Through the Window of Language: Assessing the Influence of Language Diversity on Thought

John A. LUCY


ABSTRACT: The way we understand language diversity, how languages differ in representing reality, affects our approach to understanding linguistic relativity, how that diversity affects thought. Historically, researchers divided over whether the diverse representations of reality across languages were natural or conventional, but all tacitly assumed an optimal fit between language and reality. Twentieth century anthropological linguists interested in linguistic relativity have questioned this assumption and sought to characterize “reality” without it by using domain- or structure-centered approaches. Arguments are presented favoring structure-centered approaches, along with a case illustration. A concluding discussion emphasizes the broader significance of language diversity in human development.

Key words: language and thought, Sapir-Whorf, linguistic relativity, language development, number marking

Introduction

The diversity of human languages has long given rise to speculation about its sources and consequences. Why do languages render the same reality so differently and what are the consequences of those differences for human thought? These two questions are in fact intimately related: how we understand language diversity, the ways languages differ in their renderings of reality, greatly affects our approach to understanding linguistic relativity, the effects of linguistic diversity on thought. The present paper explicates the close interrelationship between these two questions and its significance for current theoretical and empirical approaches.

The shared human capacity for natural language manifests itself in the form of many individual languages differing in important respects. Hence individual speakers do not speak language-in-general but rather one or more specific languages that differ structurally from one another. In this discussion we are concerned specifically then with structural diversity among language codes (e.g., Hopi, Chinese, Spanish) and, ultimately, their impact on thought, what we can call, structural (“linguistic”) relativity (in contrast to generic language effects or those due to specialized institutionalized practices; see Lucy 1996). At issue then are two distinct relationships: each language embodies a particular interpretation of reality and these language interpretations can influence thought about that reality (Lucy 1997a). The interpretation arises from the selection of substantive aspects of experience and their formal arrangement in the verbal code. Such selection and arrangement is, of course, necessary for every language, so the crucial emphasis here is that each language involves a particular interpretation, not a common, universal one. An influence on thought ensues when the particular language interpretation guides or supports cognitive activity more generally and hence the be-
liefs and behaviors dependent on it. Theory and research about such cognitive effects are shaped in important ways by one’s prior orientation to linguistic diversity.

**Historical Orientations to Understanding Linguistic Diversity**

There have been two dominant approaches to understanding the diverse interpretations of reality embodied in language (Aarsleff 1982). First, there are those who view the connection between language and the world as natural (or absolute). Those holding this view have two ways of explaining the observed diversity. For some, one thinks here of Leibnitz, the observed forms all represent various sorts of decline, that is, that what we see is an accumulation of human corruption upon some earlier pristine form of language, the language of Adam or the ancient Hebrews being the classic reference points, but any historical reference point suffices. The route to this uncorrupted form lies through the historical reconstruction of the original language through etymology. One can hear echoes of this view in any language ideology about the supposed slowness of contemporary speech (especially among youth) as contrasted with earlier, more correct or “logical” forms. Alternatively, others discard the historical dimension and simply imagine a pristine linguistic form lying securely within each language and that all we have to do is peel back the superficial encrustation and we will find the natural logic of language revealed. One can hear echoes of this view in all those who speak of underlying competence, deep structures, universal primitives, and the like, all waiting to be excavated from beneath the messy, misleading, and ultimately irrelevant surface diversity.

On the opposing side are those who view the connection between language and experience as artificial (or conventional) and who, therefore, regard diversity as an unavoidable feature of languages. Again, there are two ways to regard this diversity. One classic view celebrates this diversity as of local historical significance. So languages, in this view, contain in their form living traces of history and are to be treasured as repositories of the genius of a people. Thus diversity is embraced, but often with a more or less explicit evaluation of one or another language (or language type) as superior—along with the people who created it. One thinks especially of Humboldt in this regard and his claims that the inflecting languages of Europe were naturally superior as formal instruments for rendering reality (Aarsleff 1988). Alternatively, one can take a more neutral view of this process, but recognize that the lack of a uniform natural relationship presents obstacles to clear communication, especially in philosophy and science, a view classically articulated by Locke. However, in this view, the very conventional nature of language allows us to build up specialized vocabularies and professional jargons as needed to convey our views accurately. So in place of decline and corruption one sees the possibility of progress and perfection in language. Here we find much of the impetus for the rampant multiplication of technical terms that besets contemporary scholarship and for the flight to mathematical formalisms.

Disparate as these two dominant approaches are, they share the common underlying assumptions that there is a *single unitary reality* and an *ideal relation of language to it*, however elusive. Early in the twentieth century these assumptions were challenged by
Benjamin Whorf (1956; see Lucy 1992a), a linguist working in the American anthropological tradition founded by Franz Boas. Whorf questioned the existence of a single ideal relation of language to reality and in precisely this sense he also questioned our conceptualization of a unitary reality, since its qualities would vary as a function of the language used to describe it. If there is no ideal relation of language to reality, hence a fundamental scientific uncertainty about the character of that reality, then the whole problem of the relation of language to experience changes. Claims for universality in the relation of language to reality can no longer simply be presumed but require empirical proof. And no language, whether ancient or modern, received or constructed, can be judged inferior or superior, corrupted or perfected in light of its match with reality. In this view, we are lacking a language-neutral standard against which to form such judgments. And, consequently, no single language can provide through its system of categories a reliable guide to reality for the purposes of research. This is the central problem faced by contemporary researchers dealing with the question of linguistic relativity.

Contemporary Orientations to Linguistic Relativity Research

Since the appearance of Whorf’s formulation of the problem, two main strategies of empirical research have emerged aiming to solve this puzzle of how to provide a neutral ground or frame of reference for comparing languages and cognition (Lucy 1997a). Each approach has characteristic advantages and drawbacks.

One approach, which I call domain-centered, selects a domain of experience (such as color or time or space) and seeks to describe it on language-independent grounds in order first to ask how individual languages treat the domain and subsequently how speakers regard the domain in cognition. The domain-centered approach seeks to solve the comparison problem by asking how different languages partition the same domain of reality. Although the approach offers a number of advantages for comparative purposes, it tends to suffer from two weaknesses. First, the representation of the domain is typically drawn from one linguistic and cultural tradition. As such it begs the question being asked, namely, whether such representations, or even the domain itself, are universally recognized. Acknowledging this problem, some seek to anchor the description in well-established scientific concepts to help assure neutrality and objectivity. This can be illuminating, but more often one ends up with a description in terms of parameters drawn from that science and not from those semantically or structurally relevant to actual linguistic systems; this in turn can lead to a dramatic misrepresentation of the languages at issue. Further, by adopting one vision of reality, even a scientific one, as the standard for comparison, one still necessarily favors the original language and culture from which it arose. This leads, not surprisingly, to any number of demonstrations of difference in which a hierarchy quietly (re)emerges: in effect such efforts simply show how well languages do or do not represent the semantic values of the system framing the comparison. The method used for creating a neutral system based on reality thus often undermines the very possibility of fair comparison.
Perhaps the best-known example of this approach is the long tradition of research on color terms. In this research one represents reality though a selection of color chips designed to sample a color space and then asks speakers of different languages how they partition this space. The difficulties are legion: the proper definition and sampling of a color space is far from obvious; some languages don’t even have color terms and they certainly are not central to the semantic structure of any language; and languages that do have relevant terms may deploy them quite differently, such that they are often misanalyzed by those unfamiliar with the language or simply disqualified for inclusion in the study. But the most important difficulty is that whatever systems are encountered are ultimately arrayed or scaled along an evolutionary cline of how finely they represent the color space, that is, how closely they approach our system of dedicated color adjectives seen as a natural telos. Here we see the re-emergence of the old hierarchical portrayal of languages in terms of how well they match the underlying natural relation of language to reality, which is itself little more than a projection from our language (Lucy 1997a). Ironically, even under these constraints, the available evidence still indicates that variations in color language predict cognitive performance more accurately than do the purportedly underlying natural relations (Davidoff, Davies, and Roberson 1999).

The most successful recent effort at a domain-centered approach has been undertaken in the domain of space (Levinson 2003). Spatial conceptualization has been widely regarded as invariant within philosophical, psychological, and linguistic circles and yet there is variability in its linguistic encoding. For example, speakers of modern European languages tend to favor the use of body coordinates to describe arrangements of objects (e.g., “the man is to the left of the tree”). For similar situations, speakers of other languages such as Guugu Yimithirr (Australian) and Tzeltal (Mayan) favor systems anchored as cardinal direction terms or topographic features respectively (e.g., “the man is to the east/uphill of the tree”). Careful cross-linguistic comparison using a variety of innovative techniques for referential typology has been undertaken to compare “the meaning patterns that consistently emerge from domain-directed interactive discourse.” In a large number of nonlinguistic tasks it appears that speakers of different languages respond in ways congruent with their verbal practices. This research has attempted to gain the advantages of precise, extensive comparison characteristic of a domain-centered approach while simultaneously avoiding its chief pitfalls by incorporating extensive linguistic description and typology into the project from the outset. And it vigorously avoids entering into an evaluation of which type of semantic system is superior or more natural, seeing each as having characteristic costs and benefits. In these respects this effort escapes the usual weakness of domain-centered approaches. But it does so only by allowing serious slippage with regard to the original concern with linguistic structure: a single “language” may use more than one semantic approach to spatial description and languages considered the “same” in their referential usage may in fact be using radically different structural means. In the end then, the linguistic analysis and typology are not concerned so much with linguistic structure in the traditional sense but rather with patterns of linguistic usage; and a
single language structure can be used implement more than one of these usage strategies.

A second approach, which I call *structure-centered*, selects some grammatical structure (such as number or gender or aspect marking), asks how it differs across languages, and how reality might appear differently from the vantage of each relevant system. Structure-centered approaches build squarely on a long tradition of typological work in linguistics (modeled on phonology), seeking to build more neutral and structure-relevant frameworks from the outset, that is, to study languages and the realities they represent through the window of language itself. But the approach is difficult to implement: comparing categories across languages requires extensive linguistic work both in terms of local description and typological framing, and it can be extremely difficult to characterize referential entailments suitable for an independent assessment of cognition. Nonetheless, this approach most closely respects the linguistic facts and thus holds the greatest promise for identifying structural differences and directing the search for cognitive influences in appropriate directions.

The classic example of structure-centered work is Whorf’s comparison of number marking patterns in English and Hopi (Whorf 1956; Lucy 1992a). Whorf argued that English speakers measure and count cyclic experiences such as the passage of a day or a year in the same way as ordinary objects with a form and a substance. This leads ultimately by analogy to the projection of these cycles as forms for a homogeneous substance ‘time.’ By contrast, the Hopi language differentiates these cycles as a distinct type of recurrent event and are not therefore led to the same view of time. From these linguistic observations Whorf was led to identify patterns in habitual behavior that he felt bore the impress of this difference in outlook toward time. There are difficulties in Whorf’s work to be sure, notably the anecdotal quality of the characterization of effects on thought. But his approach does show how to begin with language structure and to build a characterization of reality through the window provided by the languages themselves. And, crucially, it does not entail any hierarchical evaluation of the languages with respect to a pre-given reality. My own research, discussed in the following section, seeks to develop this approach further, providing a more explicit typological anchor for the linguistic comparison and explicit psychological assessments for the cognitive comparisons.

*Structure-Centered Empirical Research: An Illustration*

I have worked for many years now exploring whether the structural differences between American English and Yucatec Maya, a language indigenous to southeastern Mexico, lead to distinctive effects on habitual cognition. This work represents the most fully developed effort at a structure-centered approach to date and can serve therefore to illustrate an approach to the relation of language, thought, and reality that builds outward from language structure. Only a sample of this work can be presented here. I describe some salient contrasts between the two languages, associated cognitive entailments, and a study assessing these entailments.
Language Contrast: Number Marking Semantics

The focus here will be on how Yucatec Maya and American English differ in their nominal number marking patterns (Lucy 1992b: 56-83). First, the two languages contrast in the way they signal plural for nouns. English speakers obligatorily signal plural for nouns semantically marked as referring to discrete objects (e.g., car, chair) but not for those marked as referring to amorphous materials (e.g., sugar, mud, etc.). Yucatec speakers are never obliged to signal plural for any referent, although they often do mark plural for animate referents.

Second, the two languages contrast in the way they enumerate nouns. For English nouns marked as having semantically discrete reference, numerals directly modify their associated nouns (e.g., one candle, two candles); for nouns not so marked, an appropriate unit (or unitizer) must be specified by a form that then takes the number marking (e.g., one clump of dirt, two cubes of sugar). Yucatec requires that all constructions with numerals be supplemented by a special form, usually referred to as a numeral classifier, which typically provides crucial information about the shape or material properties of the referent of the noun (e.g., 'un tz'íit kib' ‘one long-thin candle’, ká'a tz'íit kib’ ‘two long-thin candle’).

In essence all nouns in Yucatec are semantically unspecified as to quantificational unit almost as if they referred to unformed substances. So, for example, the semantic sense of the Yucatec word kib’ in the example cited above is better translated into English as ‘wax’ (i.e., ‘one long-thin wax’) — even though, when occurring alone without a numeral modifier in conditions other than enumeration, the word kib’ can routinely refer to objects with the form and function that we would call candles (as well as to other wax things). Given the quantificational neutrality of the noun it becomes clear why one must specify a unit when counting, since expressions such as ‘one wax’ do not make quantificational sense. By contrast, many nouns in English include the notion of quantificational ‘unit’ (or ‘form’) as part of their basic meaning — so when we count these nouns, we can simply use the numeral directly without any classifier (e.g., one candle). In short, whereas English requires such a unitizing construction only for some nouns, Yucatec requires one for all of its nouns. These complementary patterns of plural marking and numeral modification form part of a unified number-marking pattern evidenced typologically across many languages (Lucy 1992b: 61-71).

Cognitive Hypotheses and Predictions

To assess whether traces of these contrasting verbal patterns appear in speakers’ cognitive activities more generally, we need first to draw out the implications of these grammatical patterns for the general interpretation of experience. If we consider the denotational meaning of nouns referring to discrete concrete referents, that is, stable objects that maintain their physical appearance over time, then certain regularities appear from which cognitive implications can be drawn. The quantificational unit presupposed by English nouns referring to objects of this type is frequently the shape of the object. Hence use of these English lexical items routinely draws attention to the shape of a referent as the basis for incorporating it under some lexical label and as-
signing it a number value. Yucatec nouns referring to objects of this type, lacking such a specification of quantificational unit, do not draw attention to shape and, in fact, fairly routinely draw attention to the material composition of the referent as the basis for incorporating it under some lexical label. If these linguistic patterns translate into a general cognitive sensitivity to these properties of referents of the discrete type, then we can draw the following prediction: Yucatec speakers should attend relatively more to the material composition of stable objects (and less to their shape), whereas English speakers should attend relatively less to the material composition of stable objects (and more to their shape).

**Cognitive Contrast: Shape versus Material Preference**

The prediction has been tested in a variety of ways with both adult and child speakers from both languages (Lucy 1992b, Lucy and Gaskins 2001, 2003). One example will serve to illustrate the approach. Twelve speakers in each language group were shown fifteen triads of familiar objects. Each triad consisted of an original *pivot* object and two *alternate* objects, one of the same shape as the pivot and one of the same material as the pivot. So, for instance, speakers were shown a plastic comb with a handle as the pivot and asked whether it was more like a wooden comb with a handle or more like a plastic comb without a handle. The expectation was that English speakers would match the pivot to the other comb with a handle whereas the Yucatec speakers would match it with the other comb made of plastic. Speakers were shown a large number of such triads, which, across the stimulus set, controlled for size, color, function, wholeness, malleability, and familiarity.

The predicted classification preference was strongly confirmed with adult English speakers choosing the material alternate only 23% of the time and adult Yucatec speakers favoring it 61% of the time. Clearly the two adult groups classify these objects differently and in line with the expectations based on the underlying lexicogrammatical structures of the two languages. However, English-speaking and Yucatec-speaking seven-year-olds showed an identical early bias toward shape—choosing material alternates only 12% of the time. But by age nine the adult pattern was visible: English-speaking children continued to favor shape, choosing material alternates only 18% of the time whereas Yucatec-speaking children were choosing material alternates 42% of the time. Thus, the same kind of language-group difference found among adult speakers is also found in children by age nine—and the result is statistically reliable. The adult and developmental data are jointly displayed in Figure 1.
We can summarize these results as follows. Seven-year-olds show clear sensitivity to referent type independently of language group membership. Here we see they prefer shape as a basis of classification with stable objects; in related experiments, with *malleable objects* they prefer material as a basis of classification. Nine-year-olds show differential sensitivity to referent type in line with their language. This suggests that language categories increase in their importance for cognition between ages seven and nine. Adults show a consolidation into a dominant pattern for each group. We can summarize the overall pattern by saying that the two populations begin by grouping different referent types in the same way and end up grouping these same referent types in quite different ways as a function of language type.

**Common Questions**

Three questions are often raised in regard to these findings and therefore deserve some comment. First, are these same verbal and cognitive patterns evident in everyday life? Mayan speakers clearly do exhibit a great sensitivity to the material properties of objects. In the experimental tasks, they constantly evaluate the material composition of the test items before sorting them—feeling how heavy they are, poking their nails into them to test for malleability, scraping the surface to see what material might be under any paint, smelling and tasting the objects, and generally questioning or commenting on their material properties—something English speakers rarely do. In one telling case, a Yucatec woman sorted items into those that would melt if they were burned versus those that would turn to ash, attending to material in a quite striking way. A preference for material over shape also emerges with loan words from Spanish: a kerosene lantern is referred to as a *gas*, a ceramic bowl is referred to as a *porcelina*, etc. On first visits to urban areas villagers pay great attention to materials composing the
furniture and floors in homes, hotels, and restaurants. Granted, these sorts of incidents are anecdotal; but cumulatively over time they suggest the everyday vitality of a distinctive orientation to the material world.

Second, are these patterns actually due to language rather than to cultural factors such as lack of education or poverty or rural lifestyle? When one considers the whole array of studies, it seems unlikely that any single cultural ‘fact’ will account for all the results. Very specific predictions from language have consistently been confirmed, even though they involve different grammatical patterns, different cognitive functions, and different stimulus materials. No informed cultural argument has yet been made to explain these results. Children do not show these cognitive patterns until the language patterns are in place, both educated and uneducated speakers show the patterns, and adults lacking exposure to the language do not develop the usual response patterns. Finally, similar preferences for material sorting have been reported for urban Japanese speakers who also have a classifier type language but whose culture as well as educational and urban lifestyle contrast with the Maya (see Lucy and Gaskins 2003). The close linkage between language prediction and cognitive result in the absence of any plausible alternative cultural account suggests that the shaping role of language deserves to be taken seriously.

Third, is this a real linguistic relativity? After all, the preference scores are not absolute for either group. Some variability is, of course, expectable in any experiment. But the broader view taken here is not that languages completely or permanently blind speakers to other aspects of reality. Rather they provide speakers with a systematic default bias in their habitual response tendencies. Although some members of each group clearly do not recognize the alternative classification possibility, I believe they could readily be brought to do so fairly quickly, at least for a while. This said, I think the bias would nonetheless return soon after, for it serves a purpose in coordinating social action and guiding individual behavior. That a habitual bias can be recognized and even overcome for a while in special contexts for certain tasks does not render that bias unimportant. In any number of areas such as mastering a foreign language, overcoming racial and gender prejudices, following strict logical inferences, etc., we know that people can modify their behavior for a short while in some contexts, but it is another matter entirely to change habitual behavior wholesale or permanently. In short, we have evidence that language structures bear some relationship to thought, that the direction appears to be from language to thought, and the relationship appears to be robust.

**General Discussion**

Just as language universally mediates culture and mind, helping enable them in all human groups, so too it appears to play a role in producing cultural and mental diversity. The two processes go hand in hand. Only by acceptance of the conventions of one or more particular languages can we speak at all and so gain the advantages of having language support for sophisticated cultural and psychological activities. But this same acceptance of a particular language commits us to the specific conventions of that lan-
language and to their consequences for our thinking. Just as with language universals then, linguistic diversity and its influences should be viewed as natural in human life: they are not some unfortunate contextual corruption that needs to be peeled away or some intrinsic defect that needs to be hammered out of each of us. Language influences form part of the foundation of what it is to be human, that is, a species that adapts to its environment by means of diverse yet stable patterns of symbolic representation.

Comparative work on language diversity is essential in all the human sciences. The reasons for this should now be clear. If the natural process is to think in accordance with our own language, then what we take as neutral reality may in fact be a projection of the emphases of our own language. And what we take to be the meaning of a category in another language may be partly a product of our own semantic accent (Lucy 2003). In short, we risk misunderstanding the interpretations of reality implicit in other languages and their influences on thought if we do not control for our own biases. The remedy for these biases lies in taking the observed categories of other languages seriously, exploring empirically their structure and functioning. We will surely fail to progress if we simply ignore the existence of diversity or erase it through the application of interpretive approaches that effectively render other systems in terms of our own.

Expanded developmental research is also essential. Research on very young children cannot provide a full picture of the emerging relation between language and thought, let alone establish that language variation does not matter for thinking. The substantive finding that relativity effects arise in middle childhood is also theoretically illuminating. Other research on middle childhood indicates that this is a crucial period in the development and integration of higher levels of language and thought (Lucy and Gaskins 2001). We also know that this is the age at which children begin to lose their flexibility in acquiring new languages and are increasingly likely to show interference accents in languages subsequently learned (Lucy 2003) and, on the evidence above, show relativity effects. In short, during this age, substantive advances in development come hand in hand with tangible limitations in the capacity to acquire or understand other languages and with measurable effects of language codes on thought. This suggests an emerging tradeoff whereby higher levels of intellectual and social development are purchased by a deeper commitment to the mediating role of language, that is, to a particular language, one whose system of categories will then quietly shape our thought and culture thereafter. The emerging picture is that each child can achieve the fully developed humanity implicit in the general capacity for language, culture, and mind only by committing to becoming a particular sort of human, that is, one imbued with a historically specific language, culture, and mind. The deeper human universal, then, lies not so much in the substantive commonalities among these historical systems, but rather in the shared functional imperative of the tradeoff, that is, the imperative of engaging particular systems in order to consummate general development.
REFERENCES


John A. Lucy received his Ph.D. in Human Development from the University of Chicago in 1987. He taught six years at the University of Pennsylvania (Anthropology) before returning to Chicago in 1996. He is currently William Benton Professor in Comparative Human Development, Psychology, and the College, and a frequent visitor to the Max Planck Institute for Psycholinguistics. He is author of *Language Diversity and Thought* (1992), *Grammatical Categories and Cognition* (1992), and *Reflexive Language* (1993).

ADDRESS: Department of Comparative Human Development, University of Chicago, 5730 S. Woodlawn Avenue, Chicago, IL 60637 USA. E-mail: johnlucy@uchicago.edu.