

9 Grammatical categories and the **development** of classification preferences: a comparative approach

John A. Lucy and Suzanne Gaskins

University of Chicago and Northeastern Illinois University

1 Introduction

The defining characteristic of the human species is its culture-bearing capacity whereby very similar biological organisms develop and sustain extraordinarily diverse behavioral repertoires. Research on human behavior, then, must necessarily concern itself with the scope and significance of this diversity and the process of its **development** in childhood. However, contemporary psychological research often assumes instead a homogeneity of repertoire and of underlying psychological function – coupled with a concomitant assimilation of the psychological to the biological – and neglects the process of culture **acquisition**. Theories and methods developed from such a perspective neither incline their proponents to developmentally oriented comparative research nor provide a set of concepts and tools adequate to undertake it.

The reality of cultural diversity requires us to adopt a comparative perspective from the beginning as part of a coherent effort to account for the actual range of human psychological functioning and the process of its formation. Such a coherent effort demands more than simply testing whether our local findings generalize to other cultures or looking for a specific, naturally occurring equivalent for some odd manipulation we cannot perform within our own culture for one reason or another. Rather, it requires taking seriously the proposal that the human developmental process is designed to support diversity in behavioral outcome and that psychological research programs must take account of this from the outset if they are to produce adequate methods and theories. Taking a comparative perspective from the outset involves the following general steps: documenting ethnographically the range and patterning of behavioral diversity, formulating and testing for tangible psychological implications of the diversity manifest in various cultures, and then exploiting the diversity itself in order to uncover the nature of the psychological mechanisms and developmental processes at work. The ultimate objective is a body of psychological theory and method

founded on an understanding of the range of cultural diversity and therefore adequate to the full range of human experience.

Nowhere is the neglect of diversity and need for a comparative perspective greater than in the study of the relationship between **language** and thought. Psychological research on this topic since the late 1950s has generally ignored the potential cognitive significance of **language** diversity in children and adults. Yet the diversity of the world's languages is obvious and compelling: no child or adult speaks a generic **Language** understood by all, but rather one or more particular languages shared within a community. Just as theories of **language** use and **acquisition** must account for this diversity, so too must theories of the relationship between **language** and cognition. We need to know not only how a single biological organism can sustain such a diversity of languages, but also whether and how that diversity has an impact on intellectual functioning and when and how that impact takes shape. These questions simply cannot be addressed within a single **language** group nor dismissed across the board because of the presence of some linguistic or cognitive universals at an early age.

At present we know little about whether the particular **language** we speak influences the way we think, despite the obvious bearing of this issue on our theoretical understanding of mind and culture, on research methodology, and on public policy. Direct empirical research on the topic scarcely existed until the early 1990s and what did exist had often been poorly done and yielded ambiguous results (Lucy 1992a). In place of empirical research, the literature has been filled with a wide variety of speculative answers that inevitably confirm the initial theoretical predilections of the analyst. Such speculation has flourished because it has never had to accommodate to an accepted body of rigorous empirical findings about the relevant cognitive performance of speakers of diverse languages. People have freely extrapolated from research data on our own **language** or some telling personal experience with another. Additional speculation of this type seems unlikely to resolve whether structural diversity among languages has a significant impact on cognition.

The remedy for this situation lies in developing a viable empirical approach to the issue and using it. An analytic review of existing empirical research (Lucy 1992a) yielded four requirements that should be met by such research: (1) it should be comparative from the outset, presenting contrastive data on two or more **language** communities; (2) the comparison should utilize an external "reality" as the metric or standard for calibrating the content of both linguistic and cognitive categories; (3) the **language** analysis should concern one or more categories of reference having general significance in the languages; and (4) **language**-based cognitive predictions should be evaluated in light of the actual nonverbal performance of individ-

ual speakers.¹ These requirements are not especially demanding given the nature and importance of the question, but at the time of the review no existing study met all four requirements.² Therefore it was necessary to take a second step, namely, illustrating the approach by developing a new concrete case study. A study meeting these requirements (Lucy 1992b) explored the relation between **language** and thought among adult speakers of American English and of Yucatec Maya, an indigenous **language** spoken in southeastern Mexico. The study compared grammatical number marking patterns in the two languages and then tested for associated cognitive differences on classification and memory tasks involving picture and object stimuli.

In this chapter, we draw on this previous research as well as more recent work on both adults and children to illustrate what psychological research on **language** and thought needs to include in order to meet the four criteria listed above. It begins with the reality of **language** diversity, explores its significance for adult functioning, and traces its **development** during childhood. The emphasis throughout will be on the general logic and broader relevance of adopting a comparative strategy, that is, how one moves from the account of behavioral diversity through the identification of psychological correlates to a more viable theory of human functioning and **development**. In the sections to follow we will show how the grammars of the two languages differ in fundamental ways, that these differences correspond to adult cognitive preferences (that is, that there are two distinct developmental outcomes depending on the **language** spoken), and, finally, when and how these distinct patterns emerge in **development**. We conclude by revisiting the arguments about comparative methodology in terms of the particular substantive findings of the study.

2 Cultural background and **language** contrast

The Maya who form the comparison group here are subsistence farmers living in a small community in the scrub jungle of the eastern Yucatan peninsula. Descendants of the peoples who built the well-known monumental architecture of the region, today they raise corn and a variety of other crops in a slash-and-burn agricultural regime. The material culture, social organization, and ritual life of the community can best be described as traditionally “Yucatecan” – that is, they derive from a long complex accommodation between pre-Columbian Mayan, Colonial Spanish, and modern Mexican influences. Isolation and poverty combined to ensure that the village remained very traditional even by local standards well into the 1970s at which time a road was built into the community. Thereafter the pace of change quickened both because of the arrival of new influences in the form of medical clinics, agricultural projects, packaged foods, electricity, and

modern appliances such as TVs and refrigerators, and because of easier access to the outside world, in particular to the employment and market opportunities of the nearby tourist resort of Cancun.

Despite these recent cultural changes, Yucatec Maya remains the **language** of everyday life both within the community and in nearby trade towns. Most women and children are monolingual and a substantial number of men are too. Even the school curriculum is fully in Yucatec for the first three years with many children dropping out after this point. Spanish, the **language** of the outside world, is spoken increasingly often by young adults, especially men, but few individuals have anything like full fluency.

Yucatec and English differ in their nominal number marking patterns, the grammatical focus of the present study. First, the two languages contrast in the way they signal plural for nouns. English speakers *obligatorily* signal plural for a *large number* of their nouns whereas Yucatec speakers *optionally* signal plural for a comparatively *small number* of their nouns. Specifically, English speakers typically mark plural for nouns referring to animate entities and ordinary objects but not for amorphous substances (e.g. *sugar, mud*, etc.). Yucatec speakers sometimes mark plural for animate entities but only occasionally mark it for any other type of referent. These patterns of semantically contingent plural marking are consistent with well-attested crosslinguistic typological patterns (see Lucy 1992b: ch. 3).

Second, the two languages contrast in the way they enumerate nouns and this contrast derives from a deep underlying difference between the two languages. English numerals directly modify their associated nouns (e.g. *one candle, two candles*) whereas Yucatec numerals must 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'iit kib* 'one long thin candle,' *ká'a-tz'iit kib* 'two long thin candle'). Numeral classifiers of this type are a well-known grammatical phenomenon and occur in a wide variety of languages throughout the world, perhaps most notably in the languages of Asia – Chinese, Japanese, Thai, etc.

Since many classifiers have to do with shape or form, one common interpretation of them is that they represent a special emphasis on these concepts in a language's semantics. Such a view would be more plausible if the classifiers were optional, occurred in many morphosyntactic contexts, and appeared only in a few languages. But in fact they are obligatory, confined to a single morphosyntactic context, and are fairly common among the world's languages – all of which suggests that they represent an indispensable solution to a formal referential difficulty characteristic of languages of a certain morphosyntactical type.

So why have numeral classifiers? What problem do they solve? The need for them reflects the fact that *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. glossed as '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 of objects that we would call candles (as well as to other wax things). Once one understands the quantificational neutrality of the noun it becomes obvious that one must specify a unit when counting – that is, provide a classifier, since expressions such as *one wax* wouldn't 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*). Where our English nouns are quantificationally neutral like those of the Yucatec, we use the functional equivalent of a classifier construction ourselves: *one cube of sugar*, *one clump of dirt*, and so forth.

For those Yucatec lexical nouns with no close neutral analogue in English, it can be quite difficult to render the exact Yucatec sense. Actually, even for those cases where English has a somewhat equivalent neutral noun, the translation can only be approximate since in Yucatec such nouns are not in systematic contrast with a set of nonneutral nouns as they are in English. The substance reading we give *kib'* when we translate it as *wax* in order to try to make the syntactic neutrality clear isn't quite right semantically because there is no contrasting word for 'candle' in Yucatec. *Kib'* can mean either 'candle' or 'wax' and to choose either as the translation is to lose part of its regular referential meaning.

The patterns of plural marking and numeral modification just described are closely related and form part of a unified number-marking system. Hence, languages with rich, obligatory plural marking such as Hopi tend not to have numeral classifiers and those with a rich, obligatory use of numeral classifiers such as Chinese tend not to have plural marking. Moreover, for languages with both types of marking, the lexicon tends to be internally divided such that nouns requiring plural marking with multiple referents tend not to require classifiers for counting, and those requiring classifiers for counting tend not to require plurals when used with multiple referents (Lucy 1992b: ch. 2).

3 Cognitive correlates

3.1 *Original triads study*

To assess whether traces of these verbal behavior patterns appear in speakers' cognitive activities more generally, we need first to draw out the impli-

cations of these different grammatical patterns for the general interpretation of experience. We have seen that English encodes quantificational unit (or some equivalent) in a large number of its lexical nouns whereas Yucatec does not. But it is difficult to form a single generalization about the concrete denotational meaning value of such patterns because the kind of unit presupposed varies across the spectrum of lexical noun types both within and across languages. But for those lexical nouns where the contrast between the English and Yucatec is maximal – those referring to discrete objects³ – certain regularities exist from which denotational implications can be drawn.

The quantificational unit presupposed by English nouns of this type is usually the shape of an object. Hence use of these English lexical items routinely draws attention to the shape of a referent insofar as this is the basis for incorporating it under some lexical label and assigning it a number value. Yucatec nouns 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 a referent insofar as this is the basis for incorporating it under some lexical label. If these linguistic patterns translate into general cognitive sensitivity to these properties of referents of the discrete type, then *Yucatec speakers should attend relatively more to the material composition of objects (and less to their shape), whereas English speakers should attend relatively less to the material composition of such objects (and more to their shape).*

This prediction was tested with adult speakers from both languages (Lucy 1992b:136–141⁴). They were shown triads of naturally occurring objects familiar to both groups. 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. Informants were shown eight such triads and asked to decide for each pivot which of the two alternates was most like it. So, for example, speakers were shown a small cardboard box (of the type used for holding cassette tapes) as a pivot and asked whether it was more like a small plastic box of roughly the same size and shape or more like a small piece of cardboard about the size of a matchbook. The expectation was that English speakers would match the pivot to the box and Yucatec speakers would match it to the cardboard. This prediction was strongly borne out across the set of eight triads when shown to a sample of Yucatec and English speakers: eight out of ten Yucatec speakers favored the material alternates, and twelve out of thirteen of the English speakers favored the shape alternates ($p < .0007$, one-tailed Fisher exact test). Table 9.1 shows the results of this original study. Clearly the two groups classify these objects differently and in line with the expectations based on the underlying lexical structures of the two languages. Notice that both pat-

Table 9.1. *Number of English and Yucatec speakers preferring shape or material as a basis for object classification: original triad task*

Language group ^a	Classification preference	
	Shape	Material
English ($n = 13$) ^b	12	1
Yucatec ($n = 10$) ^c	2	8

Notes:

^a English versus Yucatec: $p < .0007$, one-tailed Fisher exact.

^b English versus Chance: $p < .005$, $n = 13$, one-tailed binomial test ($P = .56$ for Shape, $Q = .44$ for Material).

^c Yucatec versus Chance: $p < .025$, $n = 10$, one-tailed binomial test ($P = .56$ for Shape, $Q = .44$ for Material).

Source: After Lucy 1992a:141, table 34.

terms of classification are reasonable and neither can rightly be described as superior to the other.

3.2 *Recent extensions*

Despite the positive results, the design of this task left open several issues. First and most importantly, a number of unintended responses are possible. For example, in the triad mentioned earlier, an apparent shape or “box” match might actually arise from attending to the similarity of function as closable containers or to their size in comparison to the small piece of cardboard or to their wholeness compared to the fragment of cardboard; likewise, an apparent material or “cardboard” match might actually arise from attending to similarity of color or the greater familiarity of the two cardboard objects in contrast to the somewhat novel plastic box. Some of these alternative readings can be seen as shape and material responses in disguise, and it is likely that across the entire set of triads such unintended responses did not greatly affect the results, but it seemed desirable to improve control over these factors.

Second, the same cognitive patterns might not appear with more complex reasoning and memory tasks. It seemed worth exploring the effect of increasing the task complexity and/or difficulty. Additionally, it is possible that the simple directness of the contrast presented in the triads elicits a lexical labeling strategy: so, for example, English speakers seeing the triad just described will say “box” to themselves when they see the pivot and so

choose the box alternate.⁵ Although there would still be a **language** effect under this interpretation, it would stem more from the immediate lexical label for the particular stimulus item rather than from the broader grammatical pattern. To guard against this, the original set of triads was counter-balanced for the presence of simple lexical labels at least in Yucatec (i.e. some of the probes shared lexical labels with the substance choice and some did not), but it seemed wise to control more carefully the possibility of immediate lexical influences. A more complex task where the contrast at stake might be less obvious and less subject to the influence of immediate item labeling seemed worth undertaking.

Thus, to improve control over the stimulus materials and increase task complexity, two new assessment procedures were undertaken. The first assessment procedure replicated the triads sorting study with a more rigorously controlled set of stimulus materials. The second assessment procedure used the same sorts of stimulus materials but involved a new, more complex classification task. Both procedures were presented to twelve English and twelve Yucatec adult speakers. For full details on these and related studies, see Lucy & Gaskins (in preparation).

3.2.1 Controlled stimulus materials Many new triad types were constructed to explore with greater precision the competing bases of classification. The focus here will be on the set of twenty-one triads that we used most extensively. Each triad consisted of three objects that were deliberately selected to be familiar to both **language** groups so as to insure cultural interpretability. The dimensions of control for the new triads are detailed below, along with the results of the choices that English and Yucatec speakers made for these triads.

3.2.1.1 Control of color and size

Color does not typically find its way into grammatical distinctions, and size does so only in relatively limited ways (e.g. diminutives), although many if not most languages have forms capable of referring to these qualities. Yet these two qualities can be very salient cognitively, so we wanted to assure ourselves that they were not affecting the results. However, it proved very difficult to control color and size with naturally occurring objects familiar to both groups since many such objects do not come in a variety of colors and sizes. In particular, there is a tight linkage between color and material as is readily suggested by the way we use material names such as *gold* and *copper* as color descriptors. Our solution was to be sure that in at least a third of the triads, the color of the pivot and the material choice were different. Likewise there seems to be a linkage between size and shape such that if size is identical it seems to facilitate a shape match, and if size varies

dramatically enough it impedes a shape match: consider whether you would want to match a round wooden bead with a square wooden block of similar size or with a much larger beach ball – the larger the ball the less appealing the shape match. So we worked to assure that in at least a third of the shape matches, the size differential was greater than 10% although not vastly different. Looking at the choices that English and Yucatec speakers made across the entire set of twenty-one new triads, we found that the rate of material choices was unaffected by color redundancy and that the rate of shape choices was unaffected by an exact size match. In short, color and size cannot explain relative preference for shape or material in the results to be discussed below.

3.2.1.2 Control of function

The dimension of function, in the sense of the typical use of an object, does sometimes play a role in grammar – most notably in selection restrictions on predicate-argument agreement. Hence we took great pains to evaluate this as a basis of preference. The twenty-one triads were divided into seven sets of three, each set designed to treat the function dimension in a different way. The seven types are listed and illustrated in figure 9.1.

The two triad sets that controlled function most closely were sets 1 and 2. Set 1, the Unifunctional Wholes, consisted of three whole objects that all had a single function, thereby neutralizing function as a possible choice. Set 2, the Trifunctional Wholes, consisted of three whole objects that each had a different function thereby neutralizing function as a possible choice. The original finding was replicated: Yucatec speakers favored material choices in the triads 55.6% of the time whereas English speakers favored material choices only 23.6% of the time ($p < .009$, Kolmogorov–Smirnov two-sample test). These data are given in table 9.2, along with the data from the original triads task (recoded here by triad rather than by subject to facilitate comparison⁶) and data from the other functionally neutral triads that we turn to next. (Again, to facilitate comparison, we have reported all results here in terms of degree of material preference, but the degree of shape preference in any instance can easily be derived since shape and material preferences always add up to 100%. So, for example, English speakers' 23.6% preference for material on this task implies 76.4% preference for shape.)

Another three sets of triads (see figure 9.1) also controlled function by precluding a function match, but, in addition, one or more of the alternates had no function whatsoever. Lack of function was achieved by presenting a scrap or piece of a familiar object as an alternate where the scrap clearly had no customary function at all. (All of the pivots in these triad sets were still whole objects.) We undertook this manipulation because it appeared




MATERIAL ALTERNATE	PIVOT	SHAPE ALTERNATE
1. Unifunctional Wholes		
 plastic comb without handle	 plastic comb with handle	 wooden comb with handle
2. Trifunctional Wholes		
 cardboard matchbox	 cardboard spool	 plastic straw
3. Afunctional Shape Pieces		
 paper book	 square sheet of paper	 square piece of burlap
4. Afunctional Material Pieces		
 scrap of metal	 metal nail	 wooden pencil
5. Afunctional Shape and Material Pieces		
 scrap of plastic	 plastic plumbing joint (Y-shaped)	 wooden stick with branch
6. Function with Shape		
 wooden spatula	 wooden spoon	 ceramic spoon
7. Function with Material		
 wax taper candle	 wax egg candle	 styrofoam egg

Fig. 9.1 Examples of types of triad stimuli.

Table 9.2. *Percentage of English and Yucatec choices showing preferences for material as a basis for object classification: Function Neutral triad task*

Task type	Preference for material (percent)	
	English	Yucatec
Original Triads ^a	37.5	80.0
Function Neutral Triads		
Uni/Trifunctionals (1–2) ^b	23.6	55.6
Afunctional Pieces (3–5) ^b	22.2	64.8
All Function Neutral (1–5) ^b	22.8	61.1

Notes.

^a Recoded from table 9.1 to facilitate comparison (see n. 6).

^b English versus Yucatec: $p < .009$, $n = 12$, one-tailed Kolmogorov–Smirnov two-sample test.

that some of the strongest of the original triads used pieces of things, much like the piece of cardboard in the original example, and we wondered whether this enhanced substance choices.

Set 3, the Afunctional Shape Pieces set, consisted of three items without shared function wherein the pivot and the material alternate were recognizable objects but the shape alternate was clearly just a piece of something. Set 4, the Afunctional Material Pieces set, consisted of three items without shared function wherein the pivot and the shape alternate were recognizable objects but the material alternate was clearly just a piece of something. Set 5, the Afunctional Shape and Material Pieces set, consisted of three items without shared function wherein the pivot was a recognizable object but both the shape and material alternates were clearly just pieces of something. The results on these three sets of triads were essentially the same as those of the first two sets, but the contrast between the two groups was sharpened a bit: English preference for material fell to 22.2% and Yucatec rose to 64.8% ($p < .009$, one-tailed Kolmogorov–Smirnov two-sample test). The use of pieces rather than whole objects in triads otherwise controlled for function appears to have had no major effect on the pattern of preferences. Thus, for the purposes of further comparison, all five of these sets (fifteen triads in all) will be joined and termed the Function Neutral triads. Table 9.2 lists the composite results for the three Afunctional Pieces triad types as well as the values for the composite category of Function Neutral triads.

For all five sets of triads, the Yucatec preference for material choices was not as different from chance (50%) as the English one – which is to be

expected. Strictly speaking, the Yucatec **grammatical** pattern is neutral: nouns referring to discrete objects do not code a quantificational unit as part of their semantics but neither do they code any particular substance quality at this **grammatical** level. Therefore, on **grammatical** grounds, the Yucatec cognitive preference for shape or material should, by rights, be fairly balanced and fall near to 50% on this task. Hence the Yucatec preference for material choices is a relative preference rather than an absolute one – relative, that is, to English speakers, who package quantificational unit into nouns of this type and show a strong inclination to classify objects by shape. To the extent that the Yucatec response rate rises significantly above 50%, it suggests that the **grammatical** neutrality has been intensionalized as a general substance orientation, perhaps because relatively more lexical elements incorporate substance semantics. Such increased response rates do in fact appear on some tasks, for example, the more complex classification task discussed below.

A final two sets of triads (see figure 9.1) controlled function in yet a third way by intentionally joining it systematically with one of the two alternates. Set 6, the Function with Shape set, consisted of three whole objects wherein the pivot and the shape alternate shared function thereby aligning shape with function. Set 7, the Function with Material set, consisted of three whole objects wherein the pivot and the material alternate shared function thereby aligning material with function.

This stimulus manipulation produced dramatic effects when function lined up *against* the dominant cultural preference, moving the response rate strongly towards that of the other group.⁷ Hence the Yucatec preference for material dropped from 61.1% on the Function Neutral triads to 38.9% when function was aligned with shape and the English preference for material rose from 22.8% for the Functional Neutral triads to 72.2% when function was aligned with material (Yucatec change: $p < .021$; English change: $p < .003$, two-tailed Wilcoxon matched-pairs tests). However, when function lined up with the existing cultural preference, it did not produce statistically reliable changes. These patterns are shown in table 9.3.

Overall, it appears that when function matches are available, they can affect the results and need to be carefully controlled when assessing a relative shape versus material classification preference. The dramatic effect of function alignment on English speakers also suggests that function is relatively more important for English speakers than it is for Yucatec speakers. Indeed, for English speakers, joining material and function results in more material choices than the Yucatec speakers typically exhibit.

Taken together, these new triads bring an improved level of control to the assessment of Lucy's (1992b) original hypothesis. The relative shape and material preferences reported in the original study are not only replicated

Table 9.3. *Percentage of English and Yucatec choices showing preferences for material as a basis for object classification: Function Biased triad task*

Task type	Preference for material (percent)	
	English	Yucatec
Function Neutral Triads (1–5) ^a	22.8	61.1
Function Biased Triads		
Function with Shape (6)	8.3	38.9 ^b
Function with Material (7)	72.2 ^c	55.6

Notes:

^a Drawn from table 9.2 to facilitate comparison.

^b Function with Shape versus Function Neutral Triads: $p < .021$, $n = 12$, two-tailed Wilcoxon matched-pairs test.

^c Function with Material versus Function Neutral Triads: $p < .003$, $n = 12$, two-tailed Wilcoxon matched-pairs test.

but also prove to be robust for stimuli of this type when the dimensions of function, color, and size are controlled.

3.2.2 Complex classification task In order to increase the complexity of the triad classification task, we developed a new task, called the nine-sort task, where we asked informants to sort sets of nine objects into two piles. Four objects were made of the same material but each in a different shape; these were the material alternates. Four others were all the same shape but each made of a different material; these were the shape alternates. A ninth, pivot item was made of the same material as the material alternates and in the same shape as the shape alternates. Five different sets of nine-sorts were used in the current assessment. An example of one such set is depicted in figure 9.2 wherein the material alternates are made of cardboard, the shape alternates are tubes, and the pivot is a cardboard tube.

The twelve informants from each language were each asked to sort the nine objects into two groups. The groups were begun by placing one material and one shape alternate onto plastic trays in front of the person.⁸ Then the remaining alternates were handed to them one by one, and they were asked to make two piles so that everything in each pile was the same. The last item presented was always the pivot. If the informant had built one pile of things of the same material and another of things of the same shape, he or she was now forced to decide between material or shape in placing it. For the nine-sort set presented in figure 9.2, for example, if the informant sorted one pile of tube-shaped items and another of cardboard items, then

ALTERNATES

MATERIAL
cardboard



match-box



needle envelope



cheese box



label

SHAPE
tube



copper pipe



flashlight body



plastic pipe



bicycle handle grip

PIVOT

MATERIAL + SHAPE
cardboard + tube



toilet paper core

Fig. 9.2 Examples of a set of nine-sort stimuli.

the pivot item, the cardboard tube, which was presented last, posed a problem for the informant, for it could be put into either pile according to the categories that had been constructed during the sorting of the piles. The assignment of the ninth object is taken to reveal the informant's bias in an ambiguous situation. The nine-sort sets were pretested, and modified as necessary, to ensure that the items could be reliably sorted according to the categories of interest, namely, material and shape.

In this nine-sort task the informant is forced to form a more complex judgment over a series of items. Furthermore, since a variety of objects must be grouped together before coming to the pivot, the verbal name for any one of those objects will not suffice to guide the grouping, so any possible use of this strategy should diminish. For example, if the cardboard group includes a container and a package label, then the only way to include the cardboard tube with that group is to extract the common dimension of cardboard.

This procedure also provided a rich variety of qualitative information about how informants were approaching the classification task. Respondent's answers unfolded in sequence so one could see what items gave them pause. Informants were allowed to rearrange their groups enroute if they wished, and many did so once they had seen a larger range of the items in a given sort. Their preliminary groupings gave important clues to their other classification strategies, even though their final piles ended up being based on material and shape. And the range of grouping strategies, incidental remarks, and follow-up discussions revealed much about how they were approaching the task, particularly in the Yucatec case.

When one looks at how informants chose to place the pivot, the same group differences in preference for material and shape seen in the triads tasks emerge once again, but with a somewhat *stronger* material preference now among the Yucatec. As shown in table 9.4, adult English speakers assign the pivot item on the basis of material only 23.6% of the time, compared to 73.6% for the adult Yucatec speakers ($p < .001$, one-tailed Kolmogorov–Smirnov two-sample test). This suggests that the effects found in the earlier triad studies will be replicated (and may even be stronger) on more complicated tasks and that immediate lexical labeling with respect to a single probe is unlikely to account for the previous triad results.

The qualitative information provided by this task cannot be described in full here, but it supports the interpretation we have given to the quantitative data. The Yucatec speakers were constantly evaluating the material composition of the test items before sorting them: feeling how heavy they were, poking their nails into them to test for malleability, scraping the surface to see what the material under the paint was, smelling and tasting the objects,

Table 9.4. Percentage of English and Yucatec choices showing preferences for material as a basis for object classification: nine-sort task

Task type	Preference for material (percent)	
	English	Yucatec
Function Neutral Triads (1–5) ^a	22.8	61.1
Nine-Sorts ^b	23.6	73.6

Notes.

^a Drawn from table 9.2 to facilitate comparison.

^b English versus Yucatec: $p < .001$, $n = 12$, one-tailed Kolmogorov–Smirnov two-sample test.

and generally questioning or commenting on their material properties – and all this with familiar objects. The English-speaking Americans showed none of this sort of reaction – they could get all the information they needed by sight alone. A particularly striking example of an alternative sorting occurred with one Yucatec woman during pilot work, where we could not make sense of the principle she was using. When we asked her about her reasons during the follow-up discussion, she replied that the things on one tray would melt if they were burned whereas the ones on the other tray would turn to ash. This was a level of attention to material properties that went much deeper than we had originally imagined.

3.3 Summary

On the basis of these results, we can draw three conclusions about the relationship between language spoken and classification preference. First, classification preferences differ in these two language groups. English speakers consistently exhibit a relative preference for shape-based classifications, and Yucatec speakers consistently exhibit a relative preference for material-based classifications. Gentner & Boroditsky (ch. 8, this volume) report a similar pattern of classification differences between speakers of English and Japanese (a classifier language), suggesting that the results reported here will not prove unique.

Second, the findings are robust. They have proven replicable across many years, samples, stimulus configurations, and task types (not all of which have been reported here). It remains to be seen how widely this pattern generalizes, but the various controlled manipulations described here make it unlikely that the basic differences in shape and material preference result from artifacts of the stimuli (i.e. color, size, or function) or task design (i.e. lexical labeling, task simplicity, task directness).

Third, these relative classification preferences were predicted on the basis of the semantic patterns implicit in the **grammatical** categories of the two languages. Similar sorts of correspondences between number-marking patterns and cognitive performance have been found in a number of other tasks using quite different stimuli (Lucy 1992b: 93–136). All these factors converge to suggest that the cognitive differences stem at least in part from the **grammatical** patterns.

4 Developmental issues

Next we want to consider the formation of these language-related preferences in childhood: at what age and in what manner do they arise? The adult cognitive contrast not only gives rise to these questions, it also provides part of the key to their solution. By testing for the presence of the known adult contrast at various ages, we can locate the point in development when language and thought begin to interact in this way. Similarly, comparative work can help illuminate the developmental process itself by separating the constant from the variable, the crucial from the less relevant.⁹

Pilot work (Lucy & Gaskins 1989) indicated that the onset of these changes was between 7 and 9 years of age, so we undertook a more extensive study of these two ages using the classification tasks described above. Twelve English-speaking and twelve Yucatec-speaking children at each age were presented with the same triad materials used in the adult samples. On the Function Neutral triads (i.e. sets 1–5), both English-speaking and Yucatec-speaking 7-year-olds showed an identical early bias towards shape – choosing material alternates only 11.7% of the time. By age 9, the English-speaking children continued to favor shape, choosing material alternates only 17.8% of the time. But by this age, the Yucatec-speaking children were choosing material alternates 41.7% of the time, much like adult Yucatec speakers. Thus, the same kind of language-group difference found among adult speakers is also found in children by age 9 ($p < .033$, one-tailed Kolmogorov–Smirnov two-sample test). The group comparison as a function of age is shown graphically in figure 9.3 (adults are labeled “15+”).¹⁰

Results for the two Function Biased triad sets, where function was alternately aligned with shape or material (i.e. sets 6 and 7) did not produce the big deflections characteristic of the adult groups.¹¹ The Yucatec children’s responses on these triad types were virtually identical to their responses with the Function Neutral triads. The English children’s responses showed some small effects in the expected directions as early as age 7, but they were not statistically reliable. These results are shown graphically in figures 9.4 and 9.5.

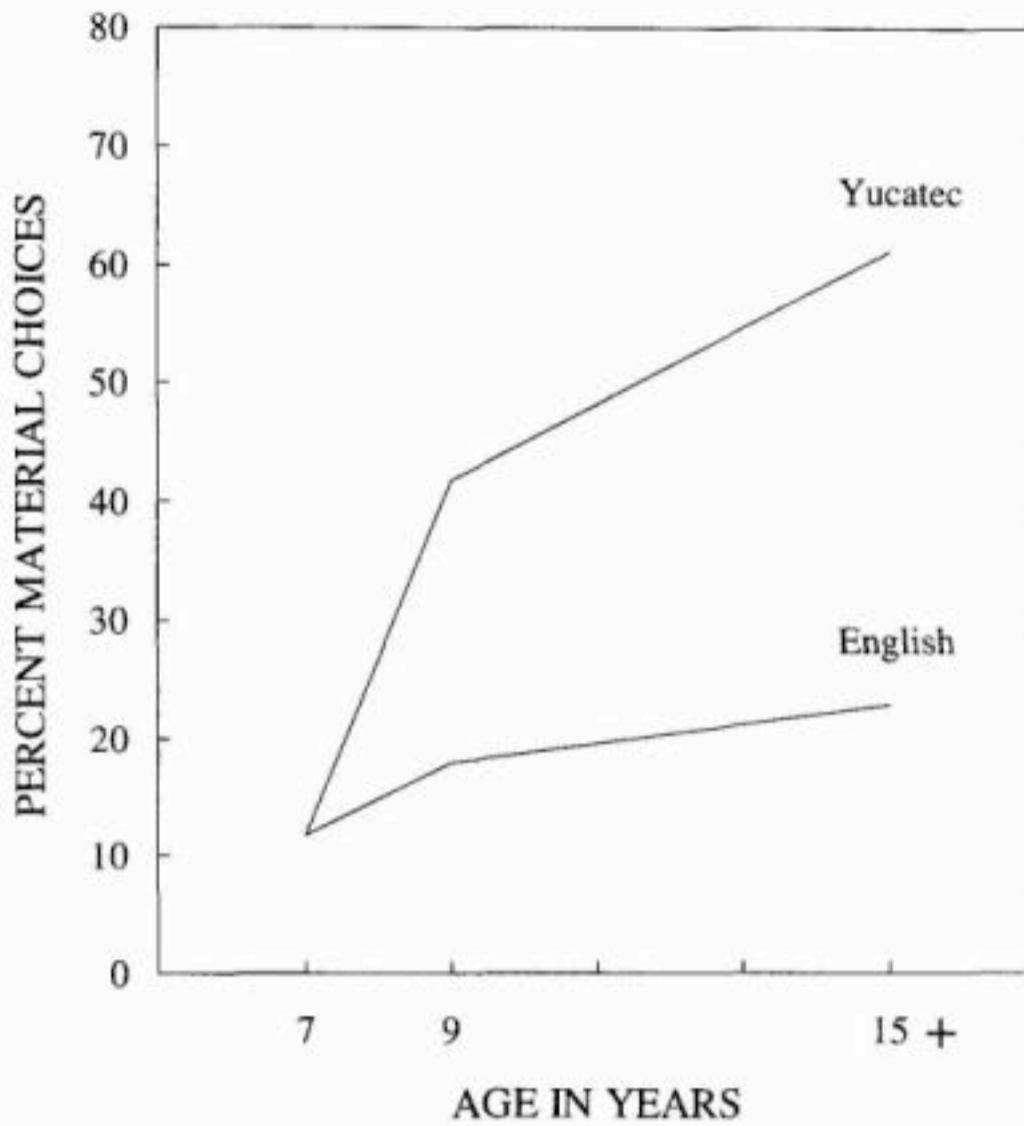


Fig. 9.3 Developmental pattern for English and Yucatec classification preferences: material vs. shape using Function Neutral triad types.

A modified version of the nine-sorts task was also administered to the same groups of children. The modified procedure was developed because children were much more likely to generate bizarre groupings with the first few objects, groupings they would then doggedly stick with rather than revise in light of subsequent objects appearing in the set. To minimize this tendency, all the alternates and the pivot were made visible to the children from the outset. In the modified procedure a paper cross (or "X") with four points was placed on the table and was encircled by an array of the eight alternate items. Then the pivot item was put in the middle of the cross, and the children were asked to pick from the array of alternates one item for each point of the cross such that the five items ending up on the cross would all be the same. This allowed the children to scan the total array of items before forming a group, and they usually had no trouble forming a plausible grouping under this arrangement. Indeed, one striking fact about both 7- and 9-year-olds

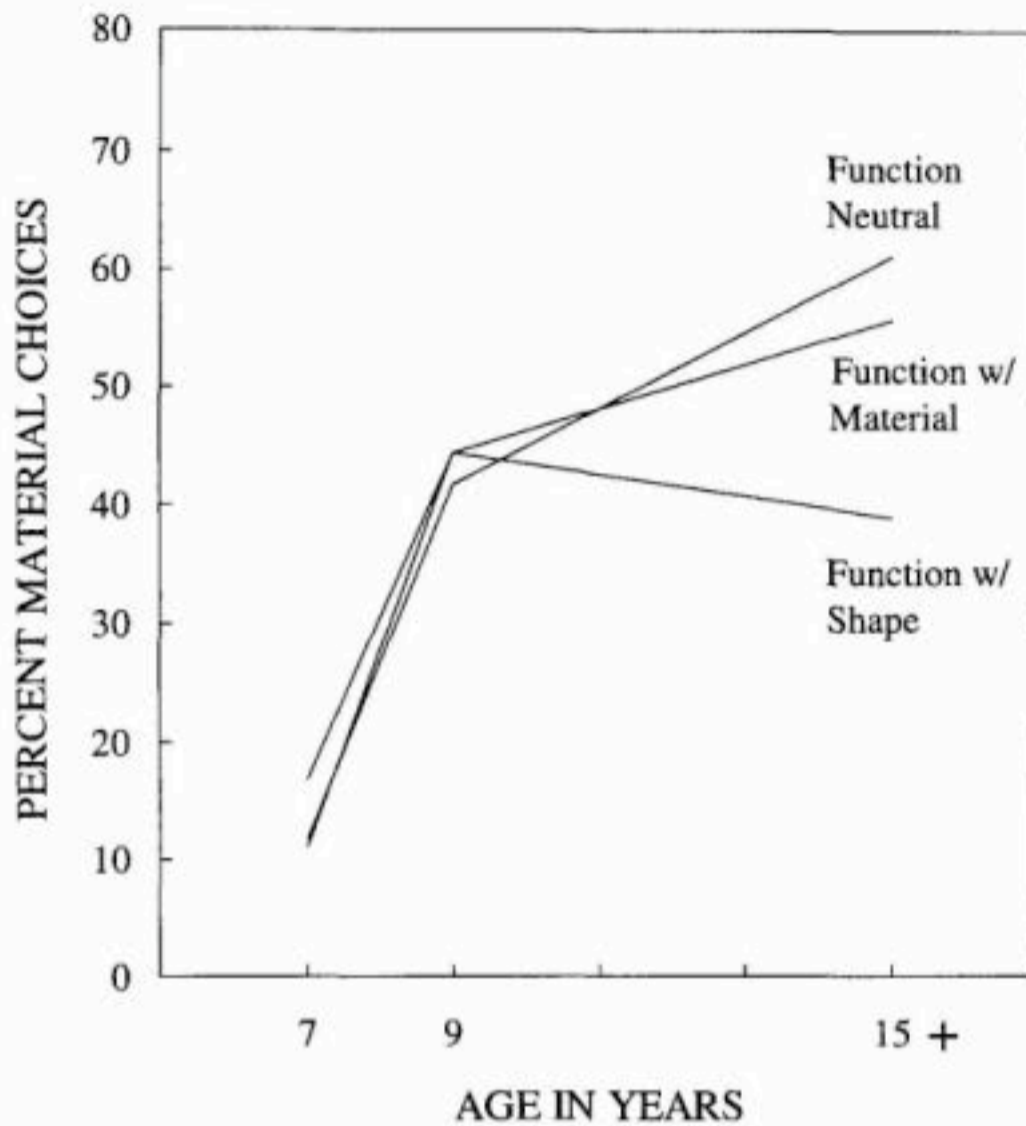


Fig. 9.4 Developmental pattern for Yucatec classification preferences: effect of joining function with shape and material.

was that they responded very decisively to the task in this form: they were sure there was a “right answer,” and that they had sorted according to it. They often didn’t even see the alternative classification possibility.

The same general pattern of Yucatec speakers showing relatively greater preference for material emerges on this task, but with one significant difference. Although the 7-year-old children in both groups favor shape as a basis of classification, the Yucatec-speaking children already demonstrate a tendency towards the adult Yucatec pattern. English-speaking children choose on the basis of material only 2.8% of the time whereas the Yucatec do so 20.8% of the time ($p < .033$, one-tailed Kolmogorov–Smirnov two-sample test). By age 9, the English-speaking children continued to favor shape, choosing material alternates only 8.3% of the time, but the Yucatec-speaking children have shifted dramatically, choosing material alternates 48.6% of the time, approaching the adult pattern, as they did on the triads

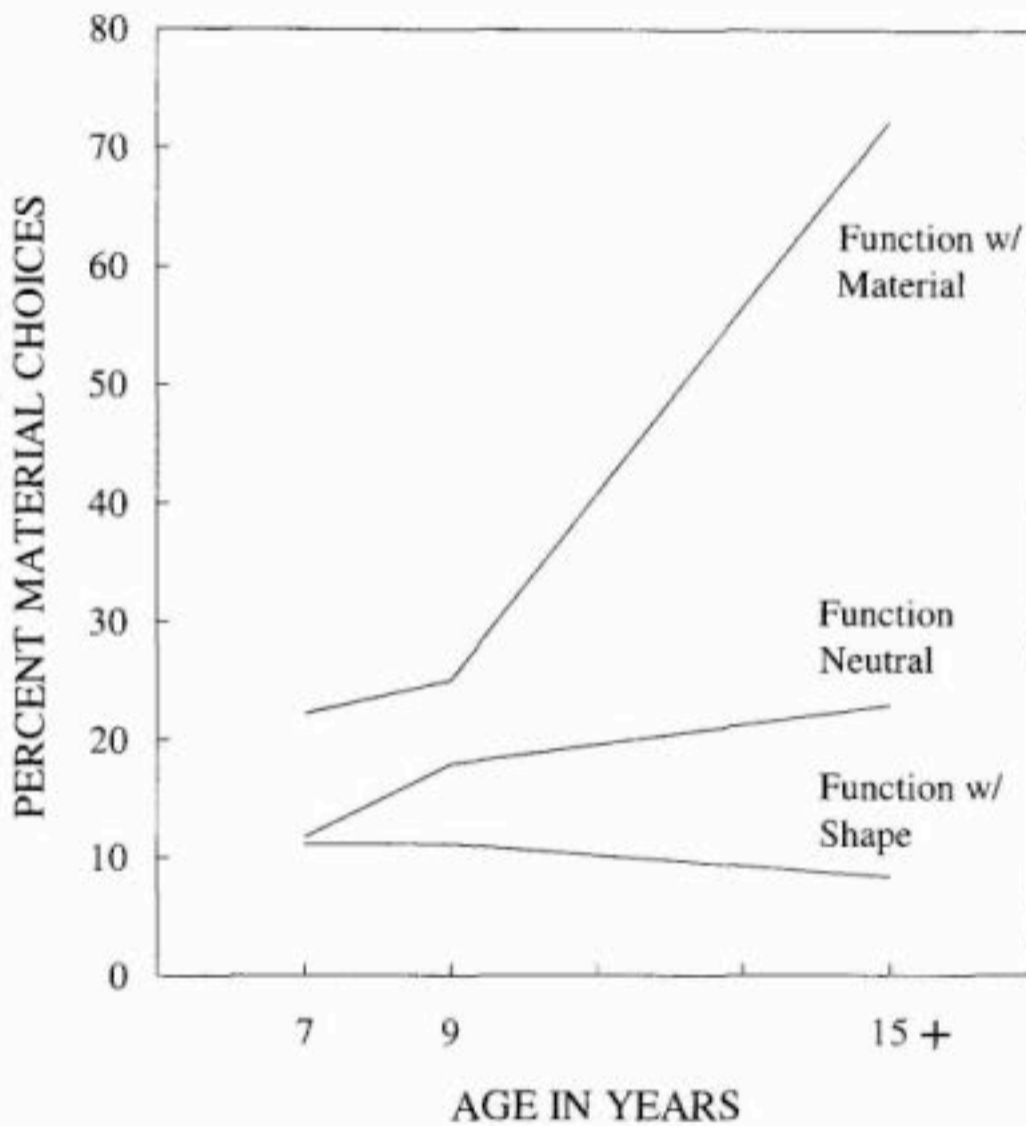


Fig. 9.5 Developmental pattern for English classification preferences: effect of joining function with shape and material.

task ($p < .009$, one-tailed Kolmogorov-Smirnov two-sample test). The group comparison as a function of age is shown graphically in figure 9.6.

Overall, the more complex nine-sort task appears to be more sensitive than the triads task to the Yucatec material bias in that the bias is stronger at all ages. The task even reveals a bias at age 7, suggesting that there may be some grammatical-category-related effects as early as this age. Nonetheless, as the graph makes clear, the major move in this direction occurs after age 7 with the Yucatec bias and the group differences becoming larger with age.

Taken as a whole, these developmental data suggest that there is a major shift towards the culturally specific adult cognitive pattern by age 9. On the basis of these results, we can draw three conclusions about the development of language-related classification preferences for these kinds of objects. First, the early tendency at age 7 is to prefer shape over material as a basis of classification for these kinds of objects regardless of language affiliation. This preference is unaffected by the introduction of function as an alterna-

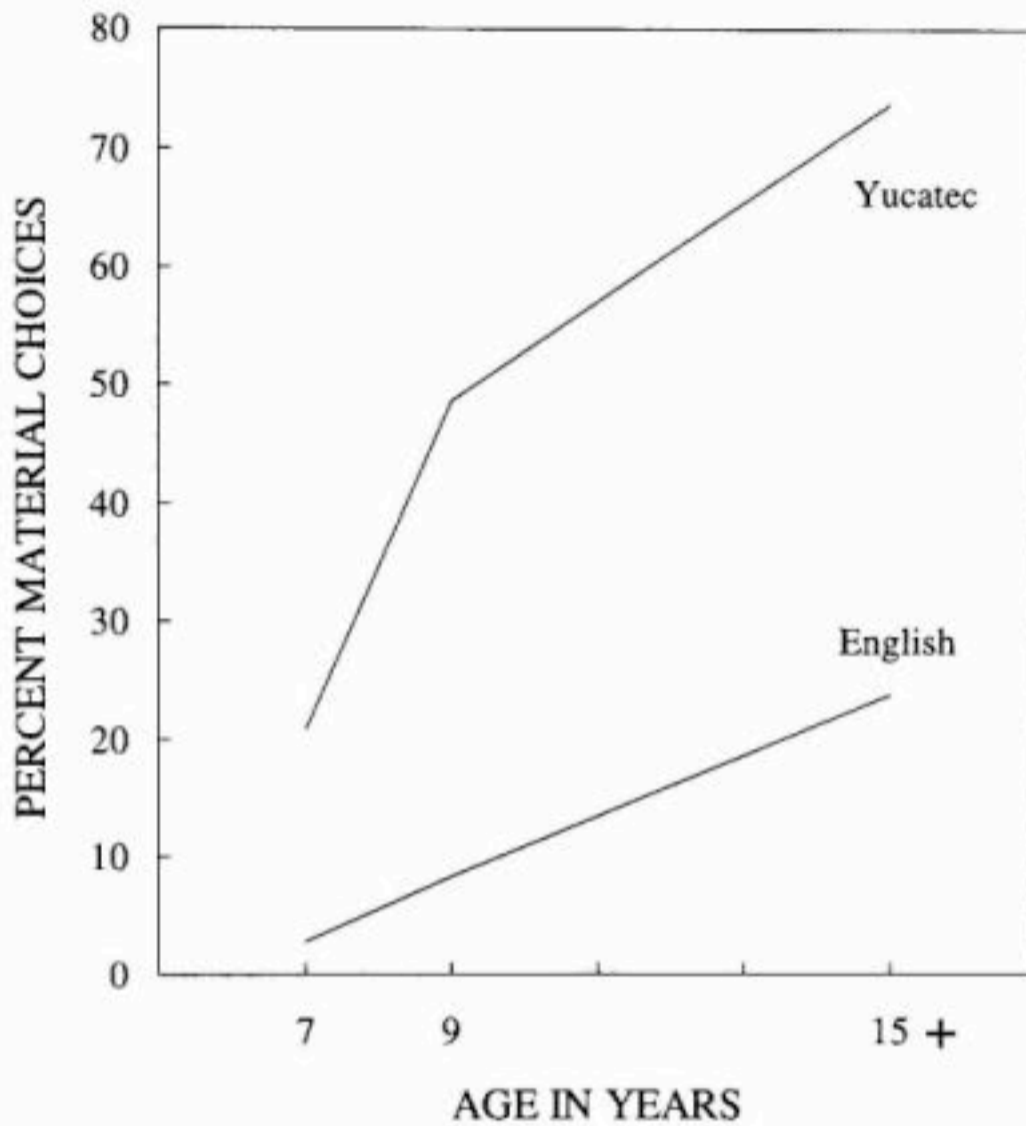


Fig. 9.6 Developmental pattern for English and Yucatec classification preferences: material vs. shape using nine-sorts.

tive basis of classification. Although Yucatec-speaking children already show some early relative inclination towards material choices, their overwhelming absolute preference is for shape.¹²

It is perhaps worth emphasizing that just as the English-speaking children have substantial command of the plural by age 7, so too do the Yucatec-speaking children have substantial command of the numeral classifier system by this age. Seven-year-old Yucatec-speaking children reliably use classifiers when counting, draw appropriate semantic distinctions among them in comprehension tasks, and will judge a number construction lacking them as faulty. However, they fall far short of having command of the full range of classifiers in comprehension and their range in production is narrower still. In short, they have the basic structural implications straight but do not yet have the full lexical range of an adult (Lucy & Gaskins, in preparation). Hence, to the extent these cognitive results derive from these basic structural characteristics of the language rather than

mastery of specific lexical items, there is no reason they could not appear at age 7. That they do not do so suggests that some rather specific reorganizations in the relation between language and thought take place between ages 7 and 9.

Second, Yucatec-speaking children at age 9 show a strong relative preference for material as a basis of classification and appear well on their way towards the adult Yucatec pattern. Although English-speaking children also show a very slight increased sensitivity to material bases of classification by age 9, their overwhelming preference continues to be for shape, in line with the adult English pattern. Because the adult preference for English speakers agrees with that shared by young children, there is little evidence for overt change in classification preference by age for the English-speaking sample. Indeed, it would be difficult to identify this age period as the crucial one without the comparative Yucatec data. But the English-speaking adults are not quite so shape-biased as the 7-year-olds, and the 9-year-olds typically show an intermediate pattern suggesting that they are beginning to respond to the same factors as English-speaking adults. In time it should be possible to design further tests to assess whether the same underlying psychological mechanisms are operating for the English samples to produce the shape preference during the various age periods, or whether there is in fact a developmental change in the basis for classification in the English sample as well, even though the overt preference for shape is uniformly expressed at all ages. More generally, additional work will be necessary to determine whether other language-related shifts occur during this age range.

Third, the results indicate that children are more rigid than adults in their responses on these classification tasks. Children in both groups begin with a strong preference for shape and shape alone. By age 9, there is a greater sensitivity to material in both groups, but its potency is heavily conditioned by language group membership. By adulthood, function plays a significant role, deflecting the baseline preference for shape versus material; but, again, the potency of these shifts depends heavily on language group membership, with function only having a significant effect when it runs counter to the baseline preference in the language. The effects of these changes appear clearly in figures 9.4 and 9.5: with increasing age for both groups there is a move away from shape (revealed by greater vertical displacement) and towards greater impact of functional manipulations (revealed by greater absolute spread in the degree of material preference across various triad types). What we see then is a shared decline in rigidity and the growth of greater classification flexibility across both language groups, but a flexibility markedly deflected by differing group norms and differentially sensitive to displacement by other factors, again as a function of those norms.

5 Conclusion

These empirical findings provide an excellent illustration of the importance of adopting a comparative approach in the study of the relation between language and thought (cf. Byrnes & Gelman 1991). By extension we would argue that such comparative work is essential in human psychology more generally because of its deeply cultural nature. A comparative approach was adopted from the outset here and it is worth reviewing exactly how this was important for the research process.

First, by documenting the range and patterning of behavioral diversity, an important set of problems and hypotheses was opened up. In particular, close study of a non-European language revealed basic differences in morphosyntactic structuring that were then interpreted in terms of the different verbal construals of experience implicit in them. We have gone on to explore the possible broader cognitive consequences of these differences – in particular, how speakers of the two languages might be interpreting reality differently beyond the act of speaking, but it is worth noting that the same contrast raises many other important questions. For example, the differences raise questions about the viability of certain proposals about how languages are structured and how children learn them. A large number of studies have appeared in recent years attempting to show that young children have an early bias towards drawing a sortal–object distinction, that is, towards believing that words refer to discrete or bounded solid objects (e.g. Markman 1990, 1991; Soja, Carey, & Spelke 1991; Clark, ch. 13 of this volume). Researchers differ on whether they think the sources of this bias are more cognitive (e.g. Bloom, ch. 6 of this volume) or more specific to language learning (e.g. Waxman 1994), but they are in accord about the utility of such a bias because they believe all languages are structured this way. Yet when we consider the many languages that differ dramatically in both their number marking systems and the relative importance of such a distinction for the language, we can immediately see that such proposals need to be tempered. What good is an early bias that has to be systematically undone by a quarter or more of the world's people in order to speak properly? Clearly some rethinking is required here.

Second, the descriptive linguistic work was used as a platform for formulating and testing for tangible psychological implications of the linguistic differences. The work with adults indicates that language-specific grammatical patterns are indeed correlated with classification preferences. Whether or not one believes that the language patterns give rise to the differences, the comparative cognitive data alone compel a rethinking of the notion of a unified developmental outcome for cognition. To the extent language does give rise to these cognitive differences, the study casts doubt on the notion,

widely held, that cognition enables and influences language but not vice versa (e.g. Fodor 1975). What is especially important to note is that the contrast between languages both gives rise to the cognitive prediction in the first place and then indicates exactly what difference we need to look for. Without the comparative contrast, the harmony of language and thought in English would both obscure the existence of the cognitive differences and make detecting them nearly impossible. (Spelke & Tsivkin, ch. 3, this volume, make a similar argument.)

Third, the study exploits the diversity in adult response in order to uncover the nature of the psychological mechanisms and developmental processes at work. The contrast between the two groups is essential to establishing when the language-linked cognitive patterns arise, what the qualitative nature of the classification preference is before the differences arise, and how responses differ in flexibility from subsequent adult forms. In short, by exploiting known differences among adults we have a powerful way to diagnose both the timing and the qualitative nature of important developmental shifts.

Regarding the changes in middle childhood, there already exists, of course, substantial evidence suggesting that thought and language are changing in new and important ways during this period. Cognitive developmentalists (Piaget & Inhelder 1969; Vygotsky 1987 [1934]) have long recognized this as the period in which the child completes a shift from dependence on more spontaneous, perceptual strategies to reliance on more systematically organized, conceptual ones. In the realm of language development, there is during this period a general reorganization of grammar along adult lines, increasing sensitivity to the syntactic and semantic presuppositions of prior discourse, and increasingly hierarchical organization of narratives (Chomsky 1969; Bowerman 1982; Karmiloff-Smith 1984; Hickmann 1993; Berman & Slobin 1994). What has not been obvious before is the degree to which these two sets of transformations may be linked in language-specific ways. Children may well draw on their increasingly sophisticated **grammatical** and discursive skills to reconfigure existing cognitive competencies into more stable, shared conceptual systems, systems that can then be relied on to correct or supplement interpretations of the world that have heretofore been partial and highly susceptible to perceptual and immediate contextual factors. In short, the child now enters the world of the adult, which is more heavily guided by systems of shared cultural meaning (Lucy & Gaskins 1994).

Both the recognition of such developmental linkages and their detailed exploration profit enormously from adopting a comparative approach. More generally, we cannot ever hope to reach an adequate understanding of human beings by ignoring behavioral diversity, nor through assimilating

it to our own norms as advanced or deficient, nor through regarding it as incomprehensible. Rather, such understanding will only come through the careful study of that diversity for what it can tell us about how people's experience of the world is mediated (developed and sustained) through languages and cultures and, in turn, how those languages and cultures are mediated (developed and sustained) by psychological processes.

NOTES

The new research reported here has been supported by the Spencer Foundation (Chicago, USA) and the Max Planck Institute for Psycholinguistics (Nijmegen, The Netherlands). Christine Kray assisted with the Yucatec sample. Kathryn Mason assisted with the English sample. We thank the volume editors Stephen Levinson and Melissa Bowerman for detailed comments on the original draft. Inge Doehring helped with the figures.

- 1 For a fuller discussion of these four requirements, see Lucy (1992a: 263–275; also 1992b: 1–2). In particular, the characterization of “reality” must not favor any one language and yet be consistent with how reality “appears through the window of language.”
- 2 The two closest candidates, after Whorf's (1956 [1939]) pioneering effort comparing Hopi and English, were Carroll's comparison of Hopi and English (Carroll & Casagrande 1958) and Heider's comparison of Dani and English (Heider 1972). Both are weak on criterion (3) in that they dealt with contrasts between small lexical sets without reference to a general structural or functional analysis of the two languages. For a full discussion, see Lucy (1992a). Since that review, an important body of new work has been produced by Steve Levinson and his colleagues (Levinson 1992, 1996; Brown and Levinson 1993; Pederson 1995; Pederson, Danziger, Wilkins *et al.* 1998). For a review, see Lucy (1997).
- 3 The term *discrete* here is being used to designate a common characteristic of referents of lexical nouns falling into a category defined by formal comparison of the two languages as discussed in Lucy (1992b: 56–83). The focus here will be entirely on this type of lexical noun. Other noun types and their referents are dealt with in separate studies (Lucy 1992b: 93–136; Lucy & Gaskins, in preparation; cf. Gentner & Boroditsky, ch. 8 of this volume).
- 4 The example described here is only one of a range of studies reported in Lucy (1992b: ch. 3) testing for various cognitive correlates of linguistic patterns.
- 5 For an example of another indirect test that eliminates this possibility and reveals the same preference pattern, see Lucy (1992b: 136–144).
- 6 Subject coding was used in the original study where the aim was to characterize individual sorting tendencies as a function of language spoken. In the present study, where the focus is on the effect of different triad types within and across languages, item coding was used instead. Since each subject contributes equally to the results and they have been expressed as percentages for the group comparisons, the statistical sample size ($n = 12$ for each group) has not been artificially inflated.
- 7 Some differences among referent types in word learning tasks as reported by Gentner & Boroditsky (ch. 8 of this volume) may have their origins in the shift-

- ing alignment of function with other features of reference. However, in the developmental studies reported below, this adult effect does not appear.
- 8 The order of presentation of alternates, including which ones were presented first, and the right-left presentation of shape and material alternates were counterbalanced within and across subjects.
 - 9 Full demonstration of the utility of the comparative approach for illuminating the developmental process itself cannot be undertaken here because it would require the introduction of several additional studies (see Lucy & Gaskins, in preparation).
 - 10 Figure 9.3 is misleading in one respect. Preliminary assessments at ages 11 and 13 among Yucatec-speaking children suggest that the 9-year-old level of performance persists and a further rise to the adult level of performance does not occur until after age 13. However, because these data are incomplete and we lack any comparable English data, we have omitted them from figure 9.3.
 - 11 See n. 7.
 - 12 The early preference for shape classification with objects of this type should not be interpreted as a general bias towards all referents. In experiments with stimuli that both languages would treat as materials, both groups show similar material preferences across this age range (Lucy & Gaskins, in preparation).

REFERENCES

- Berman, R. A., & D. I. Slobin. (eds.). 1994. *Relating events in narrative: a cross-linguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bowerman, M. 1982. Reorganizational processes in lexical and syntactic development. In E. Wanner & L. R. Gleitman (eds.), *Language acquisition: the state of the art*. Cambridge: Cambridge University Press, 319-346.
- Brown, P., & S. C. Levinson. 1993. Linguistic and nonlinguistic coding of spatial arrays: explorations in Mayan cognition. *Working Paper 24*. Nijmegen, The Netherlands: Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics.
- Byrnes, J. P., & S. A. Gelman. 1991. Perspectives on thought and language: traditional and contemporary views. In S. A. Gelman & J. P. Byrnes (eds.), *Perspectives on language and thought: interrelations in development*. Cambridge: Cambridge University Press, 3-27.
- Carroll, J., & J. Casagrande. 1958. The function of language classifications in behavior. In E. Maccoby, T. Newcomb, & E. Hartley (eds.), *Readings in social psychology*. New York: Henry Holt, 18-31.
- Chomsky, C. 1969. *The acquisition of syntax in children from 5 to 10*. Cambridge, MA: MIT Press.
- Fodor, J. A. 1975. *The language of thought*. New York: Thomas Y. Crowell.
- Heider, E. 1972. Universals in color naming and memory. *Journal of Experimental Psychology* 93: 10-20.
- Hickmann, M. 1993. The boundaries of reported speech in narrative discourse: some developmental aspects. In J. A. Lucy (ed.), *Reflexive language: reported speech and metapragmatics*. Cambridge: Cambridge University Press, 63-90.
- Karmiloff-Smith, A. 1983. Language development as a problem-solving process. *Papers and Reports on Child Language Development* 22: 1-23.

- Levinson, S. C. 1997. Language and cognition: the cognitive consequences of spatial description in Guugu Yimithirr. *Journal of Linguistic Anthropology* 7: 98–131.
1996. The role of language in everyday human navigation. *Working Paper 38*. Nijmegen, The Netherlands: Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics.
- Lucy, J. A. 1992a. *Language diversity and thought: a reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- 1992b. *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
1997. Linguistic relativity. *Annual Review of Anthropology* 26: 291–312.
- Lucy, J. A., & S. Gaskins. 1989. Language diversity and the development of thought. Paper presented at the 88th Annual Meeting of the American Anthropological Association, Washington, DC, November.
1994. The role of language in shaping the child's transition from perceptual to conceptual classification. Paper presented at the 93rd Annual Meeting of the American Anthropological Association, Atlanta, GA, December.
- In preparation. *Grammatical categories and cognitive development*.
- Markman, E. M. 1990. Constraints children place on word meanings. *Cognitive Science* 14: 57–77.
1991. The whole object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meaning. In S. A. Gelman & J. P. Byrnes (eds.), *Perspectives on language and thought: interrelations in development*. Cambridge: Cambridge University Press, 72–106.
- Pederson, E. 1995. Language as context, language as means: spatial cognition and habitual language use. *Cognitive Linguistics* 6: 333–362.
- Pederson, E., E. Danziger, D. Wilkins, S. C. Levinson, S. Kita, & G. Senft. 1998. Semantic typology and spatial conceptualization. *Language* 74: 557–589.
- Piaget, J., & B. Inhelder. 1969. *The psychology of the child*. New York: Basic Books.
- Soja, N. N., S. Carey, & E. S. Spelke. 1991. Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition* 38: 179–211.
- Vygotsky, L. S. 1987 [1934]. Thinking and speech. In R. W. Rieber & A. S. Carton (eds.), *The collected works of L. S. Vygotsky*, vol. 1: *Problems of general psychology* (trans. N. Minnick). New York: Plenum Press, 37–285.
- Waxman, S. R. 1994. The development of an appreciation of specific linkages between linguistic and conceptual organization. In L. Gleitman & B. Landau (eds.), *The acquisition of the lexicon*. Cambridge, MA: MIT Press, 229–258.
- Whorf, B. L. 1956 [1939]. The relation of habitual thought and behavior to language. In J. Carroll (ed.), *Language, thought, and reality: selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press, 134–159.