Aflalo, 2007a). Regarding the motor representation of hand actions in the lateral precentral gyrus of the human brain, yet another mapping requirement may involve the development of neurons tuned to the idiosyncratic types of hand actions that are expressed by verbs in languages.
8. Due to space limitations, I will not discuss a fourth type of visual information, namely color.
9. Another member of this family is *slink*, which is how Dr. Seuss describes the movement of the Grinch in his well-known children's story *How the Grinch Stole Christmas:* "Then he slithered and slunk, with a smile most unpleasant, around the whole room, and he took every present!" It is noteworthy that the word-initial consonant cluster *sl-* seems to have phonesthetic qualities, and that many languages actually have large inventories of ideophones that capture different manners of motion. Slobin (2004, p. 251) points out, for instance, that "Westermann's (1930) grammar of Ewe (Niger-Congo, West Africa) gives examples of 37 ideophones that can be used with the verb [for] 'walk,' with the additional information that these forms can be reduplicated and can occur with high tone for diminutives and low tone to describe motions of large entities."
10. 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), Yucatec Maya (Mayan, Mexico).
11. Determining whether the absolute system in a given language is fully abstracted from its geocentric anchor(s) requires careful fieldwork, as shown, for instance, by Schultze-Berndt's (2006) investigation of Jaminjung (Jaminjungan, Australia). This language has an absolute system based on water flow and verticality, but it "breaks down for reference beyond the drainage system which includes the territory the speakers are familiar with" (p. 105).

References

Aikhenvald, A. Y. (2003). *Classifiers.* Oxford: Oxford University Press.

Alford, J. L. van Donkelaar, P., Dassonville, P., & Marrocco, R. T. (2007). Transcranial magnetic stimulation over MT/MST fails to impair judgments of implied motion. *Cognitive, Affective, and Behavioral Neuroscience, 7,* 225–232.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4,* 267–278.

Amorapanth, P., Widick, P., & 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., & 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.

Assmus, A., Giessing, C., Weiss, P. H., & Fink, G. R. (2007). Functional interactions during the retrieval of conceptual action knowledge: An fMRI study. *Journal of Cognitive Neuroscience, 19,* 1004–1012.

Avenanti, A., Bueti, D., Galati, G., & Aglioti, S. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nature Neuroscience, 8,* 955–960.

Aziz-Zadeh, L., Fiebach, C. J., Naranayan, S., Feldman, J., Dodge, E., & Ivry, R. B. (2008). Modulation of the FFA and PPA by language related to faces and places. *Social Neuroscience, 3,* 229–238.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45,* 2883–2901.

Bak, T. H., Yancopoulou, D., Nestor, P. J., Xuereb, J. H., Spillantini, M. G., Pulvermüller, F., & Hodges, J. R. (2006). Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain, 129,* 321–332.

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18,* 513–562.

Barsalou, L. W. (2005). Situated conceptualization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 619–650). St. Louis: Elsevier.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Barsalou, L. W., & Hale, C. R. (1993). Components of conceptual representation: From feature lists to recursive frames. In I. Van Mechelen, J. Hampton, R. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 97–144). San Diego, CA: Academic Press.

Barsalou, L. W., Simmons, W. K., Barbey, A., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7,* 84–91.

Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: Theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language, 106,* 15–28.

Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipito-temporal cortex. *Current Opinion in Neurobiology, 15,* 145–153.

Beauchamp, M. S., Haxby, J. V., Jennings, J. E., & DeYoe, E. A. (1999). An fMRI version of the Farnsworth-Munsell 100-Hue Test reveals multiple color-selective areas in human ventral occipitotemporal cortex. *Cerebral Cortex, 9,* 257–263.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron, 34,* 149–159.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience, 15,* 991–1001.

Beauchamp, M. S., & Martin, A. (2007). Grounding object concepts in perception and action. *Cortex, 43,* 461–468.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *Journal of Neuroscience, 28,* 11347–11353.

Benton, A. L., & Tranel, D. (1993). Visuoperceptual, visuospatial, and visuoconstructive disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (3rd ed., pp. 165–213). New York: Oxford University Press.

Bergen, B. (2007). Experimental methods for simulation semantics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 277–301). Amsterdam: John Benjamins.

Binder, J.R., Desai, R.H., Graves, W.W., & Conant, L.L. (in press). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex.*

Bisiacchi, P., Mondini, S., Angrilli, A., Marinelli, K., & Semenza, C. (2005). Mass and count nouns show distinct EEG cortical processes during an explicit semantic task. *Brain and Language, 95,* 98–99.

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology, 58,* 47–73.

Blakemore, S.-J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision-touch synaesthesia. *Brain, 128,* 1571–1583.

Boronat, C., Buxbaum, L. J., Coslett, H. B., Tang, K., Saffran, E. M., Kimberg, D. Y., & Detre, J. A. (2005). Distinctions between function and manipulation knowledge of objects: Evidence from functional magnetic resonance imaging. *Cognitive Brain Research, 23,* 361–373.

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex, 16,* 183–191.

Bowerman, M., & Levinson, S. C. (Eds.). (2001). *Language acquisition and conceptual development.* Cambridge, UK: Cambridge University Press.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology, 17,* 2117–2121.

Bright, P., Moss, H. E., Longe, O., Stamatakis, E. A., & Tyler, L. K. (2007). Conceptual structure modulates anteromedial temporal involvement in processing verbally presented object properties. *Cerebral Cortex, 17*, 1066–1073.

Brown, C. H. (2005a). Finger and hand. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures* (pp. 526–529). Oxford: Oxford University Press.

Brown, C. H. (2005b). Hand and arm. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures* (pp. 522–525). Oxford: Oxford University Press.

Brown, P. (2006). A sketch of the grammar of space in Tzeltal. In S. C. Levinson & D. Wilkins (Eds.), *Grammars of space* (pp. 230–272). Cambridge, UK: Cambridge University Press.

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

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*, 400–404.

Burenhult, N., & Levinson, S. C. (2008). Language and landscape: A cross-linguistic perspective. *Language Sciences, 30*, 135–150.

Caetano, G., Jousmäki, V., & Hari, R. (2007). Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *Proceedings of the National Academy of Sciences, 104*, 9058–9062.

Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology, 16*, 1905–1910.

Caramazza, A., & Mahon, B. Z. (2006). The organization of conceptual knowledge in the brain: The future's past and some future directions. *Cognitive Neuropsychology, 23*, 13–38.

Chao, L. L., & Martin, A. (1999). Cortical representation of perception, naming, and knowing about color. *Journal of Cognitive Neuroscience, 11*, 25–35.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*, 913–919.

Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. *Cerebral Cortex, 12*, 545–551.

Conway, B. R., Moeller, S., & Tsao, D. Y. (2007). Specialized color modules in macaque extrastriate cortex. *Neuron, 56*, 560–573.

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., & Petersen, S.E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science, 248*, 1556–1559.

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

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press.

Cross, E. S., Hamilton, A. F., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage, 31*, 1257–1267.

Crutch, S. J., & Warrington, E. K. (2003). Spatial coding of semantic information: Knowledge of country and city names depends upon their geographical proximity. *Brain, 126*, 1821–1829.

Crutch, S. J., & Warrington, E. K. (2007). The semantic organization of mass nouns: Evidence from semantic refractory access dysphasia. *Cortex, 43*, 1031–1124.

Damasio, A. R. (1989a). Time-locked multiregional retroactivation: A systems level proposal for the neural substrates of recall and recognition. *Cognition, 33*, 25–62.

Damasio, A. R. (1989b). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1*, 123–132.

Damasio, A. R. (1989c). Concepts in the brain. *Mind and Language, 4*, 24–28.

Damasio, A. R., & Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: The convergence zone framework. In C. Koch & J. L. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 61–74). Cambridge, MA: MIT Press.

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

Damasio, H., Tranel, D., Grabowski, T. J., Adolphs, R., & Damasio, A. R. (2004). Neural systems behind word and concept retrieval. *Cognition, 92,* 179–229.

de Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C., & Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron, 46,* 671–679.

de Vega, M., Glenberg, A.M., & Graesser, A.C. (Eds.), *Symbols and embodiment: Debates on meaning and cognition.* Oxford: Oxford University Press.

Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The Scientific World Journal, 6,* 1146–1163.

Devlin, J. T., Rushworth, M. F. S., & Matthews, P. M. (2005). Category-related activation for written words in the posterior fusiform is task specific. *Neuropsychologia, 43,* 69–74.

Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R.L. (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences, 104,* 1097–1102.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness.* New York: Basic Books.

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

Enfield, N., Majid, A., & van Staden, M. (2006). Cross-linguistic categorisation of the body. *Language Sciences, 28,* 137–147.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local environment. *Nature, 392,* 598–601.

Epstein, R., DeYoe, E. A., Press, D. Z., Rosen, A. C., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsychology, 18,* 481–508.

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

Evans, V., & Chilton, P. (Eds.). (2009). *Language, cognition, and space: The state of the art and new directions.* London: Equinox.

Evans, V., & Green, M. (2006). *Cognitive linguistics.* Mahwah, NJ: Lawrence Erlbaum.

Farah, M. (2004). *Visual agnosia,* 2nd ed. Cambridge, MA: MIT Press.

Fawcett, I. P., Hillebrand, A., & Singh, K. D. (2007). The temporal sequence of evoked and induced cortical responses to implied-motion processing in human motion area V5/MT+. *European Journal of Neuroscience, 26,* 775–783.

Filimon, F., Nelson, J. D., Hagler, D. J., & Sereno, M. I. (2007). Human cortical representations for reaching: Mirror neurons for execution, observation, and imagery. *NeuroImage, 37,* 1315–1328.

Fischer, M., & Zwaan, R. A. (2008). Embodied language: A review of the role of the system in language comprehension. *Quarterly Journal of Experimental Psychology, 61,* 825–850.

Fodor, J. A. (1975). *The language of thought.* Cambridge, MA: Harvard University Press.

Foroni, F., & Semin, G. R. (2009). Language that puts you in touch with your bodily feelings: The multimodal responsiveness of affective expressions. *Psychological Science, 20,* 974–980.

Franklin, A., Drivonikou, G. V., Bevis, L., Davies, I. R.L., Kay, P., & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences, 105,* 3221–3225.

Frith, C. (2007). *Making up the mind: How the brain creates our mental world.* Oxford: Blackwell.

Gaillard, R., Naccache, L., Pinel, P., Clemenceau, S., Volle, E., Hasboun, D., Dupont, S., Maulac, M., Dehaene, S., Adam, C., & Cohen, L. (2006). Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron, 50,* 191–204.

Gainotti, G. (2006). Anatomical, functional, and cognitive determinants of semantic memory disorders. *Neuroscience and Biobehavioral Reviews, 30,* 577–594.

Gallagher, S. (2005). *How the body shapes the mind.* Oxford: Oxford University Press.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8,* 396–403.

Gallese, V., & Lakoff, G. (2005) The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22,* 455–479.

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror neuron system in humans. *Current Biology, 16,* 1824–1829.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge, MA: Cambridge University Press.

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

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*, 278–285.

Gibbs, R. W. (2006). *Embodiment and cognitive science.* Cambridge, UK: University of Cambridge Press.

Giese, M. A. (2006). Computational principles for the recognition of biological movements: Model-based versus feature-based approaches. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 323–360). Oxford: Oxford University Press.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience, 4*, 179–192.

Gilbert, A. L., Regier, T., Kay, P., & 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*, 489–494.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language, 105*, 91–98.

Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge, UK: Cambridge University Press.

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Anderson, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience, 26*, 6523–6532.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006a). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience, 26*, 4917–4921.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006b). Distinct and common cortical activations for multimodal semantic categories. *Cognitive, Affective, and Behavioral Neuroscience, 6*, 214–222.

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., & Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *NeuroImage, 32*, 906–912.

Grafton, S.T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences, 1156*, 97–117.

Graziano, M. (2006). The organization of behavioral repertoire in motor cortex. *Annual Review of Neuroscience, 29*, 104–134.

Graziano, M. S. A., & Aflalo, T. N. (2007a). Mapping behavioral repertoire onto the cortex. *Neuron, 56*, 239–251.

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

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14–23.

Grossman, E. D. (2006). Evidence for a network of brain areas involved in perception of biological motion. In G. Knoblich, I. A. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 361–384). Oxford: Oxford University Press.

Grossman, M., Anderson, C., Khan, A., Avants, B., Elman, L., & McCluskey, L. (2008). Impaired action knowledge in amyotrophic lateral sclerosis. *Neurology, 71*, 1396–1401.

Guo, J., Lieven, E., Budwig, N., Ervin-Tripp, S., Nakamura, K., & Ozcaliksan, S. (Eds.). (2008). *Crosslinguistic approaches to the study of language: Research in the tradition of Dan Isaac Slobin.* Philadelphia: Psychology Press.

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

Hamilton, A. F., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience, 26*, 1133–1137.

Hampe, B. (Ed.). (2005). *From perception to meaning: Image schemas in cognitive linguistics.* Berlin: Mouton de Gruyter.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience, 9*, 1367–1368.

Harrison, N. A., Singer, T., Rotshtein, P., Dolan, R. J., & Critchley, H. D. (2006). Pupillary contagion: Central mechanisms engaged in sadness processing. *Social, Cognitive, and Affective Neuroscience, 1*, 5–17.

Hart, J., & Kraut, M. A. (2007a). Neural hybrid model of semantic object memory (version 1.1). In J. Hart & M. A. Kraut (Eds.), *The neural basis of semantic memory* (pp. 331–359). Cambridge, UK: Cambridge University Press.

Hart, J., & Kraut, M. A. (Eds.). (2007b). *The neural basis of semantic memory*. Cambridge, UK: Cambridge University Press.

Hartley, T., Maguire, E. A., Spiers, H. J., & 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.

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.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science, 303*, 1634–1640.

Hauk, O., Davis, M. H., Kherif, F., & Pulvermüller, F. (2008a). Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *European Journal of Neuroscience, 27*, 1856–1866.

Hauk, O., Shtyrov, Y., & Pulvermüller, F. (2008b). The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. *Journal of Physiology, Paris, 102*, 50–58.

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

Haxby, J. V., Gobbini, M. I., & Montgomery, K. (2004). Spatial and temporal distribution of face and object representations in the human brain. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences III* (pp. 889–904). Cambridge, MA: MIT Press.

Heinze, S., & Homberg, U. (2007). Maplike representation of celestial *E*-vector orientations in the brain of an insect. *Science, 315*, 995–997.

Herskovits, A. (1986). *Language and spatial cognition.* Cambridge, UK: Cambridge University Press.

Hoenig, K., Sim, E.J., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience, 20*, 1799–1814.

Howard, R. J., Ffytche, D. H., Barnes, J., McKeefry, D., Ha, Y., Wodruff, P. W., Bullmore, E. T., Simmons, A., Williams, S. C. R., David, A. S., & Brammer, M. (1998). The functional anatomy of imagined and perceived color. *NeuroReport, 9*, 1019–1023.

Hwang, K., Palmer, E.D., Basho, S., Zadra, J.R., & Müller, R.-A. (2009). Category-specific activations during word generation reflect experiential sensorimotor modalities. *NeuroImage, 48*, 717–725.

Iacoboni, M., Molnar-Szakacs, I., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror system. *PloS Biology, 3*, e79.

Inoue, K. (2000). Visualizing ability and nominal classification. In G. Senft (Ed.), *Systems of nominal classification* (pp. 217–238). Cambridge, UK: Cambridge University Press.

Jackendoff, R. (2002). *Foundations of language.* Oxford: Oxford University Press.

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

James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology, 13*, 1792–1796.

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

Johnson, M. (2007). *The meaning of the body: Aesthetics of human understanding.* Chicago: University of Chicago Press.

Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., & Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *Journal of Cognitive Neuroscience, 17*, 1855–1870.

Kable, J. W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience, 14*, 795–805.

Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations. *Cognitive Neuropsychology, 20*, 525–540.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London, B, 361*, 2109–2128.

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

Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive, Affective, and Behavioral Neuroscience, 1*, 207–221.

Kellenbach, M. L., Brett, M., & Patterson, K. (2003). Actions speak louder than functions: The importance of manipulability and action in tool representation. *Journal of Cognitive Neuroscience, 15*, 30–46.

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. (2009). A neuroscientific perspective on the linguistic coding of categorical spatial relations. In V. Evans & P. Chilton (Eds.), *Language, cognition, and space: The state of the art and new directions* (pp. 139–168). London: Equinox.

Kemmerer, D. (forthcoming). Visual and motor features of the meanings of action verbs: A cognitive neuroscience perspective. In R. G. Almeida & C. Manouilidou (Eds.), *Verb concepts: Cognitive science perspectives on verb representation and processing*. Oxford: Oxford University Press.

Kemmerer, D., Chandrasekaran, B., & Tranel, D. (2007). A case of impaired verbalization but preserved gesticulation of motion events. *Cognitive Neuropsychology, 24*, 70–114.

Kemmerer, D., & Gonzalez Castillo, J. (2009). The Two-Level Theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language*.

Kemmerer, D., Gonzlez Castillo, J., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Languag, 107*, 16–43.

Kemmerer, D., Rudrauf, D., Manzel, K., & Tranel, D. (submitted). Behavioral Patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions.

Kemmerer, D., & Tranel, D. (2000a). A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology, 17*, 393–414.

Kemmerer, D. & Tranel, D. (2000b). Verb retrieval in brain-damaged subjects: 1. Analysis of stimulus, lexical, and conceptual factors. *Brain and Language, 73*, 347–392.

Kemmerer, D., & Tranel, D. (2003). A double dissociation between the meanings of action verbs and locative prepositions. *NeuroCase, 9*, 421–435.

Kemmerer, D. & Tranel, D. (2008). Searching for the elusive neural substrates of body part terms. *Cognitive Neuropsychology, 25*, 601–629.

Keysers, C., & Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in Brain Research, 156*, 383–406.

Keysers, C., Wicker, B., Gazzola, V., Anton, J.-L., Fogassi, L., & Gallese, V. (2005). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron, 42*, 335–346.

Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *Journal of Neurophysiology, 97*, 4296–4309.

Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, & Hoenig, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience, 28*, 12224–12230.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*, 16–32.

Koch, C. (2004). *The quest for consciousness.* Englewood Cliffs, NJ: Roberts and Company Publishers.

Kohonen, T., & Hari, R. (1999). Where the abstract feature maps of the brain might come from. *Trends in Neurosciences, 22*, 135–139.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience, 12,* 48–55.

Kriegeskorte, N., Mur, M., Ruff, D., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron, 12,* 1126–1141.

Laeng, B., Chabris, C. F., & Kosslyn, S. M. (2003). Asymmetries in encoding spatial relations. In K. Hugdahl & R. J. Davidson (Eds.), *The asymmetrical brain* (pp. 303–339). Cambridge, MA: MIT Press.

Lambon Ralph, M. A., Pobric, G., & Jeffries, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: Convergent evidence from rTMS. *Cerebral Cortex, 19*(4), 832–838.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16,* 217–238.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

Leutgeb, J. K., Leutgeb, S. L., Moser, M.-B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science, 315,* 961–966.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation.* Chicago, IL: University of Chicago Press.

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

Levinson, S. C. (2003a). Language and mind: Let's get the issues straight! In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind* (pp. 25–46). Cambridge, MA: MIT Press.

Levinson, S. C. (2003b). *Space in language and cognition.* Cambridge, UK: Cambridge University Press.

Levinson, S. C., & 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., & Wilkins, D. (2006). *Grammars of space.* Cambridge, UK: Cambridge University Press.

Lewis, J. W. (2006). Cortical networks related to human use of tools. *The Neuroscientist, 12,* 211–231.

Lin, L., Chen, G., Kuang, H., Wang, D., & Tsien, J. Z. (2007). Neural encoding of the concept of nest in the mouse brain. *Proceedings of the National Academy of Sciences, 104,* 6066–6071.

Lincoln, A. E., Long, D. L., & Baynes, K. (2007). Hemispheric differences in the activation of perceptual information during sentence comprehension. *Neuropsychologia, 45,* 397–405.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585–589.

Longcamp, M., Tanskanen, T., & Hari, R. (2006). The imprint of action: Motor cortex involvement in visual perception of handwritten letters. *NeuroImage, 33,* 681–688.

Lucy, J. A. (1992). *Grammatical categories and cognition.* Cambridge, UK: Cambridge University Press.

Luschi, P., Benhamoou, S., Girard, C., Ciccione, S., Roos, D., Sudre, J., & Benvenuti, S. (2007). Marine turtles use geomagnetic cues during open-sea homing. *Current Biology, 17,* 126–133.

Mahon, B.Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron, 63,* 397–405.

Mahon, B.Z., & Caramazza, A. (2005). The orchestration of the sensory-motor systems: Clues from neuropsychology. *Cognitive Neuropsychology, 22,* 480–494.

Mahon, B.Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology, Paris, 102,* 59–70.

Mahon, B.Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology, 60,* 27–51.

Majid, A., Boster, J.S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition, 109,* 235–250.

Malikovic, A., Amunts, K., Schleicher, A., Mohlberg, H., Eickhoff, S. B., Wilms, M., Palomero-Gallagher, N., Armstrong, E., & Zilles, K. (2007). Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: A probabilistic, stereotaxic map of area hOc5. *Cerebral Cortex, 17,* 562–574.

Malt, B. C., Gennari, S., Imai, M., Ameel, E., Tsuda, N., & Majid, A. (2008). Talking about walking: Biomechanics and the language of locomotion. *Psychological Science, 19,* 232–240.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25–45.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270,* 102–105.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category specificity and the brain: The sensory/motor model of semantic representations of objects. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 1023–1036). Cambridge, MA: MIT Press.

Martin, A., & Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cognitive Neuropsychology, 20,* 575–587.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature, 379,* 649–652.

McCandliss, B. D., Cohen, L., and Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7,* 293–299.

Mechelli, A., Sartori, G., Orlandi, P., & Price, C. J. (2006). Semantic relevance explains category effects in medial fusiform gyri. *NeuroImage, 30,* 992–1002.

Meteyard, L., Bahrami, B., & Vigliocco, G. (2008a). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science, 18,* 1007–1014.

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008b). Visual motion interferes with lexical decision on motion words. *Current Biology, 18,* R732–R733.

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences, 32,* 376–382.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action,* 2nd ed. Oxford: Oxford University Press.

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

Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social, Cognitive, and Affective Neuroscience, 4,* 101-109.

Murphey, D. K., Yoshor, D., & Beauchamp, M. S. (2008). Perception matches selectivity in the human anterior color center. *Current Biology, 18,* 216–220.

Murray, E., & Richmond, B. (2001). Role of perirhinal cortex in object perception, memory, and associations. *Current Opinion in Neurobiology, 11,* 188–193.

Noordzij, M. L., Neggers, S. F.W., Ramsey, N. F., & Postma, A. (2008). Neural correlates of locative prepositions. *Neuropsychologia, 46,* 1576–1580.

Noppeney, U., Josephs, O., Kiebel, S., Friston, K. J., & Price, C. J. (2005). Action selectivity in parietal and temporal cortex. *Cognitive Brain Research, 25,* 641–649.

Noppeney, U., Price, C. J., Penny, W. D., & Friston, K. J. (2006). Two distinct neural mechanisms for category-selective responses. *Cerebral Cortex, 16,* 437–445.

Okada, T., Tanaka, S., Nakai, T., Nishizawa, S., Inui, T., Sadato, N., Yonekura, Y., & Konishi, J. (2000). Naming of animals and tools: A functional magnetic resonance imaging study of categorical differences in the human brain areas commonly used for naming visually presented objects. *Neuroscience Letters, 296,* 33–36.

Oliver, M. (2001). *The leaf and the cloud.* Cambridge, MA: Da Capo Press.

Op de Beeck, H., Haushofer, J., & Kanwisher, N. G. (2008). Interpreting fMRI data: Maps, modules, and dimensions. *Nature Reviews Neuroscience, 9,* 123–135.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8,* 976–987.

Paulesu, E. J., Harrison, S., Baron-Cohen, S., Watson, J. D.G., Goldstein, L., Heather, J., & Frackowiak, R. S.J. (1995). The physiology of colored hearing: A PET activation study of color-word synaesthesia. *Brain, 118,* 661–676.

Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking.* Cambridge, UK: Cambridge University Press.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience, 8,* 636–648.

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex, 15,* 1866–1876.

Pirog Revill, K., Aslin, R. N., & Tannenhaus, M. T., & Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences, 105,* 13111–13115.

Postma, A., & Laeng, B. (Eds.). (2006). New insights in categorical and coordinate processing of spatial relations. Special issue of *Neuropsychologia, 44(9).*

Prinz, J. J. (2005). The return of concept empiricism. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science.* St. Louis: Elsevier.

Pulvermüller, F. (2008). Brain embodiment of category-specific semantic memory circuits. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 71–97). Cambridge, UK: Cambridge University Press.

Pylyshyn, Z. (1984). *Computation and cognition.* Cambridge, MA: MIT Press.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435,* 1102–1107.

Reddy, L., & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in Neurobiology, 16,* 408–414.

Ritz, T., Thalu, P., Phillips, J. B., Wiltschko, R., & Wiltschko, W. (2004). Resonance effects indicate a radical-pair mechanism for avian magnetic compass. *Nature, 429,* 177–180.

Roberson, D., Pak, H., & Hanley, R. (2007). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition, 107,* 752–762.

Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patternson, K., & Price, C. J. (2007). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, and Behavioral Neuroscience, 6,* 201–213.

Rüschemeyer, S.-A., Brass, M., & Friederici, A. D. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience, 19,* 855–865.

Sacks, O. (1995). The case of the colorblind painter. In *An anthropologist on Mars.* New York: Vintage.

Sakai, K., Watanabe, E., Onodera, Y., Uchida, I., Kato, H., Yamamoto, E., Koizumi, H., & Miyashita, Y. (1995). Mapping of the human color center with echo-planar magnetic resonance imaging. *Proceedings of the Royal Society of London, B, 261,* 89–98.

Sanches, M. (1977). Language acquisition and language change: Japanese numeral classifiers. In B. Blount & M. Sanches (Eds.), *Sociocultural dimensions of language change* (pp. 51–62). New York: Academic Press.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M. B., & Moser, E. I. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science, 312,* 758–762.

Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain, 130,* 2452–2461.

Saygin, A. P., Wilson, S. M., Hagler, D. J., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *Journal of Neuroscience, 24,* 6181–6188.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10,* 24–30.

Schulze-Berndt, E. (2006). Sketch of a Jaminjung grammar of space. In S. C. Levinson & D. P. Wilkins (Eds.), *Grammars of space* (pp. 63–113). Cambridge, UK: Cambridge University Press.

Semenza, C. (2005). The (neuro-)psychology of mass and count nouns. *Brain and Language, 95,* 88–89.

Semin, G.R., & Smith, E.R. (Eds.) (2008). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches.* Cambridge, UK: Cambridge University Press.

Senft, G. (Ed.) (2000). *Systems of nominal classification.* Cambridge, UK: Cambridge University Press.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., & Davis, A. S. (2000). The functional neuroanatomy of implicit motion perception or "representational momentum." *Current Biology, 10,* 16–22.

Senior, C., Ward, J., & David, A. S. (2002). Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition, 9,* 81–92.

Serino, A., De Filippo, L., Casavecchia, C., Coccia, M., Shiffrar, M., & Ladavas, E. (in press). Lesions to the motor system affect action perception. *Journal of Cognitive Neuroscience.*

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology, 20,* 451–486.

Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex, 15,* 1602–1608.

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia, 45,* 2802–2810.

Slobin, D. I. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107–138). Amsterdam: John Benjamins.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). Cambridge, MA: 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 & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 219–258). Mahwah, NJ: Lawrence Erlbaum.

Smith, E. E., & Kosslyn, S. M. (2007). *Cognitive psychology: mind and brain.* Upper Saddle River, NJ: Pearson Prentice Hall.

Snell-Hornby, M. (1983). *Verb descriptivity in German and English.* Heidelberg: Carl Winter.

Spiers, H. J., & Maguire, E. A. (2006). Thoughts, behaviour, and brain dynamics during navigation in the real world. *NeuroImage, 31,* 1826–1840.

Taler, V., Jarema, G., & Saumier, D. (2005). Semantic and syntactic aspects of the mass/count distinction: A case study of semantic dementia. *Brain and Cognition, 57,* 222–225.

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

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109–140.

Taylor, K. I., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2006). Binding crossmodal object features in perirhinal cortex. *Proceedings of the National Academy of Sciences, 103,* 8239–8244.

Thompson-Schill, S. L., Aguirre, G. K., D'Esposito, M., & Farah, M. J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia, 37,* 671–676.

Tomasello, M. (2003). *Constructing a language.* Cambridge, MA: Harvard University Press.

Tranel, D., Damasio, H., & Damasio, A. (1997a). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia, 35,* 1319–1327.

Tranel, D., & Kemmerer, D. (2004). Neuroanatomical correlates of locative prepositions. *Cognitive Neuropsychology, 21,* 719–749.

Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H., & Damasio, A. N. (2003). Neural correlates of conceptual knowledge of actions. *Cognitive Neuropsychology, 20,* 409–432.

Tranel, D., Logan, C. G., Frank, R. J., & Damasio, A. R. (1997b). Explaining category-related effects in the retrieval of conceptual and lexical knowledge for concrete entities: Operationalization and analysis of factors. *Neuropsychologia, 35,* 1329–1339.

Tranel, D., Manzel, K., Asp, E., & Kemmerer, D. (2008). Naming static and dynamic actions: Neuropsychological evidence. *Journal of Physiology, Paris, 102,* 80–94.

Tranel, D., Martin, C., Damasio, H., Grabowski, T. J., & Hichwa, R. (2005). Effects of noun-verb homonymy on the neural correlates of naming concrete entities and actions. *Brain and Language, 92,* 288–299.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B.H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311,* 670–674.

Urgesi, C., Moro, V., Candidi, M., & Aglioti, S. M. (2006). Mapping implied body actions in the human motor system. *Journal of Neuroscience, 26,* 7942–7949.

Vandenbulcke, M., Peeters, R., Fannes, K., & Vandenberghe, R. (2006). Knowledge of visual attributes in the right hemisphere. *Nature Neuroscience, 9,* 964–970.

Van Schie, H. T., Koelewijn, T., Jensen, O., Oostenveld, R., Maris, R., & Bekkering, H. (2008). Evidence for fast, low-level motor resonance to action observation: An MEG study. *Social Neuroscience, 3,* 213–228.

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition, 1,* 215–244.

Wandell, B. A. (2008). Colour vision: Cortical circuitry for appearance. *Current Biology, 18,* 250–251.

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron, 56,* 366–383.

Weisberg, J., van Turennout, M., & Martin, A. (2007). A neural system for learning about object function. *Cerebral Cortex, 17,* 513–521.

Whatmough, C., Chertkow, H., Murtha, S., & Hanratty, K. (2002). Dissociable brain regions process object meaning and object structure during picture naming. *Neuropsychologia, 40,* 174–186.

Wheatley, T., Weisberg, J., Beauchamp, M. S., & Martin, A. (2005). Automatic priming of semantically related words reduces activity in the fusiform gyrus. *Journal of Cognitive Neuroscience, 17,* 1871–1885.

Wible, C. G., Han, S. D., Spencer, M. H., Kubiciki, M., Niznikiewicz, M. H., Jolesz, F. A., McCarley, R. W., & Nestor, P. (2006). Connectivity among semantic associates: An fMRI study of semantic priming. *Brain and Language, 97,* 294–305.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in *my* insula: The common neural basis of seeing and feeling disgust. *Neuron, 40,* 655–664.

Wiggs, C. L., Weisberg, J., & Martin, A. (1999). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia, 37,* 103–118.

Wu, D. H., Waller, S., & Chatterjee, A. (2007). The functional neuroanatomy of thematic role and locative relational knowledge. *Journal of Cognitive Neuroscience, 19,* 1542–1555.

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition, 105,* 229–236.

Zeki, S. (1990). A century of cerebral achromatopsia. *Brain, 113,* 1721–1777.

Zeki, S. (1991). Cerebral akinetopsia (visual motion blindness): A review. *Brain, 114,* 811–824.

Zeki, S. (1993). *A vision of the brain.* Oxford: Blackwell.

Zeki, S., & Bartels, A. (1999). The clinical and functional measurement of cortical (in)activity in the visual brain, with special reference to the two subdivisions (V4 and V4 alpha) of the human color centre. *Philosophical Transactions of the Royal Society of London, B, 354,* 1371–1382.

Zeki, S., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. J. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of* Neuroscience, 11, 641–649.

Zihl, J. D., Von Cramon, N., Mai, N., Schmid, C. H. (1991). Disturbance of movement vision after bilateral posterior brain damage. *Brain, 114,* 2235–2252.

Zwaan, R. A., & Kaschak, M. P. (2009). Language in the brain, body, and world. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 368-381). Cambridge, UK: Cambridge University Press.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 224–245). Cambridge, UK: Cambridge University Press.

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

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13,* 168–171.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin and Review, 10,* 954–958.

AUTHOR INDEX

SUBJECT INDEX

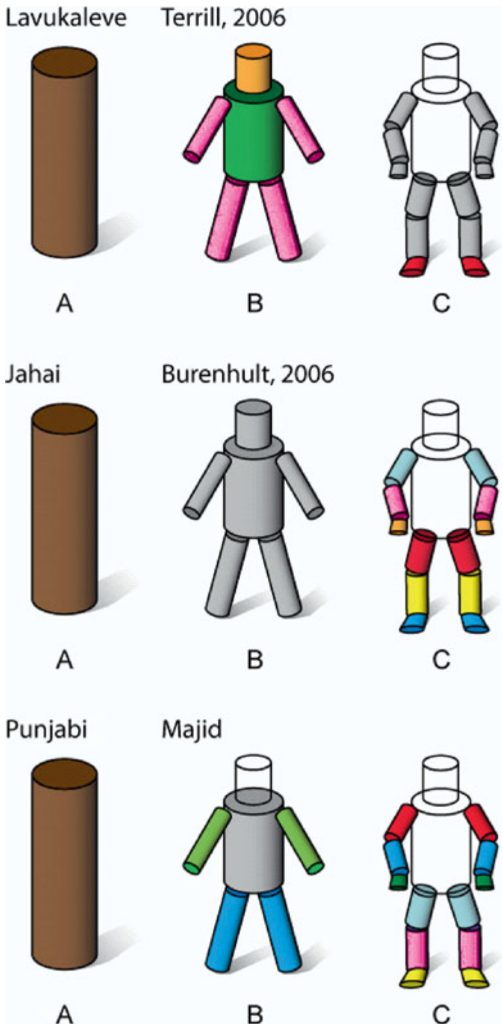


PLATE 1. Three different languages and how they name parts of the body (A–C). A gray geon means that there is no conventionalized means for talking about that body part. So, for example, in Jahai there is no word for HEAD, TRUNK, ARM, or LEG. Within a language, geons with the same color are referred to with the same word. Thus Lavukaleve speakers use *tau* to refer to ARM and LEG. Note that while Lavukalave names body parts at level B, Jahai names at level C, and Punjabi names at all levels suggesting that naming of geons at each level of the hierarchy is an independent choice. (See Figure 3.2)

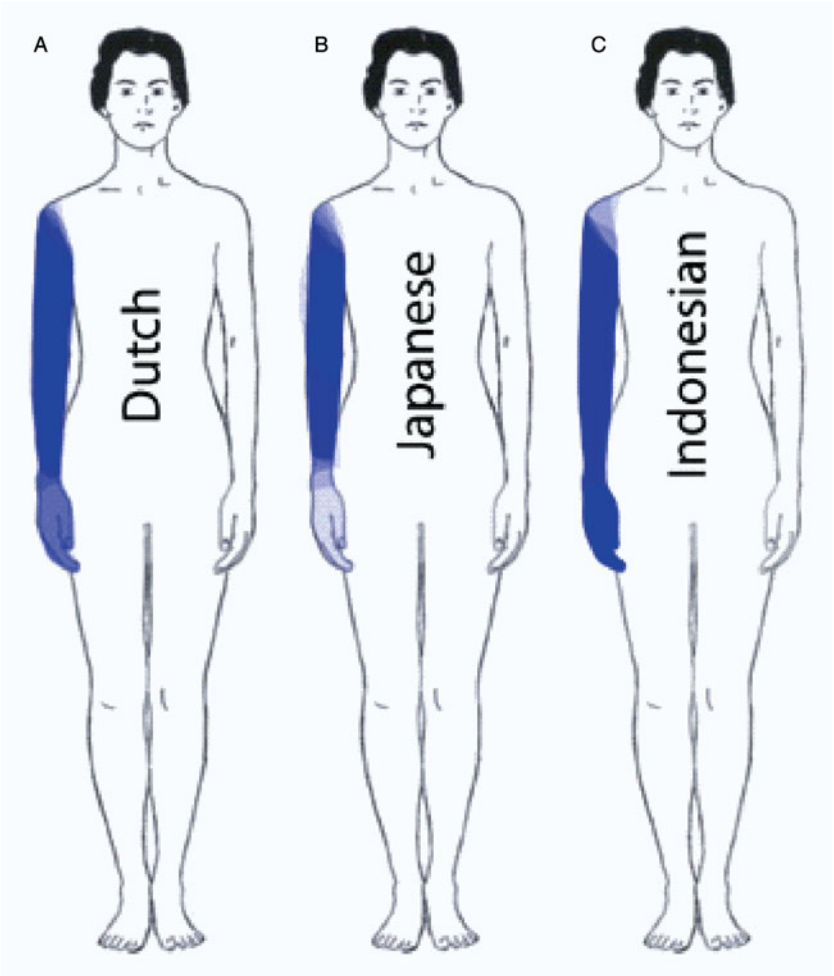


PLATE 2. Eight Dutch, Japanese, and Indonesian speakers were asked to color in parts of the body. Their responses were then layered into a single image so that points of consensus could be viewed. The darker the image, the more speakers colored in that part of the body; the lighter the image, the fewer who included that part. These are the results when Dutch speakers were asked to color in the *arm*, Japanese speakers the *ude*, and Indonesian speakers the *tangan*. (See Figure 3.3)

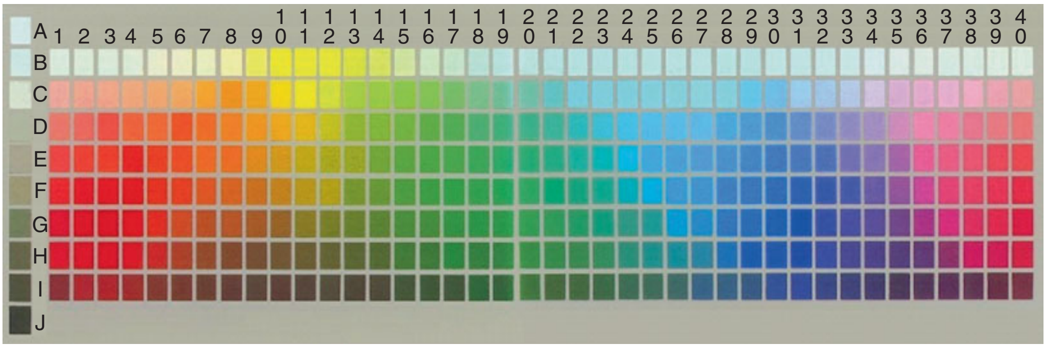


PLATE 3. Color naming grid. (See Figure 8.1)

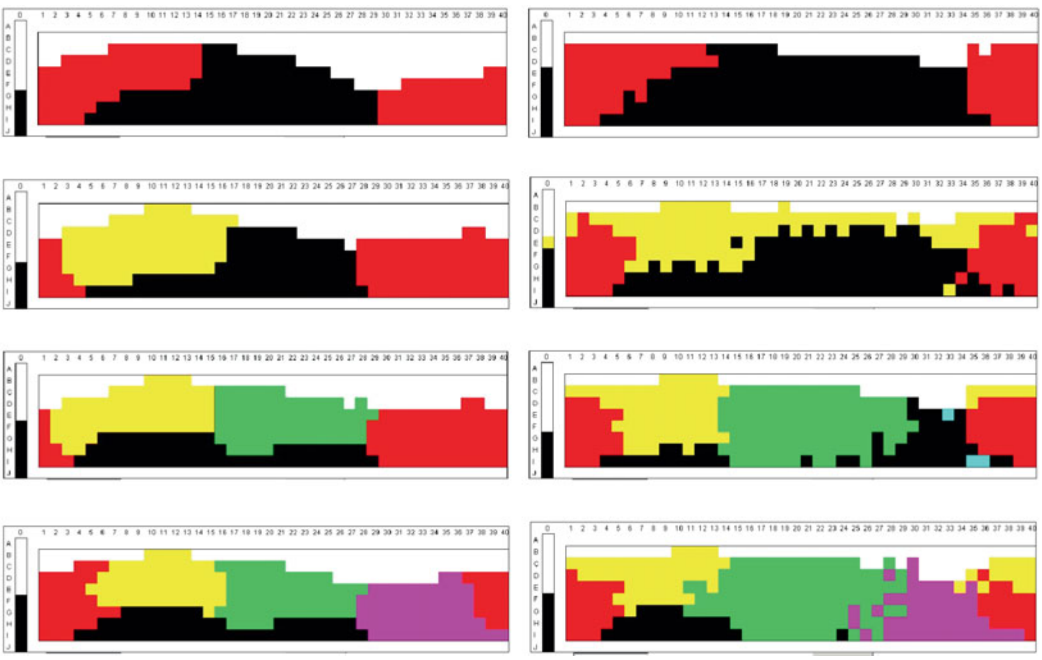


PLATE 4. Theoretical optima (left) compared with selected WCS languages (right), for $n = 3, 4, 5, 6$ categories. The WCS languages are, from top to bottom: Ejagam (Nigeria, Cameroon), Culina (Peru, Brazil), Iduna (Papua New Guinea), and Buglere (Panama). (See Figure 8.4)

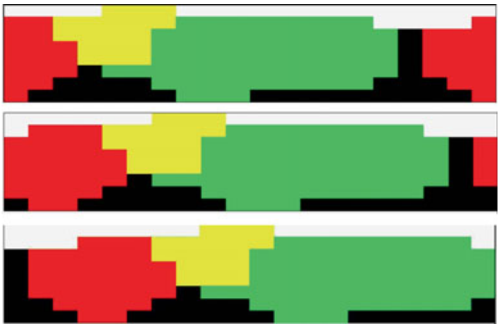


PLATE 5. Creating hypothetical color-naming systems by rotation. The top panel shows the color-naming system of Berinmo; the lower panels show the same system rotated by four (middle panel) and eight (bottom panel) hue columns. (See Figure 8.5)

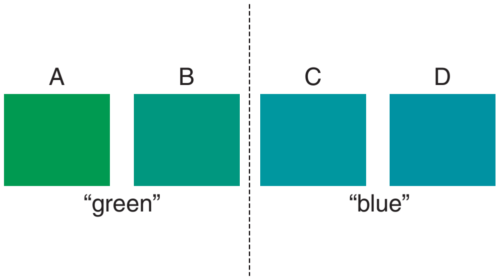


PLATE 6. Four colors (A–D) spanning the green/blue boundary—two greens and two blues. (See Figure 8.6)

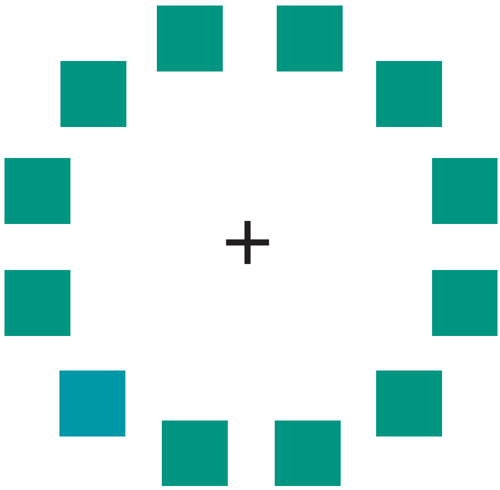


PLATE 7. Visual search task: is the odd-man-out on the left or the right? (See Figure 8.7)
(Reprinted from Gilbert et al., 2006.)

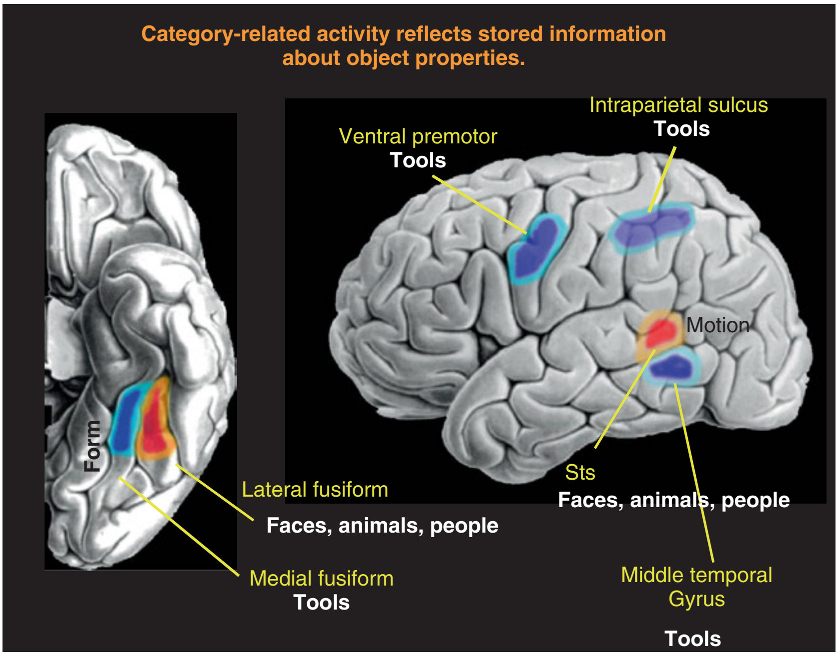


PLATE 8. Distinct activation patterns for animal (including human) and tool concepts as revealed by functional neuroimaging studies employing both pictures and printed words as stimuli. The left image shows parallel but segregated lateral (red) and medial (blue) sectors of the fusiform gyrus that may contain assemblies of conjunctive neurons for capturing the complex shape patterns of the two categories of objects. The right image shows parallel but segregated superior temporal sulcus (STS, red) and middle temporal gyrus (blue) regions that may contain assemblies of conjunctive neurons for capturing the complex motion patterns of the two categories of objects. The right image also shows ventral premotor and intraparietal sulcus regions that may contribute to visuomotor knowledge of tools. (See Figure 14.2)

(This figure was kindly provided by Alex Martin.)

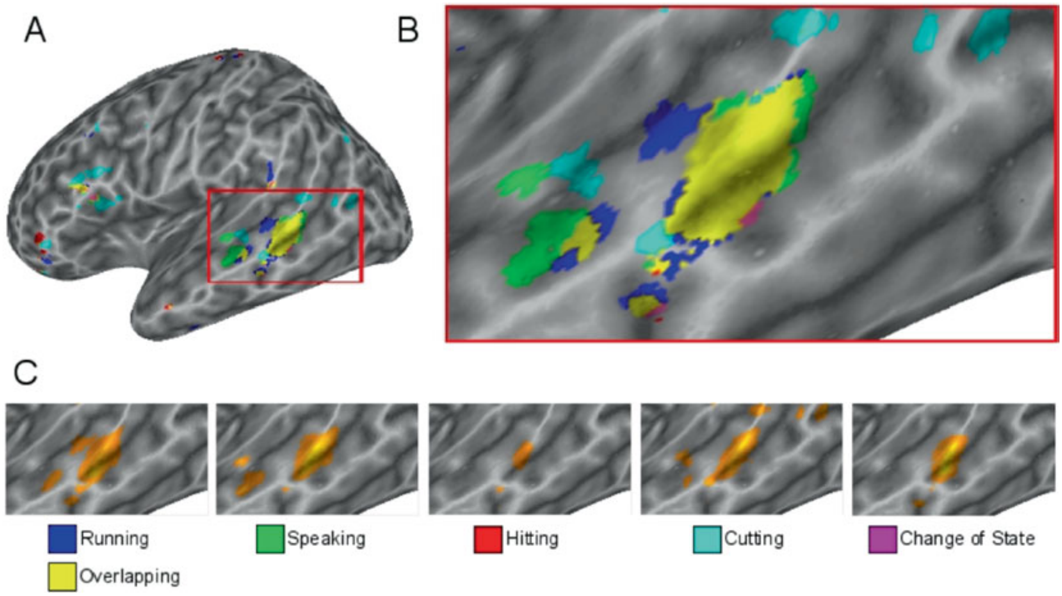


PLATE 9. Activation patterns for five classes of verbs—Running, Speaking, Hitting, Cutting, and Change of State—in the posterolateral temporal cortex (PLTC). (A) Activations rendered on the left hemisphere of an inflated brain, with the PLTC enclosed in a red box. Yellow patches signify areas of overlapping activation for two or more verb classes. Other colored patches signify areas of activation unique to particular verb classes, according to the color key in (C). (B) Enlargement of the red box in (A). (C) Activations for each separate verb class in the territory of the PLTC indicated by the red box in (A) and (B). (See Figure 14.3)

(Adapted from Kemmerer et al., 2008.)

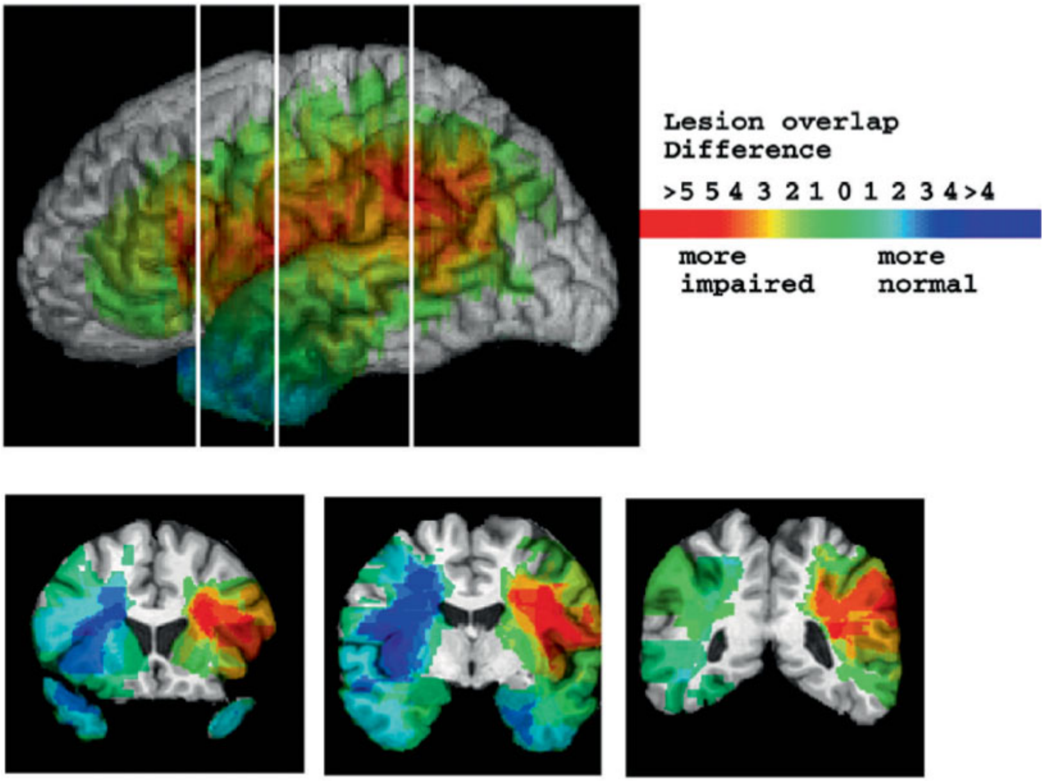


PLATE 10. Results from a lesion study in which 78 brain-damaged subjects performed the Matching Test, which assesses knowledge of English locative prepositions. For each test item (n = 50), the subject is shown three pictures of objects in various spatial relationships and is asked which picture best represents the meaning of a particular preposition. For example, in one item the preposition is *in* and the three pictures show (1) one window above another window on the outside of a house, (2) eggs in a carton, and (3) a boy on a swing. The figure shows the subtraction of lesion overlaps for 15 unimpaired subjects from the lesion overlaps for 15 impaired subjects. The color bar indicates the number of lesions in the overlap difference (the difference reached as high as 7 in the red-coded zone for "more impaired"). The top panel shows a lateral view of the left hemisphere and reveals that, relative to unimpaired subjects, impaired subjects more often had lesions in the left frontal operculum extending posteriorly into the left inferior parietal lobule, specifically the supramarginal gyrus. The vertical white lines in the top panel denote the planes of the coronal sections depicted in the bottom three panels. The leftmost line, and corresponding leftmost bottom panel, indicates the plane of the frontal operculum, and the rightmost line, and corresponding rightmost bottom panel, indicates the plane of the supramarginal gyrus. (See Figure 14.5)

(Adapted from Tranel & Kemmerer, 2004.)