Words and the Mind

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 cross-cutting 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.