Whither Whorf

Dedre Gentner and Susan Goldin-Meadow

For the last two decades, the hypothesis that language can influence thought—generally known as the *Whorfian hypothesis*—has been in serious disrepute. Admitting any sympathy for, or even curiosity about, this possibility was tantamount to declaring oneself to be either a simpleton or a lunatic. The view of most language researchers is well expressed by Pinker (1994, 65): "Most of the experiments have tested banal 'weak' versions of the Whorfian hypothesis, namely that words can have some effect on memory or categorization. Some of these experiments have actually worked, but that is hardly surprising." Devitt and Sterelny (1987, 178) express this skepticism even more strongly: "[T]he argument for an important linguistic relativity evaporates under scrutiny. The only respect in which language clearly and obviously does influence thought turns out to be rather banal: language provides us with most of our concepts." The latter quotation exemplifies the rather schizophrenic way in which the Whorfian question has been viewed. The language-and-thought question is dismissed as banal and unimportant, yet in the same breath it is stated (almost in passing) that language provides us with most of our concepts—a view far stronger than that of even the most pro-Whorf researchers.

Whorf was not the first to express the idea that language influences thought. For example, Humboldt (1836) viewed language as the formative organ of thought and held that thought and language are inseparable (see Gumperz and Levinson 1996a; Lucy 1996, for reviews). Whorf's own views were somewhat more subtle than is generally realized. Along with his well-known strong conjecture:

We dissect nature along lines laid down by our native language. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscope flux of impressions which has to be organized by our minds—and this means largely by the linguistic systems of our minds. (1956, 213)

he also considered weaker views:

My own studies suggest, to me, that language, for all its kingly role, is in some sense a superficial embroidery upon deeper processes of consciousness, which are necessary before any communication, signaling, or symbolism whatsoever can occur ... (1956, 239)

Nonetheless, the hypothesis that has come to be known as the Whorfian hypothesis, or alternatively the *Sapir-Whorf hypothesis*, states that (1) languages vary in their semantic partitioning of the world; (2) the structure of one's language influences the manner in which one perceives and understands the world; (3) therefore, speakers of different languages will perceive the world differently.

Why would anyone ever come up with the hypothesis that the language we speak shapes the thoughts we think? Consider a plausible scenario. When retelling an event, speakers of Turkish are required by their language to indicate whether they themselves actually witnessed that event (Aksu-Koç and Slobin 1986). Of course, the speaker knows whether she witnessed the event. However, she may not be interested in conveying this bit of information to the listener. Speakers of English have the option (which they often exercise) of leaving out whether they actually witnessed the event they are retelling—speakers of Turkish do not. After many years of routinely marking whether they witnessed an event, it is possible that Turkish speakers will tend to encode whether an event has been witnessed, whether or not they are talking. That is, Turkish speakers may habitually attend to this feature of the world much more than English speakers do. In other words, their way of viewing the world may have been altered just by becoming speakers of Turkish as opposed to English. This is the kind of reasoning that underlies the Whorfian hypothesis.

This strong Whorfian position was widely embraced in the 1950s and 1960s, drawing experimental support from Brown and Lenneberg's (1954) studies, which showed a positive relation between the codability of

English color terms (speakers' agreement on the names of colors) and people's ability to retain and recognize a given color from an array. The idea was that color terms influence the way in which people partition the color space, and hence their perception of color, whether or not they are talking. Other work on color within English (e.g., Lantz and Stefflre 1964) also seemed to support the hypothesis. But soon afterward Rosch published her influential paper showing that the Dani people in New Guinea, despite possessing only 2 basic color terms (dark and light), as opposed to 11 in English, nonetheless behaved on cognitive tasks as though their color categories resembled the English system ([Rosch] Heider 1972). Rosch found that the Dani's similarity groupings accorded better with English basic color terms than with their own linguistic groupings. Further, when asked to learn new categories, the Dani found the task easier when the categories were grouped around English focal colors. The implication was that the perception of color—and which colors are considered focal—is determined by the biology of human color perception and not by the language learned.

These negative findings ushered in a period of extreme skepticism concerning the possibility of linguistic influences on thought (e.g., Clark and Clark 1977; Devitt and Sterelny 1987; Pinker 1994). This skeptical view dovetailed nicely with strong zeitgeists in adjoining fields. In linguistics, the Chomskian emphasis on universals of grammar, coupled with the view that language is a separate system from general cognition and with a de-emphasis of the semantic arena, discouraged any search for a relation between language and cognition. Within cognitive psychology, there was a strong sense that concepts come first and that language merely names them: nouns name persons, places, or things; verbs name actions and events; adjectives name modifying concepts; and so on. In cognitive development, the Piagetian influence favored the same direction of influence—from thought to language. The dominant position of cognitive psychologists in the last few decades has been that (1) human conceptual structure is relatively constant in its core features across cultures, and (2) conceptual structure and semantic structure are closely coupled. Note that this view allows for no variation in semantic structure across cultures. The same view can be seen in cognitive linguistics, where the coupling between language and cognition has been

taken to be strong enough to allow semantic structure to serve as a window on conceptual structure. One lucky implication of this view for cognitive researchers was that the semantic structure of any given language—say, English—could serve as a guide to universal conceptual structure.

Strangely enough, even during this period, when discussions of language and thought were about as respectable as discussions of flying saucers, the position was enjoying a revival in folk theories of politically correct speech. Terms like *senior citizens*, *hearing impaired*, and *learning disabled* were assiduously used instead of terms like *old*, *deaf*, and *dumb*. Interestingly, academicians—even while rejecting the hypothesis in their work—joined others in our culture in behaving as though they believed that language could shape thought. Consider the example of *chairman*, now replaced by the term *chair* (suggesting that we, perhaps rather oddly, prefer the risk of confusing a human with an inanimate object over the risk of gender-specific labeling). Presumably the male-oriented label came about because men were the typical occupants of leadership positions; in this sense, our language reflected the state of the world. But why do we think it so important to change the term now? We seem to believe that calling the position *chairman* potentiates a gender bias, and that calling it *chair* can subtly change our perceptions so that we will be less likely to assume that the position should be filled by a male. Insisting upon the word *chair* seems to reflect a folk belief that changing our language can contribute to changing our cognition. Yet despite embracing—or at any rate acquiescing to—this folk belief in their personal behavior, most cognitive researchers continued to find the language-and-thought hypothesis unworthy of serious consideration in their professional life.

Recently things have changed. After decades of neglect, the language-and-cognition question has again become an arena of active investigation. Why? At least three themes can be identified. First was the brilliant work of Talmy, Langacker, Bowerman, and other language researchers who, beginning in the 1970s, analyzed the semantic systems of different languages and demonstrated convincingly that important differences exist in how languages carve up the world. For example, English and Korean offer their speakers very different ways of talking about joining objects. In English, placing a videocassette in its case or an apple in a

bowl is described as putting one object *in* another. However, Korean makes a distinction according to the fit between the objects: a video-cassette placed in a tight-fitting case is described by the verb *kkita*, whereas an apple placed in a loose-fitting bowl is described by the verb *nehta*. Indeed, in Korean, the notion of fit is more important than the notion of containment. Unlike English speakers, who say that the ring is placed *on* the finger and that the finger is placed *in* the ring, Korean speakers use *kkita* to describe both situations since both involve a tight-fitting relation between the objects (Choi and Bowerman 1991).

This evidence of substantial variability in how languages partition the world has profound consequences. It means that at least one—if not both—of the two core assumptions held by cognitive psychologists and linguists is wrong. If semantics varies crosslinguistically, then one cannot maintain that conceptual structure is universal and that semantic structure reflects conceptual structure. One could simply adopt the assumption that semantic structure and conceptual structure are independent of one another, leaving the universal view of cognition intact. However, faced with this dichotomy, a number of researchers have taken the alternative route of exploring ways in which semantic structure can influence conceptual structure.

The second theme developed from a set of theoretical arguments. These include the revival of Vygotsky's (1962) case for the importance of language in cognitive development, Hunt and Agnoli's (1991) influential review paper making the case that language influences thought by instilling cognitive habits, Miller and Stigler's (1987) research on cross-linguistic differences in number systems and their influence on learning arithmetic, and Lucy's (1994) research on the cognitive effects of classifier grammars.

The third important trend was a shift away from the focus on color to the study of domains such as space, which offer much richer possibilities for cognitive effects. Spatial relations are highly variable cross-linguistically (e.g., Bowerman 1980, 1989, 1996; Brown 1994; Casad and Langacker 1985; Levinson and Brown 1994; Talmy 1975, 1985). This suggests the possibility of corresponding cognitive variability. Further, spatial relational terms provide framing structures for the encoding of events and experience. They play a more interesting cognitive role

than color names. Finally, spatial relations, like color concepts, are amenable to objective testing in a more direct way than, say, people's concepts of justice or causality. The work of Levinson's group demonstrating cognitive differences that follow from differences in spatial language—specifically, from the use of absolute spatial terms (analogous to *north-south*) versus egocentric terms (e.g., *right/left/front/back*)—has been extremely influential in attracting renewed interest to the Whorfian question, either arguing for the effect (Levinson 1996, 1997; Levinson and Brown 1994; Pederson 1995) or against it (Li and Gleitman 2002). Interestingly, there has continued to be a line of research on color, the bête noir of the Whorfian hypothesis (Kay and Kempton 1984; Lucy and Shweder 1979). Davidoff, Davies, and Roberson (1999) have recently produced counterevidence to Rosch's claims, based on a reanalysis of her results and on further work with another New Guinea tribe.

These themes coalesced in 1991 with the Wenner-Gren Foundation Symposium in Jamaica on the topic of rethinking linguistic relativity. Its direct result was an influential volume edited by Gumperz and Levinson (1996b), and its indirect result was to spark a renewed look at issues of language and thought. One important outcome of this symposium was Slobin's "thinking for speaking" hypothesis: that language influences thought when one is thinking with the intent to use language and that this influence is not at all trivial. Variants of this idea had been considered before; for example, Pinker (1989, 360) states that "Whorf was surely wrong when he said that one's language determines how one conceptualizes reality in general. But he was probably correct in a much weaker sense: one's language does determine how one must conceptualize reality when one has to talk about it." However, Slobin was the first to focus closely on the idea of "thinking for speaking" and to delineate its implications. This version is more cautious than the grand view that language determines the way in which we perceive the world; but for that very reason it is more palatable, and perhaps more conducive to empirical testing. Moreover, it invites close consideration of the processes by which speakers link cognition and language. It also spurs related questions, such as whether speaking and comprehending are equivalent in their opportunities for linguistic influences on thought and whether language influences thinking when one is talking to oneself (a link with Vygotsky's inner speech).

The purpose of this volume is not to settle the question of language and thought—which in any case we suspect is not one question but several—but to make it clear that the question (or questions) are worth asking and to encourage theoretical and empirical research. This time around we come equipped with better analyses of the linguistic distinctions and a better understanding of the relevant psychological processes and methods for testing them. Theories and experiments are being advanced at a rapid pace. A strong sign of the health of this arena, as is clear from the chapters in this volume, is that there are now close debates on specific issues. Current research continues to find mixed results, again as demonstrated in this volume. But the depth and precision of questions has increased dramatically since the early investigations.

The topics dealt with here range broadly; they include space, number, motion, gender, theory of mind, thematic roles, and the nature and function of objects versus substances. There are even two separate chapters that raise the ante on language and cognition enough to be titled "What Makes Us Smart? Core Knowledge and Natural Language" (Spelke) and "Why We're So Smart" (Gentner). The fields represented span a broad spectrum of cognitive science: cognitive psychology, cognitive development, linguistics, anthropology, and animal cognition.

To begin, theoretical chapters by Clark, Levinson, and Tomasello introduce the relevant questions from different perspectives. The remaining chapters fall into three broad (and overlapping) categories based on their questions and methods: *language as lens*, *language as tool kit*, and *language as category maker*. The answers are far from uniform.

Under the theme *language as lens*, the question posed is whether the language we acquire influences how we see the world. This view is closest to the classical "Whorfian hypothesis" that the grammatical structure of a language shapes its speakers' perception of the world. On the affirmative side, Boroditsky, Schmidt, and Phillips argue that gender assignments, long thought to be purely grammatical, have subtle but pervasive effects on how people think about objects. A more neutral position is taken by Slobin, who argues for limited effects of the semantics of motion verbs on how people talk about—and to some extent how they think about—motion events. On the negative side, Munnich and Landau find that distinctions in spatial language do not predict deviations in spatial representation. Also on the negative side, Malt, Sloman,

and Gennari tested two possible claims of the language-and-thought hypothesis and found evidence for neither. One was that manner-path differences in the way actions are lexicalized would predict which aspects of events are likely to be retained for later recognition; the other was that nonlinguistic similarity groupings of artifact categories would reflect the semantic categories found in different languages.

Under the theme *language as tool kit*, the question posed is whether the language we acquire augments our capacity for representation and reasoning. This theme harks back to Vygotsky's view that "... learning to direct one's own mental processes with the aid of words or signs is an integral part of the process of concept formation" (Vygotsky 1962, 59; quoted in Kuczaj, Borys, and Jones 1989). All the authors in this section argue in the affirmative. Gentner suggests that relational language augments the human ability to engage in relational thought. Kuczaj and Hendry argue that teaching symbolic systems to chimpanzees leads to gains in their cognitive abilities. Inagaki and Hatano discuss linguistic and conceptual factors that influence inductive projection between animals and plants. De Villiers and de Villiers argue that acquiring the ability to use complement clauses fosters the development of theory of mind and thus the ability to pass false-belief tasks. Spelke suggests that language plays a role in providing conceptual links between initially separate modules.

Under the theme *language as category maker*, the question posed is whether the language we acquire influences where we make our category distinctions. On the affirmative side, Bowerman and Choi suggest that the acquisition of spatial semantics in a language influences infants' early categorization of spatial relations, and Lucy and Gaskins argue for the influence of classifier typology on the development of nonverbal classification. Imai and Mazuka take a more neutral stance, arguing for a limited role for linguistic typology and an important role for universal ontological knowledge on early individuation. Finally, Goldin-Meadow finds evidence for a possible universal starting point—namely, the ergative construction—that all humans may experience before learning language. Goldin-Meadow's findings point to thought before language and thus have a non-Whorfian feel. But how can the ergative construction be so basic and at the same time be so difficult for speakers

of nonergative languages to fathom? Perhaps it's because the nonergative languages that most of us speak have irrevocably altered our ergative starting point and thus, in Whorfian fashion, have influenced how we think.

There are some interesting connections among the views taken here. First, in the *language as lens* chapters, some authors dismiss language as simply a mediator of cognition, arguing that when parallels between language and thought are found, it is merely because language is used covertly in the task (in other words, the task is really a language task and thus not a good test of the Whorfian hypothesis). The chapters in the *language as tool kit* section take issue with the "merely" in this claim. They suggest that such covert uses are a manifestation of the usefulness of language in the cognitive arsenal. However, both views agree that language is a powerful mediator of cognition when we speak—and much of our lives is spent in language-related activities. We learn not just by direct experience but also by hearing or reading about the state of affairs, so at least in this sense language has the potential to shape our conceptions of the world.

Another contrast is that whereas the *language as lens* view tends to focus on obligatory elements of language, the *language as tool kit* view encompasses specific content words, such as relational terms, and special-purpose constructions such as the complement clause construction. Also, tests of the *language as lens* hypothesis tend to involve crosslinguistic comparisons; indeed, all of the chapters in this section have taken this tack. In contrast, tests of the *language as tool kit* hypothesis can also be carried out within a language, by comparing outcomes when different sets of symbolic terms are made available to populations: for example, to primates (Kuczaj and Hendry), to children (Gentner), or to deaf individuals learning language late in life (de Villiers and de Villiers). Tests of the *language as category maker* view are often crosslinguistic, comparing speakers of languages that draw the boundaries between categories in different places (Bowerman and Choi, Lucy and Gaskins, and Imai and Mazuka). However, it can also be informative to examine populations that have never been exposed to language on the assumption that these populations offer us a pre-language view of thought (Goldin-Meadow).

Finally, a developmental issue that emerges primarily in the last section of the book is the chicken-and-egg question: which comes first, the concept or the linguistic term? Scholars like Bowerman have for some time challenged the long-standing view that concepts come first and language merely names them. This question clearly calls for a developmental perspective and, indeed, each of the chapters in the *language as category maker* section examines categories over developmental time.

In the past, empirical tests of the language-and-thought question have not proven convincing to either side in the debate. We suggest this stalemate has come about, in part, because the language-and-thought question is not one question but many. Whether language has an impact on thought depends, of course, on how we define language and how we define thought. But it also depends on what we take to be the criterion for "having an impact on." Language can act as a lens through which we see the world; it can provide us with tools that enlarge our capabilities; it can help us appreciate groupings in the world that we might not have otherwise grasped. As illustrated in this book, exploring these and other possibilities requires comparisons across languages and domains, as well as comparisons across thinkers who have and have not been exposed to language. From such an agenda, we are unlikely to get a yes-or-no answer to the whole of Whorf's thesis. But if we have delineated a set of more specific questions for which the answer is no to some and yes to others, we will have achieved our goal.

References

Aksu-Koç, A., and Slobin, D. I. (1986). A psychological account of the development and use of evidentials in Turkish. In W. Chafe and J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology* (pp. 185–201). Norwood, NJ: Ablex.

Bowerman, M. (1980). The structure and origin of semantic categories in the language-learning child. In M. L. Foster and S. Brandes (Eds.), *Symbol as sense* (pp. 277–299). New York: Academic Press.

Bowerman, M. (1989). Learning a semantic system: What role do cognitive predispositions play? In M. L. Rice and R. L. Schiefelbusch (Eds.), *The teachability of language* (pp. 133–168). Baltimore: Brookes.

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

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

Brown, R., and Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology*, *49*, 454–462.

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

Choi, S., and Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *41*, 83–121.

Clark, H. H., and Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics* (pp. 515–558). New York: Harcourt Brace Jovanovich.

Davidoff, J., Davies, I., and Roberson, D. (1999). Colour categories in a Stone-Age tribe. *Nature*, *398*, 203–204.

Devitt, M., and Sterelny, K. (1987). *Language and reality: An introduction to the philosophy of language*. Oxford: Blackwell.

Gumperz, J. J., and Levinson, S. C. (1996a). Introduction to part I. In J. J. Gumperz and S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 21–35). Cambridge: Cambridge University Press.

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

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, *93*, 10–20.

Humboldt, W. von (1836). *On language: The diversity of human language-structure and its influence on the mental development of mankind* (P. Heath, Trans.). Cambridge: Cambridge University Press (1988).

Hunt, E., and Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, *98*, 377–389.

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

Kuczaj, S. A., II, Borys, R. H., and Jones, M. (1989). On the interaction of language and thought: Some thoughts and development data. In A. Gellatly, D. Rogers, and J. Slaboda (Eds.), *Cognition and social worlds* (pp. 168–189). Oxford: Oxford University Press.

Lantz, D., and Stefflre, V. (1964). Language and cognition revisited. *Journal of Abnormal and Social Psychology*, *69*, 472–481.

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

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

Levinson, S. C., and Brown, P. (1994). Immanuel Kant among the Tenejapans: Anthropology as applied philosophy. *Ethos*, 22, 3–41.

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

Lucy, J. A. (1994). *Grammatical categories and cognition*. Cambridge: Cambridge University Press.

Lucy, J. A. (1996). The scope of linguistic relativity: An analysis and review of empirical research. In J. J. Gumperz and S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 37–69). Cambridge: Cambridge University Press.

Lucy, J. A., and Shweder, R. A. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist*, 81, 581–618.

Miller, K. F., and Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development*, 2, 279–305.

Pederson, E. (1995). Language as context, language as means: Spatial cognition and habitual language use. *Cognitive Linguistics*, 6, 33–62.

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

Pinker, S. (1994). *The language instinct*. New York: Morrow.

Talmy, L. (1975). Semantics and syntax of motion. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 181–238). New York: Academic 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 the lexicon* (pp. 57–149). New York: Cambridge University Press.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press. (Original work published 1934.)

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