

Methodological approaches in the study of linguistic relativity

John A. Lucy
University of Chicago

Empirical research on linguistic relativity faces three methodological challenges. First, it must identify and characterize a language contrast where the semantic structures of different languages produce distinct referential interpretations. Second, on the basis of those language patterns it must articulate specific cognitive entailments or predictions and then assess for evidence of their presence in cognitive activity. And third, it must establish the influence or shaping role of language on the cognitive patterns by using an array of strategic assessments that make competing accounts of the cognitive patterns unlikely. This chapter reviews these challenges, outlines strategies available to address them, and provides examples of each. The examples are all drawn from research on patterns of number marking.

Keywords: bilingualism, cognitive development, comparative methods, language and thought

1. Introduction

Although few doubt the importance of language for human life, we still debate the extent to which language actually shapes thought. And since we do not speak one universal language but many different languages, one perennial issue concerns the extent to which different language shape how we think. Investigating this *linguistic relativity* proposal, or hypothesis, raises several distinct methodological challenges (Lucy 1992a).

The linguistic relativity proposal focuses on whether structural differences among languages affect thought (Lucy 1996). This focus on structural variation contrasts with the broader semiotic issue of whether speaking any language at all affects thinking and with the narrower functional issue of whether specialized uses of language affect thinking. Although a full understanding of linguistic relativity

requires attention to all three levels, the focus on structural differences lies at the heart of the methodological challenges in this area of research.

The internal logic of the proposal links language, thought, and reality in two relationships (Lucy 1997a). First, the structural properties of each language embody a particular *interpretation* of reality. The interpretation arises when substantive aspects of the speaker's experience are selected and formally arranged into configurations of referential meaning. And the interpretations vary across languages. Second, this linguistic interpretation of reality *influences* patterns of thinking about reality more generally. The influence ensues when the meanings embodied in the linguistic interpretation guide or support general cognitive activities such as attention, classification, inference, and memory – even when not engaged directly in speaking the language.

Given the internal logic of the relativity proposal, empirical research on it must confront three principal methodological challenges (Lucy 1992b). First, we must identify and characterize a relevant language contrast. A relevant contrast is one where the semantic structures of two or more languages produce distinct referential entailments that embody an interpretation of experience that could potentially influence thought about reality. Second, we must articulate specific cognitive entailments or predictions that might be visible in behavior, and then assess for evidence of their presence in cognitive activity. These related cognitive patterns must parallel the language contrast independently of the activity of speaking itself. And third, we must establish the influence or shaping role of language on the cognitive patterns. Typically this involves using an array of strategic assessments that make competing accounts of the cognitive patterns increasingly unlikely. This chapter reviews these three challenges, outlines strategies to address them, and provides examples of each. To provide cohesion and demonstrate how a range of studies can resolve issues that remain unclear in any single study, the examples are all drawn from one line of research on patterns of number marking.

2. Identifying and characterizing a relevant language contrast

The first challenge is to identify language categories that provide contrasting interpretations of reality and to characterize their meanings relative both to the language system and to their referential entailments.

The language categories most clearly identifiable as interpreting reality are those involved in the referential function of language, the language function that foregrounds the surrounding “context” (Jakobson 1960). By virtue of their manifest classification of the world, referential categories provide the most direct entry point for exploring the relation language to experience (Lucy 1992b). These

categories include the morphological, lexical, and grammatical categories that lie at the heart of language as a system and that make language distinctive from other forms of communication (Lucy 1992a).

To characterize the meanings of referential categories, we must attend to both sense and denotation.¹ Each referential category has a sense value, that is, it always participates in a system of categories such that its place in that system contributes to its meaning value and it contributes to the value of other elements (Lyons 1977; also Lucy 1992a, 1994, 1997b, 2010). Second, each referential category has a denotation value, that is, it stands in certain regular relationships with referents in the world such that these relations contribute to its meaning and its use then has referential entailments (Lyons 1977; also Benveniste 1971). The two components of meaning always stand in dialectal relationship and actively shape each other, both psychologically within the individual speaker and historically within the language community, to produce a uniquely linguistic structure that can then inform other behavior (Lucy 2010).

Understanding the dialect of sense and denotation matters not only theoretically, in order to characterize the full meaning of language categories, but also methodologically, in order to counter the inevitable tendency to bias the analysis in terms of one's own language. When we focus on the meaning of a single form, or a small set of forms, in isolation from its place in the language as a whole, we risk tacitly applying (or filling in) the missing systemic values from those in our own language. Likewise, when we privilege either denotation or sense at the expense of the other, we are making the unwarranted assumption that they mesh in the same way as they do in our own language. Put in other terms, researchers inevitably bring a *semantic accent* to the task of linguistic comparison, that is, a tendency to interpret individual forms as if they formed part of the analyst's own language system (Lucy 2003, 2004, 2010, 2011). The analytic remedy for this lies in explicitly focusing the comparison between languages on systems of meaning, attending to both sense and denotation, ideally framed within a typological perspective that provides a neutral framework for characterizing differences (Lucy 1992b).

To illustrate this structural approach to contrastive linguistic analysis we can compare the patterns of number marking in American English and Yucatec Maya, an indigenous language of southeastern Mexico (Lucy 1992b: 56–83). First, the two languages contrast in the way they signal plural number for nouns. English signals plural obligatorily for phrases semantically marked +animate or +discrete

1. These two components of meaning are often termed *sense* and *reference* (Lyons 1968). Here I follow Lyons (1977: 176) in distinguishing *reference*, an utterance-dependent meaning, from *sense* and *denotation*, as stable components of meaning associated with the category forms themselves.

(e.g. *the dogs, the chairs*, etc.) but not for those marked –discrete (e.g. *the sugar, the mud*, etc.). Yucatec does not draw this distinction formally: speakers are never obliged to signal plural for any referent. However, they may opt to mark plural and often do for animate referents. Thus, the two languages agree in frequently marking plural on animate phrases and rarely marking plural on non-discrete phrases, but they disagree on how to handle discrete entities that are discrete but not animate.

Second, the two languages contrast in the way they enumerate nouns. English is again split such that for noun phrases marked as semantically discrete, numerals directly modify their associated nouns (e.g. *one candle, two candles*); for noun phrases not so marked, an appropriate unit must be specified by a form that then takes the number marking (e.g. *one clump of mud, two cubes of sugar*). Yucatec is again continuous in that all constructions with numerals must be supplemented by a special *numeral unitizer* form (traditionally referred to as a *numeral classifier*) that typically provides crucial information about the shape or material properties of the referent of the noun (e.g. *'un-ts'ít kib'* 'one long-thin candle', *káa-ts'ít kib'* 'two long-thin candle').

The need for these unitizers reflects the fact that all nouns in Yucatec are semantically unspecified as to quantificational unit – almost as if they referred to unformed substances. Hence most nouns can occur with a variety of unitizers to specify a quantificational unit, as illustrated in Example (1), which displays how various unitizers can occur with the numeral *'un* 'one' and the noun *háas* 'banana' such that each combination signals a different referent. The noun *háas* itself is

- | | | |
|-----|-----------------------|---|
| (1) | <i>'un-ts'ít háas</i> | 'one one-dimensional banana (i.e. the fruit)' |
| | <i>'un-wáal háas</i> | 'one two-dimensional banana (i.e. the leaf)' |
| | <i>'un-kúul háas</i> | 'one plant(ed) banana (i.e. the plant/tree)' |
| | <i>'un-kúuch háas</i> | 'one load banana (i.e. the bunch)' |
| | <i>'um-p'ít háas</i> | 'one bit banana (i.e. a bit of the fruit)' |

semantically neutral among these meanings, although context typically makes the referent clear. Likewise, 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 it can routinely denote a candle. By contrast, many nouns in English include the notion of quantificational 'unit' (or 'form') as part of their meaning – so when we count these nouns, we can simply use the numeral directly without any unitizer (e.g. *one candle*). Thus, whereas Yucatec requires a unitizer for all of its nouns when enumerating, English only requires one for some nouns, those without inherent quantificational unit.

These patterns of pluralization and numeral unitization are complementary. English nouns that have an inherent quantificational unit do not require a unitizer,

but do require a plural mark where relevant, whereas nouns lacking an inherent quantificational unit require a unitizer, but no plural marking. Yucatec nouns all lack a quantificational unit, hence all require unitizers and none require a plural marker – although one can be applied optionally when speakers wish to emphasize some multiplicity of referents in the utterance. This complementary distribution is portrayed in Table 1, where the optional plural marking for animates in

Table 1. Regular plural and unitizer marking patterns in Yucatec Maya and American English (optional Yucatec plural marked in parentheses)

Language	Referential noun phrase semantics		
	+Animate +Discrete	–Animate +Discrete	–Animate –Discrete
English	plural	plural	unitizer
Yucatec	unitizer (plural)	unitizer	unitizer

Yucatec is signaled by parentheses. Further, this complementarity is evident across many languages. Those with rich, obligatory plural marking tend not to have obligatory unitizing constructions and those with rich, obligatory use of numeral unitizers tend not to have plural marking. In languages with both types of marking, the lexicon tends to be internally split, as in English, such that noun phrases requiring plural marking with multiple referents tend not to require unitizers for counting, and those requiring unitizers for counting tend not to require plurals when referring to multiple referents. And across languages there is an ordering relationship such that some referents are more likely to have plural marking and others to have unitizer marking – and it is this typological ordering that motivates the feature analysis in the table above. The wide distribution and systematic nature of this interrelationship also suggest that this complementary distribution responds to language-internal semiotic dynamics rather than to local social or environmental factors.

3. Articulating and assessing related patterns in cognitive activity

The second challenge is to articulate a cognitive prediction from the language patterns and then to assess individual speakers for evidence of the predicted patterns in a way that is free of language interference in the assessment process.

3.1 Articulating a cognitive prediction based on language patterns

Predicting cognitive behavior from language behavior requires identifying patterns of meaning in language that require speakers to attend to the world in certain ways. As indicated in the previous section, obligatory and/or ubiquitous referential categories compel speakers to attend to these meaning values and their denotational correlates whenever they speak. We then ask whether traces of these patterns of referential meaning appear in cognitive classifications used in the general interpretation of experience.

We can illustrate the process of developing a cognitive prediction with number marking in English and Yucatec (Lucy 1992b: 85–90). English requires the plural to be signaled for a wider range of referent types than does Yucatec. The most common denotational correlate of the plural is whether one or more than one of a given referent is present. Thus English speakers must attend to the number and kinds of objects in order to signal number as required, whereas Yucatec speakers need not, though of course they may signal number if they wish. If this linguistic pattern translates into a general sensitivity to number in other cognitive activities, then English speakers should habitually attend to number for a wider array of referent types than should Yucatec speakers. These referent types have thus far been characterized in terms of referential features but can now be given notional interpretation for cognitive purposes: [+animate, +discrete] referents are typically animals, [–animate, +discrete] referents are typically objects, and [–animate, –discrete] referents are typically materials. English and Yucatec agree in pluralizing animals and in not pluralizing materials. However, they disagree with respect to objects: English marks plural for them whereas Yucatec does not. The prediction then is that English and Yucatec speakers engaging in cognitive activities that require attention to number will both attend to number for animals, both ignore number for materials, but differ from each other with objects, as shown in Table 2.

Table 2. Predicted attentiveness to number as a function of referent types in Yucatec Maya and American English

Language	Referent types		
	Animals	Objects	Materials
English	attend	attend	ignore
Yucatec	attend	ignore	ignore

Yucatec requires unitizers for a wider range of referent types than does English. The typical denotational correlate of the need for a unitizer varies across lexical noun types both within and across languages. If we focus first on nouns referring

to *stable objects* that typically maintain their physical appearance over time, the unit presupposed by English nouns is frequently the shape of the object hence English speakers must routinely attend to the shape of a referent in order to incorporate it under a lexical label. Yucatec nouns by contrast do not presuppose a unit and thus do not require attention to shape but rather to the material composition of the referent in order to incorporate it under a lexical label. If this linguistic pattern translates into a general sensitivity to these features for referents of this type, then English speakers should attend relatively more to the shape of referents and Yucatec speakers should attend relatively more to their material composition. By contrast, for *malleable objects* that can change shape but retain their cohesion without the assistance of a container, neither English nor Yucatec presupposes a quantificational unit and both routinely require attention to the material composition of a referent in order to incorporate it under a lexical label. Thus both English and Yucatec speakers should attend relatively more to the material composition of malleable objects than they do for stable objects. The prediction, shown in Table 3, is that during cognitive activities that require attending to shape versus material, English and Yucatec speakers should respond to malleable objects in the same way but disagree in their treatment of stable objects. Alternatively, looking within each language group, the prediction is that English speakers will show a cognitive split vis-à-vis the two types of objects whereas Yucatec speakers will show cognitive continuity across them.

Table 3. Predicted attentiveness to shape versus material as a function of referent type in Yucatec Maya and American English

Language	Referent types	
	Stable	Malleable
English	shape	material
Yucatec	material	material

3.2 Assessing for the presence of predicted cognitive patterns

Contemporary research that aims to assess cognitive predictions drawn from language needs to meet several key requirements. First, we must assess the cognitive activity of individual speakers. Although Whorf (1956) identified the importance of the “microcosm” within each speaker as the locus of relativity effects, he did not formally assess the thought processes of any individual speaker in his comparative work but rather relied on comparisons of broad cultural patterns of behavior. Since the work of Brown and Lenneberg (1954) it has become standard

to assess a sample of individuals engaging in one or more controlled cognitive tasks to assure that the cognitive patterns are present in individual speakers and to improve the precision of the research. A second requirement introduced by Brown and Lenneberg (1954) is that the assessment be “nonlinguistic”, by which they meant that the behavior serving as evidence of cognitive patterns should not itself be verbal so as to assure that the response arises from general cognition rather than the specific requirements of the response modality. Several other requirements have been introduced (Lucy 1992b) to help make the comparisons culturally valid and fair: assessments should contrast patterns of activity rather than absolute responses, they should employ familiar or readily interpretable tasks and materials, and they should use a variety of tasks and materials. Unfortunately, many studies still contrast absolute performance using novel materials on a single task, making interpretation difficult.

The process of cognitive assessment can be illustrated by drawing again on number marking research. One prediction was that English and Yucatec speakers should differ in their attentiveness to number as a function of referent type. To test this, speakers were presented with a line drawing of an everyday scene and asked to judge which of several alternate drawings most resembled it. Each alternate was identical except that one element had been changed: the number of an animal, an object, or a material. English speakers should judge the material alternate as most similar and reject alternates that changed the number of animals or objects, referents for which they routinely attend to number. By contrast, Yucatec speakers should divide their choices evenly between the object and material alternates and reject the alternate that changed the animals, referents for which they routinely attend to number. As shown in Figure 1, both predictions were born out. English speakers always chose the material alternate and never the animal or object alternates. Yucatec speakers divided their choices roughly evenly between the object and material alternates and only rarely chose the animal alternate. This task meets the requirements discussed above in that it assesses individual speakers, elicits a nonlinguistic response, and utilizes readily interpretable tasks and materials. Further, the results are robust across different stimuli, instructions, and tasks (Lucy 1992b).

The second prediction was that English and Yucatec speakers should differ in their attentiveness to material versus shape more for some referent types than for others. This prediction was tested in a number of tasks, only one of which will be described here. Speakers in each language group were shown triads of familiar objects to classify. Each triad consisted of an original object and two alternate objects, one of the same shape as the original and one of the same material as the original. So, for instance, speakers were shown a plastic comb with a handle as the original and asked whether it was more like a wooden comb with a handle, that

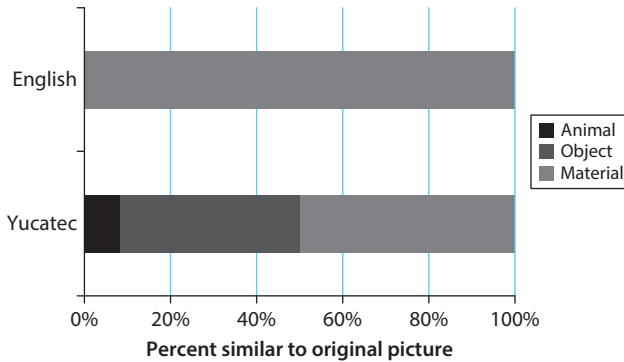


Figure 1. Comparison of English and Yucatec similarity judgments with pictures varying numbers of different referent types

is, a shape match, or more like a plastic comb without a handle, that is, a material match. For each such triad the speaker was asked “Is this [pointing to original] more like this [pointing to one alternate] or more like this [pointing to the other alternate]?” As shown in Figure 2, the predictions were confirmed. For stable objects, where the two languages differ, Yucatec speakers preferred material over two and a half times more than English speakers whereas for malleable objects, where the two languages agree, the difference between the two was attenuated and not statistically reliable. Again, this task assesses individual speakers, elicits a nonlinguistic response, and utilizes readily interpretable tasks and materials. And the results are robust across a range of stimuli, instructions, and tasks (Lucy and Gaskins 2001, 2003).

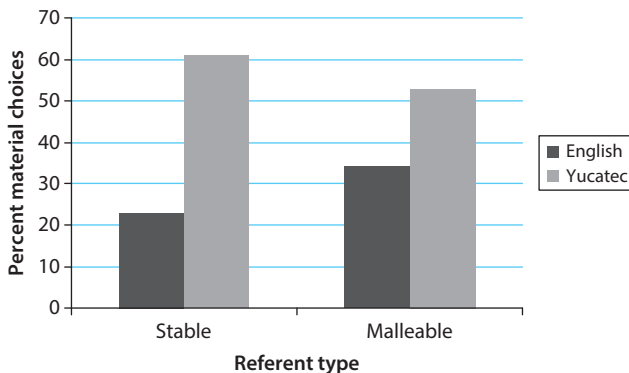


Figure 2. Comparison of English and Yucatec preference for material in similarity judgments for different referent types

3.3 Addressing concerns about language interference

Some critics argue that a nonverbal cognitive response may not be sufficient to show an influence of language on thought because of possible experimental artifacts created by the use of language in the assessment instructions or the direct use of language in cognitive activity.

3.3.1 *Use of language in assessment instructions*

The use of language in the assessment instructions might create an artifact if the verbal instructions cue speakers of different languages to apply distinct language categories to the task. The observed results might then arise not from general cognitive preferences that would occur in ordinary situations but rather from the immediate suggestive influence of the verbal instructions in the assessment. Concerns about effects of particular instructions can be addressed by varying the instructions appropriately. Worries about more general effects of using any verbal instructions can be addressed by developing nonverbal instructions.

For example, in the triad sorting task used in research on number marking, the way the instruction “more like” is rendered in the two languages might imply something different for the two groups, cueing different responses such that speakers of the two languages are not, indeed cannot be, receiving the same instructions. Likewise, in work with children, the meaning of the instructions may change during language acquisition. Inversely, highly specific instructions might lead to responses that conceal actual differences: asking which items are “the same” in the context of slight differences in shape might direct all respondents to material choices. Thus it can be difficult to assure that instructions are equivalent and appropriate precisely because language so readily influences behavior.

One solution is to pre-test alternative verbal forms used in direct instructions to assure a neutral attitude towards the stimuli. The aim is functional equivalence in posing a comparable choice rather than structural equivalence between the instructions. Thus, to approximate the English “more like” instruction with Yucatec speakers we pre-tested in detail several alternatives, not seeking a literal translation of the English but rather instructions that would engender a neutral choice for both groups. Two forms emerged as useful in Yucatec. One was a construction using a loan word from Spanish, which can be glossed as ‘more it’s-equal’, and another purely Yucatec construction, which can be rendered ‘its-twin’. Other pre-tested instructions tended to bias respondents more towards material or shape. No approach can assure full equivalence, but this approach provided a way to minimize verbal bias.

Alternatively, one can design the task so that a response is elicited indirectly such that the instructions do not invoke the classification at stake. Over the years, recognition memory tasks have been the preferred method of pursuing this

option. For example, we have used one procedure where we lay out several sets of triads of the type described above in a grid-like array on a table, ask a speaker look over the array for a minute, and then to turn around and wait for a minute. While they are turned away, we switch two items with the same shape or with the same material and then ask the speaker to turn back around and identify or locate the items that have been changed. Here the key question to the respondent “what changed?” makes no use of “more like” or any appeal to reflective judgment of similarity. As expected, Yucatec speakers had more difficulty recognizing shape switches and English speakers had more trouble recognizing material switches (Lucy and Gaskins in prep.). Although such indirect tasks can be difficult to design, they minimize direct translation problems and provide an important supplement to direct assessments.

Finally, one can design an assessment that does not rely on verbal instructions. This can address concerns about the effects of specific instructions and concerns that using any language in the task might suggest to respondents that they apply the categories of their language. Such an assessment requires nonverbally training a respondent into a task procedure before introducing the relevant contrasts. So, for example, the experimenter can model a triad task by laying out a triad of the type described above but with a transparently obvious match based on exact identity and then making the appropriate choice by moving one alternate next to the original. The experimenter then can reset the triad and indicate through gestures that it is the respondent’s turn to do the same. If the respondent’s choice is not correct, the intended selection can be demonstrated and another, fresh demonstration made, until the procedure is clear. Once speakers can make the expected choices on their own, additional triads can be introduced where the choices are not exactly identical, but differ in one or another peripheral respects (e.g. size, color, etc.), to build up the idea of making an approximate match. And then, finally, triads can be introduced that involve those alternates that force a choice between shape and material. This procedure minimizes verbal cuing or an overall language set. It also allows the assessment procedure to be used with populations with limited verbal abilities (e.g. children, the deaf, language learners, etc.). Results from using such procedures match those produced using verbal instructions, suggesting that the verbal instruction are not shaping the results (Lucy and Gaskins in prep.).

In sum, cognitive patterns associated with language can be identified through a variety of nonlinguistic assessment procedures. Such procedures must not only avoid direct verbal responses but also assure that instructions are equivalent, or are indirect enough not to shape the results, or are operationalized in nonverbal form. Further, since any given assessment may go awry in unforeseen ways, an array of procedures can collectively assure that a particular use of language in the assessment process will not itself account for the results.

3.3.2 *Use of language in cognitive activity*

Speakers responding nonverbally may draw directly on language forms to guide their thinking – employing grammatical categories to guide classification, using lexical labels to help memory, talking silently to themselves as they reason, etc. If the specifics of their language shape their thinking, this can be regarded as evidence for linguistic relativity. However, some researchers discount this as evidence for relativity. They require instead evidence for effects on “non-linguistic representations” (e.g. Li, Dunham, and Carey 2008). They exclude from consideration any effect of language *on* thought that might be due to the direct involvement of language *in* thought.

So why exclude direct uses of language in cognition as evidence for relativity? Those favoring this view note that there are aspects of cognition that precede the phylogenetic emergence of language. Typically included here would be perceptual processes, simple memory, and elementary cause-effect reasoning. These are taken to be basic and universal – and not affected by language. The introduction of language might supplement them, thus shaping eventual behavior, but this leaves unanswered the question of whether these basic processes themselves are affected. In their view, only effects on these basic processes would constitute a linguistic relativity challenging the universality of cognition.

This exclusion on theoretical grounds introduces a fresh methodological requirement into linguistic relativity research. If thought is regarded an integrated *practical activity*, any impact of specific language properties on such activity counts as a relativity effect, regardless of the mechanisms involved. And the methods already described produce abundant and decisive evidence of such relationships. But if thought is decomposed into individual *processing mechanisms* and the only impacts that matter are those affecting the most basic mechanisms that do not require the direct use of language, then methodologically, this “moves the goalposts”. Now we must demonstrate language effects where direct use of language forms or processes in cognition can be precluded by design.

Setting aside the theoretical cogency of this approach, much existing evidence already meets this methodological requirement. For example, in the triad classification task there have been worries that rather than a broad influence of grammatical structure on thinking there might be a narrow, low-level use of lexical labels applicable to the stimuli to help guide choices. Yet using triads where the same lexical item applies to all three items or triads where each item would receive a different lexical label has no effect on the results (Lucy and Gaskins 2001:263–69). Likewise, when speakers sort much larger number of items where no single, object-specific lexical label can mediate across the sets, the differences between the two language groups become even stronger (Lucy and Gaskins 2001:269–72). And when large numbers of stimuli are presented simultaneously, as in the memory

task, such that a lexical labeling (or semantic feature) strategy becomes virtually impossible, effects still appear. Finally, using novel stimuli for which there are no readily available lexical items does not eliminate the effect (Li, Dunham, and Carey 2008: 495) – although such contrived stimuli can create other problems (Lucy and Gaskins 2003: 475). None of this implies that lexical labeling does not occur or that it does not affect thought, but only that it cannot be the sole factor at work since the relationship appears when use of the labels is blocked. Likewise, some have argued that when respondents receive verbal instructions, they draw on global statistical patterns of noun marking to respond (e.g. Li, Dunham, and Carey 2008). But such a strategy seems unlikely when exposed to an array of items at once with indirect instructions or when verbal instructions are absent. In many triad studies then, speakers exhibit language-specific cognitive preferences for shape or material even when language forms are not invoked directly in the task.²

There remains legitimate disagreement about the necessity of this additional methodological requirement. Insofar as the language someone speaks can be shown to have significant relationships to their cognitive activity, there is a relativity effect. And this effect remains completely valid even if another cognitive mechanism remains unaffected by language. The lack of language relationship in one area does not warrant ignoring significant relationships elsewhere. Indeed, a range of language effects on thought now seems likely: some appearing only when responding verbally, others only when language is directly involved in processing, and still others where language has an effect even when not directly involved. And, of course, there are likely activities where language has no detectable effects at all. There is no reason to rule out one or another relation as less worthy of attention. We have moved beyond the question of whether there exist any relations between a specific language and practical cognition to questions about the range and types of effects.

4. Establishing the shaping role of language

Establishing the shaping role of language on thought requires an ensemble of correlational approaches that collectively make other explanations unlikely. Four principle approaches have been used. First, internal assessment design can strengthen a causal interpretation when the predictions from language are highly

2. Standard methods to assess lexical effects on cognition have occasionally been used in relativity research: creating verbal interference (e.g. Lucy and Shweder 1988), training a new category (e.g. Casasanto, Fotakopolou, Pita, Boroditsky in rev.), and testing for hemispheric specialization (e.g. Regier and Kay 2009).

specific, strong, and consistent. Second, comparative studies with additional languages can rule out competing causal factors. Third, developmental studies with children can establish the temporal priority of the language pattern over the cognitive pattern. And finally, studies with second language learners can reveal the cognitive effects of increased exposure to second language patterns.³

4.1 Internal assessment design

The internal design of an assessment can help establish the shaping role of language by limiting competing explanations. A broad association between a single feature of language and some global pattern of thought might arise from any number of factors. By contrast, specific, robust predictions from a cohesive language pattern to specific features of thought will be difficult to account for in other ways and warrant serious attention. Reversing the causal direction requires accounting for the specific cognitive differences independently of language and how they produce the language differences. Appeals to third factors shaping both language and thought must plausibly predict the same specific patterns.

Two strategies help establish the specificity and robustness of an association. First the linguistic pattern and cognitive pattern need to exhibit a *distinctive and tightly linked signature*. The two languages must have structural patterns that contrast in a well-defined way and each language's pattern must link tightly with the locus, shape, and strength of the associated cognitive patterns. Second, the observed associations must exhibit *consistency across assessment procedures*. That is, the associations should be robust across variations in the stimulus materials, cognitive tasks, activities, and sample groups. Such consistency reduces the risk that some artifact of the assessment procedure or context has produced the association and increases the likelihood that the association is strong and general. Hence "Any alternative account of the results will have to account equally well for the overall pattern of results across a diversity of tasks and not just one result in an individual experiment" (Lucy 1992b: 149).

In the study of number marking in English and Yucatec, for example, the prediction was highly specific. The contrast in plural marking displayed a distinctive contingency on lexical semantics as a function of animacy and discreteness. Accordingly, the predicted and observed cognitive performance was tightly linked to this pattern: attending to number for animals, ignoring it for materials, and differing for objects. Further, this same pattern of lexical semantics successfully

3. These strategies encompass the Bradford Hill (1965) causal criteria of Strength, Consistency, Specificity, Plausibility, Temporality, and Exposure Gradient.

predicted that English speakers would attend more to shape with stable objects than for malleable objects and that Yucatec speakers would treat the two kinds of objects in more similar ways. Any alternative account has to confront the full array of results. If the cognitive patterns are taken as shaping language, then we must account for why English and Yucatec speakers differ in their attentiveness to number and shape independently of language. And if we appeal to some environmental factor such as culture or education (e.g. Mazuka and Friedman 2000) to explain this cognitive difference, then it must be shown that these factors actually shape cognition and do so as a function of object type. In short, the tight linkage of signature patterns makes competing accounts less plausible.

Likewise, these number marking studies have proven robust across a variety of assessment procedures. The assessments have explored grammatical and lexical categories, employed verbal and nonverbal instructions, and utilized verbal and nonverbal responses. The tasks have utilized picture stimuli and everyday objects in a variety of combinations and arrays. The assessments have tapped cognitive activities such as attention, classification, similarity, and memory. And the results have held up across genders, generational cohorts, and community locale over a number of years (Lucy and Gaskins in prep.). All this makes it likely the association stems from language rather than assessment artifacts.

4.2 Comparative studies with additional languages

Comparative studies with additional languages provide another important methodological approach that can help rule out competing explanations (Lucy 1992a: 273–76). The typological approach used in the cross-linguistic comparison should also reveal a range of languages of similar and contrasting types (Lucy 1992b: 56–83). If language is indeed shaping cognition, then those speaking languages of similar types should exhibit similar patterns of thought. And to the extent that these other languages are spoken in communities that differ in many ways – environmental, institutional, cultural, etc. – then comparison with these languages creates the possibility of evaluating the contribution, or lack of contribution, of these factors to the cognitive findings. Since a wide range of factors can be expected to influence thought, language, and the assessment process itself, there is no expectation that the match will be perfect in every case, but only that clear traces of the language patterns should consistently appear across a range of languages of the appropriate types.

Turning again to number marking for illustration, Athanasopoulos (2006) compared attentiveness to number as a function of referent type in English and Japanese using the same assessment task and a subset of the same materials described above. Japanese is structurally similar to Yucatec (see Downing 1996),

hence should show a similar pattern of results in contrast to English. This was indeed the case as shown in Figure 3: Japanese speakers respond like Yucatec speakers with a balanced selection of object and material alternates in contrast to English speakers who favored material alternates almost twice as often; both groups rarely selected the animal alternates.

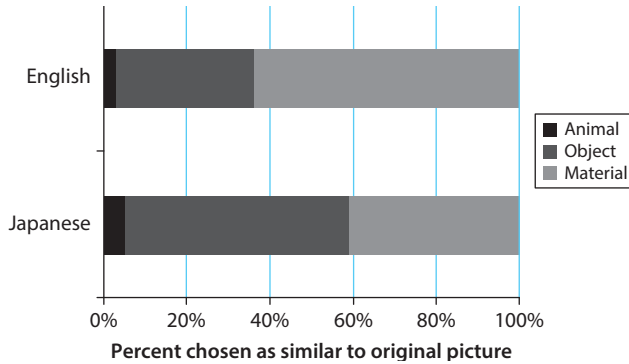


Figure 3. Comparison of English and Japanese similarity judgments with pictures varying numbers of different referent types (drawn from data in Athanasopoulos 2006)

Likewise, there have been several studies of shape versus material preference as a function of referent type in languages similar to Yucatec. For example, as shown in Figure 4, Athanasopoulos (2007) found that Japanese speakers prefer material alternates more for stable referents than do English speakers and these differences are greatly attenuated for malleable referents. Others, using somewhat different materials and instructions,⁴ have found a similar pattern of material preference with stable objects for Japanese (Imai 2000; Imai and Mazuka 2003)⁵ and Mandarin Chinese (Li, Dunham, and Carey 2008).

4. Many studies use Imai and Gentner's (1997) three-way distinction among referent types: complex solids, simple solids, and non-solids. Complex solids conflate function with shape, eliciting fewer material responses. However, nonverbal tasks contrasting non-solids with simple solids (or all solids) produce results similar to those with the original malleable and stable referent types.

5. Mazuka and Friedman (2000) did not replicate the Japanese preference for material. However, they did not control for function matches or provide a within-group contrast of referent types, so it is difficult to evaluate their results.

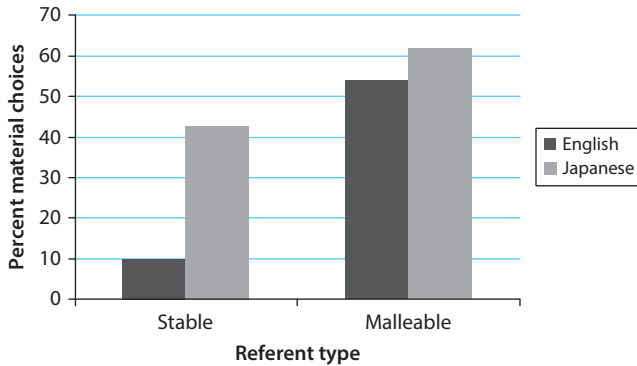


Figure 4. Comparison of English and Japanese preference for material in similarity judgments for different referent types (drawn from data in Athanopoulos 2007)

These comparative studies help rule out competing explanations for the cognitive differences. When similar findings appear in typologically similar languages located in societies with distinct environments, histories, cultures, degrees of modernity, etc. it becomes unlikely that such external contextual factors give rise to the cognitive patterns. Likewise, these comparative assessments extend the range of investigators and assessment procedures, lessening the likelihood that the results are due to poor internal design or investigator bias. Indeed, several of these replications come from researchers with different agendas and, in some cases, from those critical of the idea of linguistic relativity.

4.3 Developmental studies with children

Developmental studies can help establish the temporal priority of language-specific patterns over associated cognitive patterns. Since all normal children begin with similar intellectual and verbal capacities, differences between groups must emerge as part of the process of enculturation. Insofar as language patterns precede the associated thought patterns in development, this suggests that language is likely the shaping factor. Such a developmental sequence cannot decisively rule out that other factors are shaping both, but when conjoined with good assessment design and comparative research, it can help establish causal precedence. It can also provide important information about the timing of and mechanism for the interaction of language-specific patterns with thought.

Developmental research using the triads sorting task described above illustrates this approach (Lucy and Gaskins 2001, 2003). The triads were administered

to American English and Yucatec Maya children at ages seven and nine. For stable objects, as shown in Figure 5, English- and Yucatec-speaking seven-year-olds showed an identical early bias toward shape – rarely choosing material alternates. But by age nine the adult pattern was visible: English-speaking children continued to favor shape, choosing material alternates infrequently, whereas Yucatec-speaking children were now choosing material alternates much of the time. Thus, the same kind of language-group difference found among adult speakers is also found in children by age nine. For malleable objects, as shown in Figure 6, where we expect the two groups to look alike, we find that both English-speaking and Yucatec-speaking seven-year-olds make a substantial number of material choices and that they continue to do so at age nine. Overall, the similarity of response found among adult speakers for referents of this type also appears in children.

Putting these results together, we see that seven-year-olds show clear sensitivity to referent type independently of language group membership. That is, there is a shared cognitive orientation: both groups show a relative preference for material as a basis of classification with malleable objects and relative preference for shape as a basis of classification with stable objects. By contrast, nine-year-olds show differential sensitivity to referent type along adult lines: their classification preferences differ where the languages differ and correspond where the languages correspond. This suggests that language categories increase in their importance for cognition between ages seven and nine, that is, that category patterns in the linguistic structure become important in a new way. Thereafter, Yucatec responses converge towards material choices and English responses towards shape choices

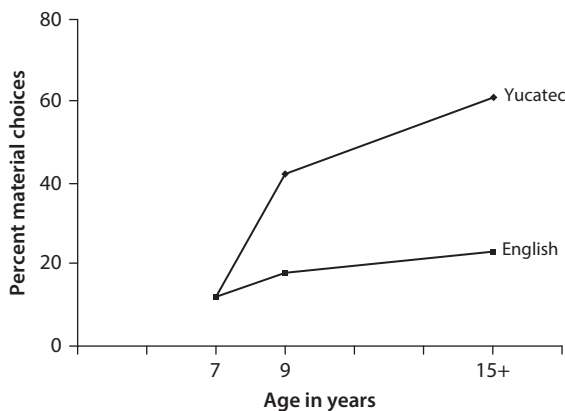


Figure 5. Developmental pattern for English and Yucatec classification preferences with stable objects: material versus shape (from Lucy 2004: 13)

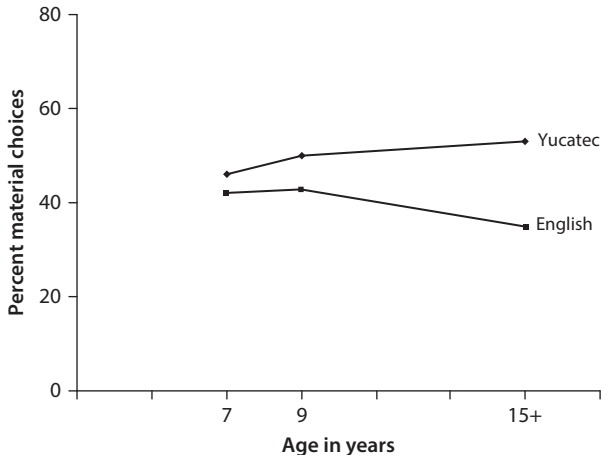


Figure 6. Developmental pattern for English and Yucatec classification preferences with malleable objects: material versus shape (from Lucy 2004: 15)

as a function of the structure of their language. Similar results showing that non-verbal classification preferences in early childhood give way to distinctive patterns in adulthood in line with language patterns have been found for Japanese (Imai 2000) and Mandarin Chinese (Li, Dunham, and Carey 2008).

Crucially, from a methodological point of view, the central components of the verbal number marking system have been in place for many years for these children. Just as English-speaking children have substantial command of plurals by age seven, so too do Yucatec-speaking children have substantial command of numeral classifiers by this age. Children in both groups reliably comprehend and use the appropriate forms and will judge constructions misusing them as faulty. This is not to say that children in either group have yet mastered all the details of their language, but that there is no question whatsoever that the basic structural characteristics of the number marking system in each language are firmly in place before the cognitive patterns appear. The details of the mechanism leading to this cognitive shift remain unclear, but it clearly depends on exposure to the language patterns rather than the other way around.⁶

6. Differential exposure to language can also be assessed by using deaf subjects. However, since access to deaf individuals is not always possible and evaluation of deaf performance presents its own special difficulties, the use of such special populations has not been included here as a general method.

4.4 Studies with second language learners

Studies with bilingual speakers⁷ can be used to help evaluate the role of language patterns in shaping cognition. If speaking a first language affects thinking, then speaking a second language may also have effects on cognition. Such effects may depend not only on the mere presence of a second language, but also on the nature and extent of mastery of the second language in relation to the first, its structural and functional similarity to the first, and the psychological and social coordination of the two by the speaker. Recently, there has been interest in using linguistic relativity research to explore these bilingual processes (Cook and Bassetti 2011; Han and Cadierno 2010; Jarvis and Pavlenko 2008; Pavlenko 2011). Our interest is in the reverse direction, namely, the extent to which the performance of bilingual speakers can be used methodologically to inform relativity research. In particular, assessing the impact of second language mastery on cognition provides another very precise way to verify the causal priority of language over cognition or other factor. And studying interference in second language learning itself can provide a way to evaluate the importance of quantitative (statistical) exposure versus qualitative (structural) understanding in how language and thought relate.

4.4.1 *Impact of second language mastery on cognition*

Although mastery of a second language is known to produce general effects on cognition (e.g. in metalinguistic awareness, executive function, creativity, etc.), less is known about how specific meaning structures used in a second language might affect thinking (Bassetti and Cook 2011: 143). If exposure to a different language moves cognitive performance in the direction of monolingual speakers of the second language, this suggests that language is a shaping force, so long as one can rule out competing factors that might have shaped both language and thought. Such learning effects might arise as a simple function of amount of exposure or as a more complex function of the qualitative mastery of specific structures.

The cognitive effects of learning a second language with a contrasting structure can be illustrated in the area of number marking. Athanasopoulos (2006) used the picture-sorting task described above to explore the impact on Japanese speakers of learning the English plural marking pattern. As shown in Figure 7, he found that Japanese bilingual responses moved in the direction of the monolingual English speakers and that more advanced learners showed a stronger effect

7. The term bilingual is employed here to include a wide variety of second language users, including multilingual speakers, insofar as these have been used in relativity research. For discussions of the complexities of defining bilingualism, see Hoffman (1991) and Grosjean (1998).

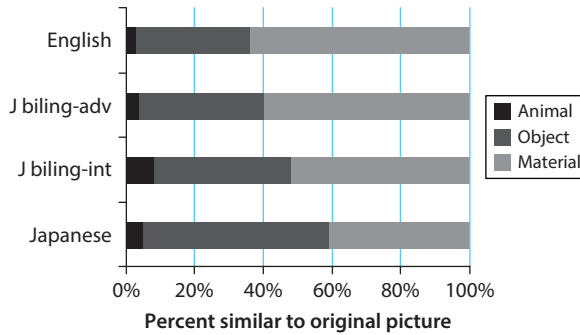


Figure 7. Comparison of Japanese monolinguals, intermediate and advanced Japanese-English bilinguals, and English monolinguals on similarity judgments with pictures varying numbers of different referent types (drawn from data in Athanasopoulos 2006)

than intermediate learners. He cannot say for certain whether these responses “are purely non-linguistic or whether such processes are guided by implicit verbal descriptions”, but he concludes that the results clearly “support the view that language influences cognitive dispositions by directing speakers’ attention to specific features of stimuli” (2006: 95).

Likewise, Athanasopoulos (2007, 2011) assessed bilingual performance on the triad-sorting task designed to test for material versus shape preference. As shown in Figure 8, Japanese learners of English moved toward the English pattern by attending less to material for stable objects (his “count” objects). But there was no

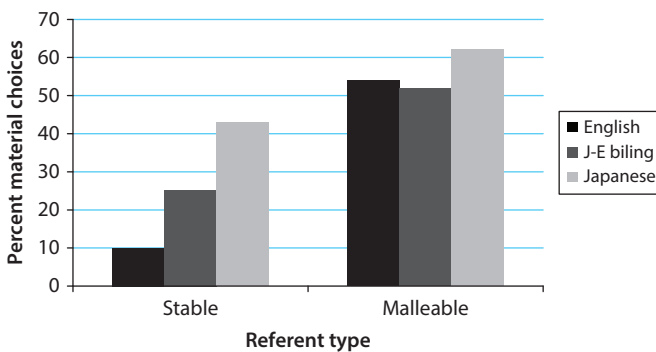


Figure 8. Comparison of Japanese monolingual, Japanese-English bilingual, and English monolingual preference for material in similarity judgments for different referent types (drawn from data in Athanasopoulos 2007)

effect for malleable objects (his “mass” substances) where English and Japanese do not differ. In short, these Japanese learners of English perform in ways that show a clear influence of English number marking patterns precisely where those patterns differ from Japanese.

Studies with bilinguals raise anew the issue of whether the language used in the task might influence the results. Indeed, Grosjean (1998) and others have shown the language modality active in a bilingual can be affected by the assessment conditions. However, Athanasopoulos (2007, 2011) explicitly controlled for language by dividing his bilingual sample such that one group received instructions in English from a non-Japanese administrator and another received instructions in Japanese from a Japanese native speaker. He found no significant differences between the two groups in nonlinguistic performance. Instead, the best predictor of performance was the degree of second language proficiency, which “suggests that language may affect habitual thought at a deeper, more permanent level” (2007: 698).⁸

Another concern that reappears with bilinguals is that exposure to various cultural factors might account for performance differences being attributed to language. This is a legitimate concern given the close relation between language and culture, especially in the lexicon and in patterns of use. However, Athanasopoulos (2007, 2011) also explored this issue and found that when he controlled for second language proficiency, as measured by a general test and a targeted grammaticality judgment task, the length of stay in the second language cultural environment, did not predict differences in nonlinguistic performance. Length of cultural exposure mattered through its effect on proficiency.⁹

These bilingual studies provide a level of methodological precision “impossible in monolingual adult speakers” (Athanasopoulos 2007: 698). The cognitive effects appear at points of structural contrast and they increase with degree of exposure. Further, they are not dependent on the language used in the assessment or on general cultural exposure. These results make it likely the associations arise from language rather than from thought or some other contextual factor. And they make it unlikely the associations arise from direct cues in the language of assessment. Currently, however, they do not distinguish effects due to gradient statistical exposure from those due to conceptual reorganization.

8. Barner, Inagaki, and Li (2009) found that bilingual Mandarin speakers altered their word extensions depending on the language of the task, although their study lacks a monolingual Mandarin baseline. Cook, Bassetti, Kasai, Sasaki, and Takahashi (2006) however do not find such results with Japanese bilinguals.

9. Cook et al. (2006) found some effect on word extension for Japanese bilinguals with length of residence in the second language environment.

4.4.2 *Interference in second language learning*

Interference arises when the categories of a language create difficulties in learning another language. From one point of view, such interference would not represent a linguistic relativity effect because the effect is on language activity, not on cognition generally. But from another point of view, such interference is not language internal, since it does not concern the influence of language patterns on the use of the same language but rather the ability to master another independent language. From this vantage, the second language is simply another part of experienced reality. And insofar as the difficulties of acquiring it arise not merely from lack of exposure, but also from the presence of a prior language-shaped sensibility about experience, they can be regarded as a kind of relativity effect.

I have used the term *semantic accent* for the sort of interference effect that is mediated by a speaker's pre-existing language categories and associated view of reality, whether this effect appears among ordinary language learners or professional linguists (Lucy 2003, 2004, 2010, 2011). However, many researchers theorize such interference effects by using Slobin's formulation of *thinking for speaking*, "a special form of thought that is mobilized for communication" in our native language "while we are speaking" and may therefore affect "one's mastery of the grammatical categories of a foreign language" (1987: 436). In considering second language effects, Slobin puts special emphasis on categories that "cannot be experienced directly in our perceptual, sensorimotor, and practical dealings with the world" but which language alone requires us to make (1996: 91). By contrast, my use of the term *accent* also seeks to capture those aspects of linguistic categories with referential entailments and those with continuities across verbal and nonverbal modalities. If language learning depends not only on statistical exposure and functional entrainment, but also on grasping the conceptual structure of a language and the associated view of reality, study of such learning provides a methodological avenue into assessing the role of structural differences in giving rise to relativity.

One number marking study that explored interference as a form of relativity, albeit within a "thinking for speaking" framework, is Han's (2010) longitudinal case study of an adult Chinese speaker. Despite years of experience with English and a high level of attainment, this speaker still exhibited patterns of difficulty with plurals and articles. Over the eight years of Han's study, in both naturalistic production and systematic translation tasks, these difficulties persisted despite increased exposure and use. The speaker's errors stemmed from applying English number marking forms using a heuristic based on marking of "specificity" in Chinese, which diverges from English number and definiteness marking. Thus, noun phrases that contain an explicit quantifying expression (e.g. *two*, *several*, *much*, etc.) were regarded as specific and hence pluralized. Utterances lacking

such an expression, for example, generic constructions, were systematically left unmarked. Likewise, when Chinese would include a demonstrative or number word, the construction was construed as specific, and the appropriate article applied, but not otherwise. So his usage “patterned largely after his L1 ... in spite of his long-term experience with the target language” (Han 2010: 181).¹⁰ These heuristics allowed him to approximate the correct patterns statistically yet fall short of genuine mastery of the structures that give rise to them.

Overcoming this first language interference requires a “conceptual *restructuring*” that “entails not just mapping individual forms onto individual meanings, but rather, integrated mapping of a cohort of forms” (Han 2010: 178). This cohort of forms in English includes a lexicon conceptually structured with respect to quantification so as to encode diverse referents in ways that interact with the requirements of formal number marking and definiteness. Tracing the degree of restructuring to nonlinguistic cognitive performance provides a way to distinguish the effects of structural mastery from statistical approximation in linguistic relativity, something that cannot be accomplished using global measures of proficiency alone. Thus studies of structural interference can not only provide direct evidence for relativity effects in language learning, they can also clarify the structural factors shaping learners’ nonlinguistic cognition.

5. Conclusion

The methodological challenges to investigating linguistic relativity emerge directly from the internal logic of the proposal, namely, that the structural properties of each language embody a particular interpretation of reality that influences patterns of thinking. The primary challenges thus concern language structure, cognitive interpretation, and direction of influence.

Languages not only render experience into categories for the purposes of reference, they also enrich that experience by bringing to it the structural meanings that have emerged from the dialectic of sense and denotation (Lucy 2010). The first methodological challenge is to identify, characterize, and contrast these structures of referential meaning. Studies of linguistic relativity that ignore or evade structure, for example by focusing on a set of lexical items in isolation, miss the heart of the proposal. The concern with structure also guides the development of cognitive assessment and establishing direction of influence: the patterns of

10. Japanese learners of English show similar difficulties with quantification (Jarvis and Pavlenko 2008: 138).

cognitive activity should tightly match the language structure, appear in structurally similar languages, follow the emergence of the language structures in child development, and interact in predictable ways with the learning of structurally distinct languages. Thus, ignoring language structure also forfeits much of the methodological power necessary to show how a formal semantic structure yields substantive cognitive entailments that have a functional impact on speakers.

Drawing out the cognitive interpretations latent in language structure and designing a way to test them requires recognizing the referential commitments entailed in language structures and then imagining how those commitments might appear in other cognitive activities. An incorrect analysis of the language, whether arising from semantic accent or associated view of reality, can lead to mistaken or skewed cognitive predictions. Indeed, even understanding the structural logic of the language, we may still have difficulty imagining the alternative way of viewing reality. And once we have a prediction, all the usual difficulties remain: developing an assessment that contrasts relative rather than absolute performance, across individuals, controlling for language use in the task, all in a way that is culturally valid and fair. In essence, to the extent that the relativity proposal is valid, we must work through our own semantic accent and understanding of reality just to get to a point where we can grapple with the challenges of assessment design.

Finally, establishing the direction of influence requires an integrated ensemble of approaches that varies key elements in strategic ways. The language analysis should produce a distinctive set of interlocking predictions. The cognitive assessments must test those predictions with a variety of materials, instructions, and tasks to assure that results are not due to artifact. These results must also track the language predictions within a language group and across populations speaking different languages. Further, the language patterns should precede the cognitive patterns among children acquiring language and show effects as a function of among those learning a second language. Such an ensemble of methods allows us to rule out various competing hypotheses and collectively make an alternative explanation for the whole set of results unlikely. Thus a demonstration of a correspondence between a language form and some cognitive activity can only really be persuasive when it forms part of an ensemble of supporting findings.

Research meeting these challenges is now appearing, as the examples provided here demonstrate. This research makes clear that the diverse interpretations of experience in languages do influence thought. The task now before us is to use these methods to assess the various types and full scope of these effects, as well as their power and limits.

References

- Athanasopoulos, Panos. 2006. Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition* 9: 89–96. DOI: 10.1017/S1366728905002397
- Athanasopoulos, Panos. 2007. Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes* 22: 689–699. DOI: 10.1080/01690960601049347
- Athanasopoulos, Panos. 2011. Cognitive restructuring in bilingualism. In A. Pavlenko, ed., *Thinking and Speaking in Two Languages*, 29–65. Bristol, Buffalo, & Toronto: Multilingual Matters.
- Barner, David, Shunji Inagaki, & Peggy Li. 2009. Language, thought, and real nouns. *Cognition* 111: 329–344. DOI: 10.1016/j.cognition.2009.02.008
- Bassetti, Benedetta & Vivian Cook. 2011. Relating language and cognition: The second language user. In V. Cook & B. Bassetti, eds., *Language and Bilingual Cognition* 143–190. Abingdon: Routledge, Taylor, and Francis Group.
- Benveniste, Emile. 1971. The nature of the linguistic sign. In *Problems in General Linguistics*, 43–48. Miami, FL: University of Miami Press.
- Brown, Roger, & Eric Lenneberg. 1954. A study in language and cognition. *Journal of Abnormal and Social Psychology* 49: 454–62. DOI: 10.1037/h0057814
- Casasanto, Daniel, Olga Fotokopolou, Ria Pita, Lera Boroditsky. In prep. How deep are effects of language on thought? Time estimation in speakers of English and Greek. *Cognition*.
- Cook, Vivian J., & Benedetta Bassetti. 2011. *Language and Bilingual Cognition*. Abingdon: Routledge, Taylor, and Francis Group.
- Cook, Vivian J., Benedetta Bassetti, Chise Kasai, Miho Sasaki, & Jun Arata Takahashi. 2006. Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism* 10: 137–152. DOI: 10.1177/13670069060100020201
- Downing, Pamela. 1996. *Numeral Classifier Systems. The Case of Japanese*. Amsterdam & Philadelphia: Benjamins. DOI: 10.1075/sidag.4
- Grosjean, F. 1998. Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1: 131–149. DOI: 10.1017/S136672899800025X
- Han, ZhaoHong, & Teresa Cadierno, eds. 2010. *Linguistic Relativity in SLA: Thinking for Speaking*. Clevedon: Multilingual Matters.
- Han, ZhaoHong. 2010. Grammatical morpheme inadequacy as a function of linguistic relativity: A longitudinal case study. In Z. Han & T. Cadierno, eds., *Linguistic Relativity in SLA: Thinking for Speaking*, 154–182. Clevedon: Multilingual Matters.
- Hill, Austin Bradford. 1965. The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine* 58: 295–300.
- Hoffman, Charlotte. 1991. *An Introduction to Bilingualism*. London: Longman.
- Imai, Mutsumi. 2000. Universal ontological knowledge and a bias toward language-specific categories in the construal of individuation. In S. Niemeier & R. Dirven, eds., *Evidence for Linguistic Relativity*, 139–160. Amsterdam & Philadelphia: Benjamins.
- Imai, Mutsumi, & Dedre Gentner. 1997. A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition* 62: 169–200. DOI: 10.1016/S0010-0277(96)00784-6

- Imai, Mutsumi, & Reiko Mazuka. 2003. Reevaluating linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In D. Gentner & S. Goldin-Meadow, eds., *Language in Mind: Advances in the Study of Language and Thought*, 429–464. Cambridge, MA: MIT Press.
- Jakobson, Roman. 1960. Closing statement: Linguistics and poetics. In T. Sebeok, ed., *Style in Language*, 398–429. Cambridge: MIT Press.
- Jarvis, Scott, & Aneta Pavlenko. 2008. *Crosslinguistic Influence in Language and Cognition*. New York: Routledge.
- Li, Peggy, Yarrow Dunham, & Susan Carey. 2009. Of substance: The nature of language effects on entity construal, *Cognitive Psychology* 58: 487–524. DOI: 10.1016/j.cogpsych.2008.12.001
- Lucy, John A. 1992a. *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis* [Studies in the Social and Cultural Foundations of Language 12]. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620843
- Lucy, John A. 1992b. *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis* [Studies in the Social and Cultural Foundations of Language 13]. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620713
- Lucy, John A. 1994. The role of semantic value in lexical comparison: Motion and position roots in Yucatec Maya. *Linguistics* 32: 623–656. DOI: 10.1515/ling.1994.32.4-5.623
- Lucy, John A. 1996. The scope of linguistic relativity: An analysis and review of empirical research. In J.J. Gumperz & S.C. Levinson, eds., *Rethinking Linguistic Relativity*, 37–69. Cambridge: Cambridge University Press.
- Lucy, John A. 1997a. Linguistic relativity. *Annual Review of Anthropology* 26: 291–312. Palo Alto: Annual Reviews Inc. DOI: 10.1146/annurev.anthro.26.1.291
- Lucy, John A. 1997b. The linguistics of ‘color’. In C. Hardin & L. Maffi, eds., *Color Categories in Thought and Language*, 320–346. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511519819.015
- Lucy, John A. 2003. *Semantic Accent and Linguistic Relativity*. Conference on Cross-linguistic Data and Theories of Meaning, Catholic University of Nijmegen, The Netherlands, 20 May.
- Lucy, John A. 2004. Language, culture, and mind in comparative perspective. In M. Achard & S. Kemmer, eds., *Language, Culture, and Mind*, 1–21. Stanford: Center for the Study of Language and Information Publications.
- Lucy, John A. 2010. Language structure, lexical meaning, and cognition: Whorf and Vygotsky revisited. In B. Malt & P. Wolff, eds., *Words and the Mind: How Words Capture Human Experience*, 268–288. Oxford: Oxford University Press.
- Lucy, John A. 2011. Language and cognition: The view from anthropology. In V. Cook & B. Bassetti, eds., *Language and Bilingual Cognition*, 43–68. Abingdon: Routledge, Taylor and Francis Group.
- Lucy, John A., & Suzanne Gaskins. 2001. Grammatical categories and the development of classification preferences: A comparative approach. In S. Levinson & M. Bowerman, eds., *Language Acquisition and Conceptual Development*, 257–283. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620669.011
- Lucy, John A., & Suzanne Gaskins. 2003. Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner & S. Goldin-Meadow, eds., *Language in Mind: Advances in the Study of Language and Thought*, 465–492. Cambridge, MA: MIT Press.
- Lucy, John A., & Suzanne Gaskins. In prep. Language development and linguistic relativity. *Language and Cognition: A Journal of Language and Cognitive Science*.

- Lucy, John A., & Richard Shweder. 1988. The effects of incidental conversation on memory for focal colors. *American Anthropologist* 90: 923–31. DOI: 10.1525/aa.1988.90.4.02a00100
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139165570
- Lyons, John. 1977. *Semantics. Volume 1*. Cambridge: Cambridge University Press.
- Mazuka, Reiko, & Ronald S. Friedman. 2000. Linguistic relativity in Japanese and English: Is language the primary determinant in object classification? *Journal of East Asian Linguistics* 9: 353–377. DOI: 10.1023/A:1008356620617
- Pavlenko, Aneta. 2011. *Thinking and Speaking in Two Languages*. Bristol, Buffalo, & Toronto: Multilingual Matters.
- Regier, Terry, & Paul Kay. 2009. Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences* 13: 439–446. DOI: 10.1016/j.tics.2009.07.001
- Slobin, Dan. 1987. Thinking for speaking. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 13: 435–445.
- Slobin, Dan. 1996. From “thought and language” to “thinking for speaking”. In J. J. Gumperz & S. C. Levinson, eds., *Rethinking Linguistic Relativity*, 70–96. Cambridge: Cambridge University Press.
- Whorf, Benjamin. 1956. The relation of habitual thought and behavior to language. In J. B. Carroll, ed., *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*, 134–159. Cambridge, MA: MIT Press. (Original work published 1941.)