

How Does Language Affect Thought?

Chris Swoyer, Department of Philosophy, University of Oklahoma

Penultimate version of a paper forthcoming in an interdisciplinary volume on Language and Cognition with Taylor & Francis.

There has been a long, often passionate, debate over the ways in which language affects thought. The claim that a person's language influences how they experience or think about the world is known as the "linguistic relativity hypothesis" or "linguistic relativism". Such influences are causal, and because many different aspects of language could in principle influence many different aspects of thought, the linguistic relativity hypothesis is really a family of claims about these potential influences.

My aim here is to sketch a context for thinking about linguistic relativity. I begin with a very short history of linguistic relativity doctrines (which I shall denominate collectively as the "linguistic relativity hypothesis"), stressing the influence of the ambient intellectual climate on its formulations and fortunes. I

then note several quite distinct versions of the doctrine, examine problems that arise in testing them, and discuss avenues for future work.

There are many snares in the neighborhood. Many participants in the debates over linguistic relativity oversimplify the views of their opponents. The problem is exacerbated because relativistic theses often come in two forms: A bold and arresting version, which is proclaimed, and a weaker, less vulnerable version, which is defended—with the first having a tendency to morph into the second when under attack. Moreover, although relativistic lines of thought often lead to quite implausible conclusions, there is something seductive about them, and even when the arguments are weak they have captivated a wide range of thinkers from a wide range of traditions.

Discussions of relativism are also frequently marred by all-or-none thinking: Either virtually everything is relative or virtually nothing is. But usually the question is whether there is a space for an interesting and plausible version of relativism between claims that are banal (the Babylonians did not have a counterpart of the word “telephone” so they did not think about telephones) and those that are dramatic but almost certainly false (those who speak different languages see the world in totally different ways). And it could turn out that some

versions of the thesis are true while others are false.

Although I will not defend specific empirical claims about the hypotheses here, the discussion will suggest several morals. First, even if all humans are biologically endowed with a rich set of linguistic and cognitive universals, there may still be room for interesting ways in which differences in language could lead to differences in thought. Second, as of now, many versions of the relativity hypothesis (and many natural languages) have not been tested at all. In the relatively few cases where they have, the methodology is not always impeccable nor the results univocal. The results thus far are limited, qualified, and piecemeal. Finally, tests of interesting versions of the hypothesis are **very difficult** to perform. What is needed now is less polemics and more detailed empirical work involving as many different methods as possible.

LINGUISTIC RELATIVITY

What Counts as a Substantive Difference?

People could agree that two languages differ in some way or that two groups engage in rather different forms of reasoning, yet disagree as to whether the

difference is big enough to matter. Take concepts, for example. Some concepts are much more **central** to our thought than others. For example, our concepts of “causation”, “physical object”, “person”, “space”, and “color” are more central to our thought than our concepts of “avocado”, “pickup truck”, and “toothpick”. More interesting versions of relativity involve larger differences like the former rather than the latter. Such debates pit those who see a glass as half full against those who see the same glass as half empty. Whether a difference is large enough to be of interest depends heavily on how large the differences has been thought to be by previous thinkers. Against the background of Whorf’s extreme claims of linguistic relativity, many recently discovered differences seem rather trivial. Given the recent fashion for innate capacities and cognitive universals, the same differences appear more substantive.

A Preliminary Statement of the Linguistic Relativity Hypothesis

There are around four to five thousand languages in use today, each quite different from many of the others. Differences are especially pronounced between languages of different families, e.g., Indo-European languages like English and German and Latin, on the one hand, and non-Indo-European languages like Hopi and Japanese and Swahili, on the other.

Many thinkers have urged that large differences in language lead to large differences in experience or thought. They may even hold that each language embodies a worldview, with quite different languages embodying quite different views, so that speakers of different languages think about the world in quite different ways or even, in a common and pungent—if hazy—metaphor, may live in “different worlds”.

Such suggestions have an intuitive appeal for many readers, but questions about the impact of language on thought are **empirical questions** that can only be answered by **empirical investigation**. Despite considerable progress in the last quarter century, the enthusiasm of partisans on both sides of the debate often far outstripped the available evidence.

A (Slightly) More Careful Statement of the Hypotheses

Interesting versions of the linguistic relativity hypothesis embody two claims:

I. Linguistic Diversity: Languages can differ in substantial ways from one another.

II. Linguistic Influence on Thought: Features of a person’s language influence

how they think, and they influence it in systematic ways.

The thesis of linguistic diversity can be construed uncontroversially. Even if all human languages share numerous abstract linguistic universals, there are often large differences in their syntactic structures and their lexicons, as anyone who has learned a second language can attest. The second claim is more controversial, but because linguistic forces could shape thought in various ways and to varying degree, this thesis comes in more and less plausible forms.

As a first approximation, we can think of language as the independent variable and cognition as the dependent variable. We must replace these general notions with much more fine-grained features of language and thought, however, to obtain testable versions of the general hypothesis. We should try to answer three questions:

1. **Which aspects** of language influence **which aspects** of thought in a systematic way?
2. What **form** does this influence take?
- 3 How **strong** is the influence?

For example, certain features of a language's syntax (e.g., whether there is a distinction between intransitive verbs and adjectives) or its lexicon (e.g., what

color words or spatial vocabulary it contains) might be hypothesized to influence perception, classification, or memory (e.g., in recall tests) in clearly specifiable ways.

LINGUISTIC INFLUENCE ON THOUGHT: A VERY BRIEF HISTORY

Background

Current thought about linguistic relativity has its roots in debates that began in late-eighteenth and nineteenth-century Germany, particularly in the work of Johann Georg Hamann (1730-88), Johann Wilhelm von Humboldt (1767-1835), and especially Gottfried Herder (1744-1803). That work was part of the Romantic reaction to various Enlightenment ideas.

We can view the debates as staking out positions along a continuum between two poles, betwixt two ideal types. At one end of the spectrum, we find the views of Leibniz, Hume, Voltaire, Condillac and other Enlightenment figures who believed in the *constancy of human nature* or, more to the point here, the *constancy of the basic mechanisms and concepts of human thought*. True, they allowed that there might be interesting differences between various languages

(like English and Hebrew). But these differences were seen as a patina over shared basic concepts and modes of thought, and with intellectual care and ingenuity they could be peeled off to reveal the cognitive uniformities underneath.

Many later thinkers, particularly anthropologists, took a more empirical route to a similar destination, often taking a cue from the German anthropologist Adolf Bastian's (1826-1905) postulate of *the psychic unity of mankind*. And far more sophisticated variations on these ideas have been popular in recent decades among proponents of substantive linguistic and cognitive universals (linguistic and cognitive features that are the biological endowment of **all** normal human beings).

At the other end of the spectrum we have various versions of the linguistic relativity hypothesis. From this perspective there **are** striking differences among some (not necessarily all) languages, and at least some of these differences lead to non-trivial differences in how their users perceive and think about the world.

Many of the early champions of this view, including the Romantics, were exceptionally erudite, with a command of an array of diverse languages. Later champions of linguistic relativity based their claims on more direct empirical

contact with the users of different languages rather than just the texts they left behind.

Few thinkers occupy either extreme of the spectrum, but many are much closer to one end than the other. Roughly speaking, the relativists dominated the Western intellectual climate in the first half of the twentieth century and their opponents dominated the second half.

The Big Names

It will be easier to see why the linguistic relativity hypothesis captivated so many thinkers if we briefly consider the more arresting claims of Edward Sapir (1884-1936) and Benjamin Lee Whorf (1897-1941). Sapir was an American anthropological linguist who, like many American anthropologists of his day, was a student of the Dean of American anthropologists, Franz Boas. Whorf, a businessman and amateur linguist, was a student of Sapir. Unlike many earlier champions of linguistic relativity, Sapir and Whorf based their claims on first-hand encounters with the languages and cultures they described, and this gave their accounts a vividness earlier discussion typically lacked. A few quotations will convey the flavor.

Sapir

In 1929 Sapir averred:

Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society. It is quite an illusion to imagine that one adjusts to reality essentially without the use of language and that language is merely an incidental means of solving specific problems of communication or reflection.

(Sapir, 1929, p.209).

Our language affects how we perceive things:

Even comparatively simple acts of perception are very much more at the mercy of the social patterns called words than we might suppose. ...We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation.

(ibid., p.210).

But the differences do not end with perception:

The fact of the matter is that the 'real world' is to a large extent unconsciously built up on the language habits of the group. No two languages are ever sufficiently similar to be considered as representing the same social reality. The worlds in which different societies live are distinct worlds, not merely the same worlds with different labels attached.

(Ibid., p. 209).

Whorf

The linguistic relativity hypothesis gained its widest audience, and notoriety, through the work of Whorf, whose collected writings became something of a relativistic manifesto. Even by the rather lax standards of early discussions of the hypothesis, Whorf is unclear and inconsistent, sliding back and forth between very brash claims and more guarded ones. Debate continues about his considered views, but there is little doubt that his bolder claims were what captivated many readers.

When languages are similar, Whorf tells us, they are not likely to issue in

dramatic cognitive differences. But languages that differ markedly from English and other Western European languages (which Whorf calls, collectively, “Standard Average European”) often **do** lead their speakers to think in very different ways, even to the point of having very different worldviews.

We are thus introduced to a **new principle of relativity**, which holds that all observers are not led by the same physical evidence to the same picture of the universe, unless their linguistic backgrounds are similar, or can in some way be calibrated. ...The **relativity of all conceptual systems**, ours included, and their dependence upon language stand revealed.

(Whorf, 1956, p. 214f, bold added).

We dissect nature along lines laid down by our native languages. 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 kaleidoscopic flux of impressions which has to be organized by our minds—and this means largely by the linguistic systems in our minds. [...] no individual is free to describe nature with absolute impartiality but is constrained to certain modes of interpretation even while he thinks himself most free (Ibid.,

pp. 213-214).

And

...users of markedly different grammars are pointed by their grammars toward different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world.

(Ibid., p. 221).

In yet a third essay “facts are unlike to speakers whose language background provides for unlike formulation of them” (Ibid., p. 235). Indeed,

[Western] Science ...has not yet freed itself from the illusory necessities of common logic which are only at bottom necessities of grammatical pattern in Western Aryan grammar; necessities for substances which are only necessities for substantives in certain sentence positions ...

(Ibid., pp. 269-270).

Both Whorf's and Sapir's discussions brim with metaphors of coercion: Our thought is “at the mercy” of our language, it is “constrained” by it; no one is free

to describe the world in a neutral way; we are “compelled” to read certain features into the world. Here the influence of language on thought is almost preternaturally strong.

Linguistic relativism was defended by numerous other thinkers from many backgrounds. Here is a vivid encapsulation from the philosopher Ernst Cassirer (1874-1945):

...the distinctions which here are taken for granted, the analysis of reality in terms of things and processes, permanent and transitory aspects, objects and actions, do not precede language as a substratum of given fact, but that language itself is what initiates such articulations, and develops them in its own sphere.

(Cassirer, 1923/55 [this is 1923/55 in the references], p.12).

Fred C Swoyer 4/14/10 8:39 AM
Deleted: 1946

Because the linguistic relativity hypothesis came to prominence through the work of Sapir and Whorf, and it is often called the “Sapir-Whorf Hypothesis” or simply the “Whorf Hypothesis” in deference to them. I will stick with the label “linguistic relativity”, however, for although it is not perfect, it makes it easier to separate the hypothesis from the details of Sapir’s and Whorf’s views. The basic ideas can even be generalized, as the work of Nelson Goodman (1978) suggests,

to the claim that symbol systems—including computer languages, conventions for diagrams, even styles of painting—influence perception and thought, but I will focus on natural languages here.

Linguistic relativity hypotheses were popular among many American anthropologists during the first half of the twentieth century, and some anthropologists (who seem less affected by nativist trends than other social scientists) still endorse it. The hypothesis also received succor from behaviorism, the dominant approach in psychology (and to a lesser extent other social sciences) during this period. Many behaviorists found the hypotheses congenial, because they thought that many aspects of human behavior and thought were learned (“conditioned”) rather than innate, so that people with quite different learning histories might well end up with quite different modes of thought. However, with the emphasis on behavior, inner episodes of thought were often held to be beyond the reach of science, or even beyond the pale, so behaviorists could easily slide into the view that nothing remains for language to influence, and the relativity hypothesis becomes a non-issue.

The Demise of Linguistic Relativity

A half-century after Whorf, the linguistic relativity hypothesis had degenerated into the poster child for shoddy empirical work coupled with speculative excess.

What happened?

Cognitive science

Cognitive science happened. As a result of several mutually reinforcing trends beginning in the mid-fifties, behaviorism began to wither and was eventually replaced by cognitive psychology and, more recently, cognitive science (an interdisciplinary approach to the study of cognition). One stimulus for this was the development of the computer and the information sciences, and (despite their differences) most cognitive scientists came to view the human mind as an information encoding and processing system and mental operations as functionally-specifiable computational mechanisms which process information.

This picture is compatible with various versions of linguistic relativity, for example with the view that language and concepts are acquired by very general learning mechanisms through which we might acquire quite different languages or styles of thought. The rise of the cognitive sciences retored the study of inner mental processes (like perception, attention, memory, decision making) to

Fred C Swoyer 4/14/10 8:40 AM

Deleted: ,

Fred C Swoyer 4/14/10 8:40 AM

Deleted: **it did** restore

respectability. It also killed behaviorism with its anti-nativist tendencies, and it required a fairly rich picture of our biological cognitive endowment, which at least allowed for a rich set of innate linguistic and cognitive universals.

Chomsky and nativism

A second major cause of the passing of the linguistic relativity hypothesis was the work of the linguist Noam Chomsky who has argued for over half a century that human beings could only learn natural languages if they had a good deal of innate linguistic equipment to guide their way (e.g., Chomsky, 2000). Chomsky characterized this equipment in different ways over the years, but the abiding theme is that unless infants entered the world with such a biological inheritance, they could never progress beyond the sparse set of utterances they hear to the rich linguistic ability they achieve.

After all, in just a few years all normal children acquire the language that is spoken by those around them. They pick up a highly complex and virtually unbounded ability to distinguish sentences from non-sentences and to understand and utter a virtually unlimited number of sentences they have never encountered before on the basis of the utterances they hear and the feedback (rarely in the

form of corrections) they receive. The problem is that children's data are very unsystematic and sparse compared to the systematic and nearly unbounded linguistic competence they achieve in just a few years.

Hence, the argument continues, the child needs help to progress from this impoverished input to the rich output (viewed as the acquisition of a recursive grammar [in what sense is the grammar recursive rather than some rules of the grammar?—in my field a grammar is called recursive just in case some of its rules are— but it would be fine to say “the acquisition of a grammar with recursive rules”] for a complex natural language). This help can only be provided by something innate that constrains and guides the child in their construction of the grammar. Furthermore, because any child can learn any human language, the innate endowment must put constraints on which of the countless logically possible languages are humanly possible (otherwise the data would be compatible with too many possible languages for children to single one out). In recent years this line of thought is sometimes reinforced by considerations drawn from formal learning theory (e.g., Matthews & Demopoulos, 1989).

If the features of human languages are limited by such innate, language-acquisition mechanisms, there is less scope for the large differences among

languages that the more extreme linguistic relativists have imagined.

Furthermore, if there are innate cognitive universals, as many have also urged (e.g., Brown, 1991; cf. Chiang & Wynn, 2000; Spelke & Newport, 1998; Xu & Carey, 1996), the scope narrows even further. But it is now increasingly recognized that even quite rich linguistic and cognitive universals could leave room for interesting linguistic and cognitive variations within the space of humanly possible languages and modes of thought (example of source?). The question is whether or not they do.

Cognitive modules

Inspired in part by Chomsky's claims on behalf of an innate language acquisition "module" in the brain, many cognitive psychologists went on to argue that the human mind is composed of a number of distinct modules for processing various types of information. Although the classic account of a module (Fodor, 1983) is now generally thought to be too strong, many cognitive scientists still believe that the human mind is composed of relatively independent modules, perhaps many of them. Candidates include modules for acquiring the syntax of one's native language, for recognizing human faces, and even for detecting cheaters (source?).

If such cognitive modules exist, some of the empirical issues about linguistic relativity will translate into issues concerning the ways in which various modules can influence one another. Moreover, champions of modules tend to see many of our cognitive capacities as innate, which tends to run counter to the relativity hypothesis. However, many claims about specific cognitive modules are very difficult to test, and some are based on little more than the Just-So stories of evolutionary psychologists. The issues here are empirical and the jury is still out. Still, even many of those skeptical of modules in general admit that there is a pretty good case for the existence of one or more language modules, and so some version of linguistic nativism certainly seems far more plausible than it did sixty years ago.

Finally, as we will see shortly, the demise of linguistic relativity hypotheses also resulted partly from a few empirical results that were, until recently, often taken to show that the facts on the ground simply did not support them.

TESTING THE HYPOTHESIS

Preliminaries

There are two general strategies for assailing the linguistic relativity hypothesis. First, one might argue that natural languages are not nearly as different from one another as is often supposed. To be sure, there are many surface differences, many features that make a second language difficult to learn, but at a more abstract level, these are just surface variations on an underlying core of similarity (involving different settings of the same several dozen Chomskyan parameters, for example). If this is true, many versions of the relativity thesis are nipped in the bud. Perhaps large differences in languages **would** have led to large differences in thought, but because there are not large differences among human languages, such possibilities do not arise. Here the question of what counts as a difference large enough to be interesting—the half-empty-half-full problem—surfaces yet again.

The second general strategy for attacking the linguistic relativity hypothesis is to argue that although natural languages **do** sometimes (maybe often) differ in substantive ways, these differences do **not** lead to substantive differences in cognition. Either languages are a crust over human thought that can, with luck, be pried off with careful experimentation to reveal species-wide cognitive universals underneath, or else, **if** there are large cognitive differences among human groups, they are due to something other than language (e.g., non-linguistic aspects of

culture).

Example: Color language and color cognition

The more extreme versions of the relativity hypothesis have little support (e.g., Malotki, 1983), but this leaves room for more modest versions. The first step in assessing relativity hypotheses is to take clearer and much more manageable versions of the hypothesis into the lab or out into the field to test. For example, instead of the grand claim that the lexicon of one's native language influences all aspects of thought, one might test the claim that the vocabulary of one's language influences how they will perceive, classify, or remember plants or animals or types of snow.

As it happened, much of the most punctilious investigation of the relativity hypothesis in the 1950s and 1960s involved color language and color cognition. At the outset this was an area where linguistic relativity seemed plausible. On the one hand, there is nothing in the physics of light (e.g., in facts about surface spectral reflectances) that suggests drawing boundaries between colors at one place rather than another; in this sense, our segmentations of the spectrum are arbitrary. On the other hand, it was well known that different languages had color

terms which sliced the color spectrum in different spots. So, since nothing in the physics of color could determine how humans thought about color, it was natural to conclude that a person's color cognition follows the grooves laid down by their color language.

Color was also an auspicious object of study because investigators could use Munsell color chips (a widely used, standardized set of chips of different colors) or similar stimulus materials with subjects all over the globe. This assured that whatever differences they found in their cognitive (dependent) variables really did involve the same independent variable, **color** (as anchored in the chips), rather than some more nebulous construct.

Brent Berlin and Paul Kay (Berlin & Kay, 1969; cf. Kay & McDaniel, 1978) did the most significant work on color. It did much to raise the quality of empirical work on the linguistic relativity hypothesis. And together with much subsequent work it suggested to many that the more robust versions of the linguistic relativity hypothesis were false when it comes to color. Indeed, this and related work (much of it involving ethnobiological and kinship terms) led many to conclude that language did not influence thought in any interesting way.

We now know that colors may be a rather special case, however, for although there is nothing in the physics of color that suggests particular segmentations of the spectrum, the opponent-process theory of color vision, now well confirmed, tells us that there are neurophysiological facts about human beings which influence many of the ways we perceive colors (e.g., source?). We do not know of any comparable innate mechanisms that would channel thought about causation or space or social traits into similarly deep, cross-cultural grooves. There may well be similarities in the ways human beings think about such things, but we cannot conclude this from the research done on color.

Most versions of the hypothesis have never been tested

At least until recently, a close look would have shown that the case of each side of the debate was shaky. Often the only consideration cited in favor of linguistic relativity hypotheses was to point to a difference between two languages and assert that it adds up to a difference in modes of thought. But this simply assumes what needs to be shown, namely that such linguistic differences **give rise to** cognitive differences. On the other hand, refutations of the hypothesis often target implausibly extreme versions or proceed as if refutations in one domain (e.g., color language and color cognition) show that it is false in others like spatial

language and cognition (whereas in fact it remains open whether, say, differences in spatial vocabulary influence navigation or the interpretation of directions).

Extreme versions of the linguistic relativity hypothesis are dead, and good riddance. But several things have recently breathed new life into more circumspect versions. First, at least some of the empirical work thought to undermine linguistic relativity has been criticized on methodological grounds.

Second, we have seen great progress over the last quarter century on both the methodological and the theoretical sides in the study of language and cognition.

Third, highly-regarded empirical work (often employing these new tools) suggests that some aspects of language **do** influence some aspects of thought (one notable example is the work of Melissa Bowerman and her collaborators; e.g., Majid, Bowerman, Kita, Haun, & Levinson, 2004). There is even some literature suggesting that more sweeping negative conclusions about color cognition were premature (e.g., Davies & Corbett, 1997; Winawer et al., 2007).

Finally, as we will now see, it has become increasingly clear that many aspects of language and thought have not been tested for relativity at all.

Why the hypotheses is so difficult to test

Deciding to reopen the case for the linguistic relativity hypothesis is one thing. Adequately testing it is another. There are several reasons why it is very difficult to pin down the influence of specific aspects of a language (or language use) on specific aspects of thought.

Pinning down the independent and dependent variables

First, despite the efforts of able theorists and experimenters, there is still little consensus about many features of **our own** language and thought, which of course makes it difficult to get clear about those of quite alien cultures. To be sure, we have learned a great deal in the last half century about the phonology and syntax of a large number of languages. But in the latter case we have almost been too successful, for we now have numerous approaches to the syntactic theory, along with seemingly endless disputes among their adherents.

When we turn to other aspects of language, the situation is far worse. For example, it seems reasonable to suppose that **if** a language does affect the way its

users think, the meanings of its words and phrases are a strong candidate for having such an impact. But the fragmentary nature of current syntactic theories is nothing compared to the chaos surrounding current theories of linguistic meaning (semantic theories). There are numerous, hotly contested accounts of what meaning **is** and few signs of movement toward agreement about even the most fundamental issues. Nor is there any theoretical consensus about the nature and mechanisms of language **use**. Indeed, there is even debate over what a language **is**, **what one learns** when they learn one, and **what sort of thing would even count as a solution** to these problems.

Things are little better when we turn to cognitive (dependent) variables. Although there has been impressive progress in many areas one would expect to bear on the linguistic relativity debate, much remains up for grabs. For example, there are large and irreconcilable differences about the nature and mechanisms of human judgement and decision making (e.g., Swoyer, 2002) or inductive inference (e.g., Feeney & Heit, 2007)

Or consider concepts. Concepts are relevant here, because at least in the case of linguistically encoded concepts (roughly those with “names”), differences in the lexicon could plausibly be supposed to lead to differences in concepts. But there

is much dispute about what concepts even are; indeed, there is probably less accord now than there was twenty years ago. The only thing most theorists can agree on is that none of the theories now on display offers a satisfactory account of all concepts, and in some cases it's not clear they can account for any (e.g. Margolis & Laurence, 1999). This is especially relevant to the relativity hypothesis because concepts do more than classify. Classification is rarely an end in itself, and concepts underlie all of our higher mental processes, including inference, prediction, planning, learning, and explanation. So if a difference between two languages leads to a difference in the concepts their users deploy, this difference would ramify throughout their modes of thought (I'm not sure I understand the meaning; neither am I (VC) suggest rephrasing).

Fred C Swoyer 4/14/10 8:42 AM

Deleted:

Fred C Swoyer 4/14/10 8:43 AM

Deleted: their cognitive economy

Even if we assume that we are clear enough about the nature of concepts to investigate specific examples, we are not all that clear about many of these. “Causation”, for example, is one of our most central concepts (and one that concerned Whorf). But despite classical psychological work by Piaget and Michotte [do we need refs for Piaget and Michotte?—I didn't put them in since it's pretty classic work but only marginally relevant to work on language and I was trying to keep it limited to that] and much “conceptual analysis” by philosophers, careful empirical study of casual learning and causal reasoning is

only now getting under way, and there is still uncertainty about just what our concept (or concepts) of causation amounts to (e.g., Penn & Povinelli, 2007). Hence, it will be difficult to compare our causal cognition (if it is unified enough to be profitably compared with anything else) with that of a quite different culture. In short, we have at best a tentative handle on many of the central linguistic and cognitive variables that seem likely to be relevant to linguistic relativity hypotheses.

Interactions, aggregation, and confounding

Testing relativity hypotheses is also difficult because most things that happen, and certainly many things involving language and cognition, are the effects of multifarious causal influences that interact with each other in delicate and non-linear ways. Even if we had good models (which we do not) for how certain linguistic variables affect cognitive variables if nothing else interfered, in the real world something always does. When matters are this complex, it is extremely difficult to avoid confounded variables.

In the physical sciences it is sometimes possible to create conditions that shield disturbing influences so that we can ascertain the source, magnitude and direction

of a force whose influence is swamped outside the lab. Once we have learned this, we know that the force will add together with other forces, in conformity with the rules of vector addition for physical forces, and that an object acted on by this sum of forces will behave (by accelerating in accordance with Newton's second law) in just the way that it would if this resultant force vector were the only force acting on it. We typically cannot determine all the forces in the real world, of course, and when it comes to predicting where a leaf blowing in the wind will land, the natural scientists' forecasting powers are not that far ahead of the cognitive scientists. Still, in vector addition we have a simple theoretical picture about how physical forces interact.

Unfortunately, we have nothing remotely comparable in the case of cognitive mechanisms. When several linguistic or other cognitive mechanisms operate at the same time they might reinforce each other, cancel each other out, exhibit some sort of interference effect, or interact in a way that depends on twenty-two of the other psychological processes then taking place. Hence, even if we can learn about the behavior of specific mechanisms in the lab, this does not tell us how those mechanisms would behave over a range of circumstances in the outside world. Here we encounter a virulent example of the familiar trade-off between studying a realistic situation (in the field) and control (in the lab), a

trade-off between internal validity (control) and external validity (realism and generalizability).

When the topic is linguistic relativity, we need to do much of our work in the field, where we have even less grip on what causes what. There has been impressive progress in recent years on isolating causes in such conditions (Pearl, 2009, is state of the art), but in practice the problems remain daunting. If we do discover a difference in the color perception of two groups, is it due to language or to some other variable(s) that is entangled with the language users' culture, physiological traits (e.g., eye pigmentation), or the like? Put another way, a language is a vital part of a culture, and many aspects of a person's culture are likely to covary with aspects of their language in ways that are difficult to tease apart. It is always possible that a generalization will become so hedged in by qualifications that it evaporates into a mist of ever higher-order interactions as experimenters probe an ever wider range of cases (C causes E, unless B is present, though if we have C, B, and D we do get E, unless ...).

In light of such difficulties, it is scarcely surprising that some empirical results bearing on linguistic relativity seem to conflict with others or that there are even claims about wholesale failures to replicate earlier results (e.g., January & Kako,

2007). None of this means that testing the relativity hypothesis is impossible, but it does suggest that a series of related tests will often be needed to establish a firm conclusion. On the positive side, we now have much better tools and methods for studying language and cognition (from developmental psychology through psychology of aging), and a much better understanding of the issues involved than people did even a couple of decades ago. We do not always know what accounts are right, but we now know that many earlier accounts were wrong, and we know about various phenomena that any adequate account would need to explain. So we seem to be on the right track.

RELEVANT VARIABLES

The pendulum swung from relativistic views during the first half of the twentieth century to anti-relativistic views during the second. Now it is drifting back to a let's-go-look-and-see middle. It may help fix ideas if we conclude with a brief look at a few more specific variables that currently appear relevant in testing hypotheses about linguistic relativity.

Linguistic (Independent) Variables

Various aspects of language could easily affect cognition.

Grammar Languages can differ in their grammar or syntax. Many earlier discussions of the linguistic relativity hypothesis focused on grammar and lexicon as independent variables. For instance, many of Whorf's contentions, e.g., his claims about the way Hopi thought about time, were based on (what he took to be) large-scale differences between Hopi and Standard Average European that included grammatical and lexical differences (e.g., Whorf 1956, p. 158).

To take a simpler example, typical word order may vary between languages. In English, the common order is subject, verb, object. In Japanese, it is subject, object, verb. In Welsh, verb, subject, object. And of course there are many subtler grammatical differences between languages. It should be noted that grammar does not mean the prescriptive grammar we learned in grammar school but the syntactic structure of a language. In this sense, a grammar comprises a set of rules (or some equivalent device) that can generate all and only the sentences of a given language. Actual tests will have to be more specific still (this is also true for the variables mentioned below), focusing on specific, manageable aspects of grammar. For instance, languages can differ in whether they make a distinction between intransitive verbs and adjectives. Does this affect any aspects of their

users mental lives?

Lexicon Different languages have different lexicons (vocabularies, roughly), and the lexicons of different languages may classify things in different ways. For example, the color lexicons of some pairs of languages segment the color spectrum at different locations.

Semantics Different languages might have different semantic features (over and above differences in lexical semantics).

Pragmatics It is increasingly clear that context plays a vital role in the use and understanding of language (as well as in cognition; e.g., Philip & Aydede, 2008; Swoyer, 2002). So it is possible that differences in the way speakers of different languages **use** their languages in concrete settings affect their mental life.

Metaphor Different languages employ different metaphors or employ them in different ways.

Cognitive (Dependent) Variables

Thought Language might influence many different aspects of thought. Most empirical work has focused, appropriately enough, on those aspects of cognition that are easiest to assess without relying on language. This is important because we otherwise risk finding influences of one aspect of language on some related aspect **of language**, rather than on some aspect of thought. Commonly studied cognitive variables include perceptual discrimination, availability in memory, and classification. But we also need to ask whether features of one's native language influence one's judgement and decision making, problem solving, inductive inference, or various aspects of social cognition, e.g., classifying people in terms of various traits, or explaining behavior more by citing traits or situations.

Differences in language might also affect more general styles of thought (cf. e.g., Nisbett, Peng, Choi, & Norenzayan, 2008). Again, several recent theorists have proposed dual-process accounts of cognition. Different writers develop this distinction in different ways, but the basic idea is that human beings have two quite different cognitive subsystems (or two types of subsystems). There is an "explicit" subsystem that is largely conscious, symbolic, verbal, rule-governed, serial, flexible and capable of reflection. But there is also an "implicit" subsystem that is largely nonconscious, associative, impulsive, affective and that reacts automatically to stimuli (e.g., Chaiken & Trope, 1999; Sloman, 1996). Aspects of

language might well influence one system more than the other or influence the two systems in different ways.

We already know that cultural differences (which are entangled with language) can affect social cognition (e.g., Norenzayan, Choi, & Nisbett, 2002). As we employ more tools and techniques and approaches, we may also encounter some quite unexpected things. For example, it now appears that the effects of linguistic relativity are stronger in the right visual field than in the left (Aubrey, Gilberta, Regierd, Kaye, & Ivrya, 2008), something no one would have even considered several decades ago.

Example of Influence By way of example, certain features of syntax or of the lexicon might exert a causal influence on certain aspects of visual perception (e.g., on which colors we can discriminate), classification (e.g., on how we sort things by color), or long-term memory (e.g., on which differences among colors we remember most accurately) in clearly specifiable ways. If there is such an influence we would also like to know what mechanisms mediate it, but until we have a better idea whether such difference exist, we are not well positioned to answer deeper questions like this.

Figure One presents a sort of Chinese menu [Eh? New expression on me; does it means you combine anything from column 1 to column 2? Yes—it's a common expression here, but I now see not everywhere. It could be changed to "presents a set of choices" To me a Chinese menu is ordering by number!] illustrating a few (of the many) families of variables that are plausibly hypothesized to be relevant to linguistic relativity theses. Here, one or more variables in the left-hand column (in the proper context) might be thought to influence one or more variables in the right-hand column. Of course, even these variables are too general for a direct test, and in actual empirical work we would consider quite restricted examples of them. For example, one would not test memory in general but, say, some quite specific recall task after priming.

[INSERT FIGURE 1 ABOUT HERE]

There are many combinations of variables that could matter, and when we observe that they might only be active under certain conditions (e.g., it might require priming to elicit a certain memory effect), there seem to be hundreds of things to test. The point is not to emphasize such numbers but to indicate how many aspects of the linguistic relativity hypothesis there are and how implausible glib and general claims about them are likely to be.

CONCLUSION: WHERE ARE WE NOW?

In addition to careful studies of more variables and more languages, we need to employ as many different methods and approaches as possible. The skills of the experimental psychologist, the field linguist, and the scholar of ancient languages all are relevant. Cross-cultural psychology (e.g., Berry, Poortinga, & Pandey, 1996) and cognitive anthropology (e.g., D'Andrade, 1995) will play an ever more prominent role, as will neuropsychology (with its array of methods of brain imaging and, one expects, more powerful tools on the horizon) and biological psychology. Among the approaches here, those that focus on bilinguals will surely be very important. If quite different methods converge on similar conclusions, we will know that we have found something worth knowing.

Given our strong nativist Zeitgeist, it is important to stress that substantive linguistic and cognitive universals are entirely compatible with substantive linguistic and cognitive differences between languages and cultures (and subcultures and other groups). Universals mean that the variations will be played out within a restricted space, but that space may well allow dramatic and unexpected differences. Conversely, finding interesting linguistic or cognitive differences between two groups would not rule out the existence of interesting

linguistic or cognitive universals.

I have not tried to provide a comprehensive overview of the issues surrounding the linguistic relativity hypotheses or to cite much of the field's voluminous

literature. My aim has been to provide orientation and to motivate **the conclusions**

that questions about the impact of a variable on cognition are empirical, **and they**

are questions about **what causes what and how it does so [not sure I**

understand; I feel there is a missing 'and so are' or the like]. Such questions

can only be answered once we specify **which aspects** of an independent variable

influence **which aspects** of thought and what form that influence takes. Such

hypotheses can vary greatly in specificity, strength, and scope. And because small

samples make for weak inductions, a comparison of more than a handful of

linguistic communities is needed to draw any firm conclusions. Furthermore,

testing a specific version of the hypothesis requires a combination of

skills, including those of a good ethnographer, linguist, and experimental

psychologist. Progress will be slow, often painful, sometimes hard to discern. But

that's the thing about science.

REFERENCES

Aubrey L., Gilberta, B., Regierd, T., Kaye, C.A., & Ivrya, R.B. (2008). Support

Fred C Swoyer 4/14/10 9:14 AM

Deleted: following claims: Questions

for lateralization of the Whorf effect beyond the realm of color discrimination.

Language Learning, 58, 63-79.

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*.

Berkeley, CA: University of California Press.

Berry, J.W., Poortinga, Y.H., & Pandey J. (1996). *Handbook of cross-cultural psychology: volume 1, theory and method*, 2nd Ed. Needham Heights, MA: Allyn & Bacon.

Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.

Cassirer, E. (1923/1955). *Philosophie der symbolischen Formen*, Berlin: Bruno Cassirer, vol. 1; trans. R. Manheim, *The philosophy of symbolic forms, volume 1: Language*, New Haven, CT: Yale University Press.

Chaiken, S. & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chiang, W. & Wynn, K. (2000). Infants' tracking of objects and collections.

Cognition, 77, 169-195.

Chomsky, N. (2000). *New horizons in the study of language and mind*.

Cambridge: Cambridge University Press.

D'Andrade, R.G. (1995). *The development of cognitive anthropology*. Cambridge: Cambridge University Press.

Davies, I.R.L. & Corbett, G.C. (1997). A cross-cultural study of colour-grouping:

- Evidence for weak linguistic relativity. *British Journal of Psychology*, 88, 493-517.
- Feeney, A. & Heit, E. (2007). *Inductive reasoning: Experimental, developmental, and computational approaches*, Cambridge: Cambridge University Press.
- Fodor, J.A. (1983). *The modularity of mind*. Boston: MIT Press.
- Goodman, N. (1978). *Ways of worldmaking*. Indianapolis: Hackett.
- January, D. & Kako, E. (2007). Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001). *Cognition*, 104, 417-426.
- Kay, P. & McDaniel, C.K. (1978). The linguistic significance of meanings of basic color terms. *Language*, 54, 610-646.
- 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 Sciences*, 8, 108-114
- Malotki, E. (1983). *Hopi time: A linguistic analysis of the temporal concepts in the Hopi language*. Berlin: Mouton.
- Margolis, E. & Laurence, S. (1999). *Concepts: Core readings*. Cambridge: MIT Press.
- Matthews, R. J. & Demopoulos, W. (1989). *Learnability and linguistic theory*. New York: Springer.
- Nisbett, R.E., Peng, K., Choi, I., & Norenzayan, A. (2008). Culture and systems

- of thought: Holistic versus analytic cognition. J.E. Adler & L.J. Rips (Eds.), *Reasoning: Studies of human inference and its foundations* (pp. 956-985). New York: Cambridge University Press.
- Norenzayan, A., Choi, I., & Nisbett, R.E. (2002). Cultural similarities and differences in social inference: Evidence from behavioral predictions and lay theories of behavior. *Social Psychology Bulletin*, 28, 109-120.
- Pearl, J. (2009). *Causality: Models, reasoning and inference*, 2nd ed. Cambridge: Cambridge University Press.
- Penn, D.C. & Povinelli, D.J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology*, 58, 97-118.
- Sapir, E. (1929). The status of linguistics as science. *Language*, 5, 207-214.
- Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22.
- Spelke, E. S. & Newport, E. (1998). Nativism, empiricism, and the development of knowledge, In R. Lerner (Ed.), *Handbook of child psychology, Vol. 1: Theoretical models of human development*. 5th ed. New York: Wiley.
- Swoyer, F. C. (2002). Judgment and decision making: extrapolations and applications, in R. Gowda & J. Fox (Eds.), *Judgments, decisions, and public policy* (pp. 9-45). Cambridge: Cambridge University Press.

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

Philip, P. & Aydede, M. (2008). *The Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.

Winawer, J., Witthoft, N., Frank, M., Wu, L., Wade, A., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences USA*, 104, 7780–7785.

Xu, F. & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.