The Development of Color Categories in Two languages: a longitudinal study.

Debi Roberson

University of Essex

Jules Davidoff

Goldsmiths College

Ian Davies

University of Surrey

Laura Shapiro

Goldsmiths College

MS # 2003-131-r-r-r
Received Aug 3/04

Abstract

This study unites investigations into the linguistic relativity of color categories with research on children's color category acquisition. Color naming and comprehension, together with children's ability to remember colors were tracked in two populations over a three-year period. A group of children from a semi-nomadic equatorial African culture, whose language contains five color terms, was compared to a group of English children. Despite large differences in visual environment, language and education, there were notable similarities in the pattern of term acquisition. Children from both cultures appeared to acquire color vocabulary slowly and with great individual variation. Those who knew no color terms made recognition errors based on perceptual distance, and the influence of naming on memory increased with age for both groups. The results suggest that an initial perceptually driven color continuum is progressively organized into category sets appropriate to each culture and language.

The development of color categories in two languages: A longitudinal study.

Although the physiological basis of color vision is essentially the same for all humans with normal trichromatic color vision (Mollon, 1999), there is considerable diversity in the way that different languages segment the continuum of visible colors. Some languages are reported to use as few as two terms to describe all colors (Rosch, 1972); others use many more (Kay, Berlin & Merrifield, 1991; MacLaury, 1987). The variability exists even for those terms deemed by Berlin and Kay (1969) to be 'basic' (monolexemic, present in the idiolect of all observers and not subsumed within the meaning of other terms). Once one considers secondary terms there is far greater diversity. Nevertheless, despite these diverse naming systems there are empirical data that support the proposal for pan-human cognitive universals in color categorization transcending terminological differences (Heider & Olivier, 1972; Rosch, 1972). Rosch (1973) further argued that these universal categories are each based on the same focal colors regardless of the number of terms in the speaker's language.

The view that there is between-culture agreement as to the fundamental names (basic color terms or BCTs) and that these derive from the structure of the visual system, which is shared between all humans, remains prevalent. Indeed, some recent studies of languages having eleven basic categories (Sturges & Whitfield, 1997; Guest & Van Laar, 2000; Lin, Luo, MacDonald & Tarrant, 2001) have provided some support for the pre-eminence of category centers, or 'foci'. However there is no correspondence between BCTs and any processes yet found in the visual system (Boynton, 1997, Webster et al., 2000; Valberg, 2001); nameability is an important feature of color sets, independent of any perceptual qualities of focality (Guest & Van Laar, 2002) and cross-cultural studies have found no increased salience for the proposed universal ‘focal’ colors (Davidoff, Davies & Roberson, 1999; Jameson & Alvarado, 2003a). The present study, therefore, reopens the debate on innate, pan-human cognitive universals in color categorization (the Universalist hypothesis) by turning to a new source of evidence. It examines the acquisition of color terms.

The studies reported here bring together two well-researched fields of investigation into color categorization. The first field concerns a broad range of recent studies that have reported effects of language on a variety of cognitive tasks (number systems: Gumperz & Levinson, 1997; spatial relations: Levinson, 1996; Bowerman & Choi, 2001; artifact categories: Malt & Johnson, 1998; Malt, Sloman & Gennari, 2003; modes of motion: Gennari, et al., 2000; time: Boroditsky, 2001; shape: Lucy, 1992; Roberson, Davidoff & Shapiro, 2002 and grammatical gender: Sera et al., 1994, 2002). In particular, a number of studies with adults have shown effects of language on color categorization (Kay & Kempton, 1984; Pilling, Wiggett, Özgen, & Davies, 2003; Roberson & Davidoff, 2000; Roberson, Davies & Davidoff, 2000). For example, Roberson et al. (2000) reported studies with adults in a remote hunter-gatherer tribe in Papua New Guinea. Those studies found substantial differences in perceptual judgments and memory performance between a language with eleven BCTs and one with only five (Berinmo). Contrary to the influential view that language and cognitive organization can be independent (Rosch, 1972, 1973), these differences suggest that language not only facilitates memory performance, but also affects the perceived similarity of stimuli (the Relativist hypothesis), a result also found by Kay and Kempton (1984).

A second, and largely independent, field of study has examined the tardy acquisition of color terms by children (Rosch, 1971; Andrick & Tager-Flusberg, 1986; Mervis, Bertrand & Pani, 1995; Soja, 1994; Sandhoffer & Smith, 1999). These studies considered the acquisition of color terms in English speaking children for whom the set of basic terms to be acquired would be just those that are presumed to be universally present before the correct acquisition of the linguistic terms. Thus, given the evidence that primary categories appear to be in place at 4 months of age (Bornstein, Kessen & Weiskopf, 1976), the relatively late acquisition of color terms has puzzled many researchers. Bornstein (1985) found that three-year-old children were slow to learn paired associates to colors relative to abstract shapes, and Sandhoffer and Smith (1999) found that young children were both slower to comprehend color terms than size terms and to match items on the basis of color (but see Pitchford & Mullen, 2001).

Estimates of the age at which children acquire a minimum color vocabulary (4 basic terms) have consistently dropped from the 7-8 years of age estimated by Binet and Simon (1908; cited in Bornstein, 1985) to 2-3 years (Shatz et al., 1996; Andrick & Tager-Flusberg, 1986). Indeed, several studies have argued that, with constant intensive training, children as young as 1.5 years can produce and use some color terms accurately (Cruse, 1977; Mervis et al., 1995). However, Rice (1980) reported that hundreds of trials were needed to achieve this outcome, compared to the single presentation learning demonstrated for object terms (Carey, 1978), and without such intensive input, estimates of the age at which children acquire color terms falls between two and six years, depending on the number of terms examined and the measures of knowledge taken. These studies used a variety of methodologies and either examined only a single measure (naming or comprehension), or the acquisition of only a small subset of terms over a very short period of time. We therefore introduce strict controls on our response measures, examining both naming and comprehension systematically over a three year period in order to establish a more reliable measure of children's color term acquisition.

A related and unanswered question, raised initially by Bornstein (1985), concerns how color term acquisition differs in speakers of different languages. In the framework of a presumed innate, universal fixed set of color categories, Bornstein (1985) predicted that acquiring color terms would be even more difficult for children learning a language in which the universal set must be over-written by a new set, even if there were fewer terms to be learnt. They might have to assimilate their existing hue-based universal categories into a new and orthogonal set of linguistic categories based on another dimension, such as lightness in the case of the Dani reported by Rosch (1972). Similarly, Bowerman and Choi (2003) suggest that, the more robust and pre-potent the pre-linguistic organization of the perceived world is, the greater the resistance that language acquisition would have to overcome in order to re-structure mental life. Thus, while Davidoff et al. (1999) presented compelling evidence for linguistic influence on color categories in adults, the acquisition of a set of name categories that are different to the presumed set of innate, universal categories might show a different developmental pattern to that of English speaking children.

We set out to address these questions in a study that included two groups speaking different languages. One group of young English children were tested initially before they entered pre-school and subsequently through three years of formal education. The other group of children were also followed longitudinally and belonged to the Himba: a semi-nomadic cattle-herding tribe in northern Namibia. Himba people speak a dialect of the Herero language but cultural isolation over the last hundred years has resulted in a variety of cultural and linguistic differences. While villages have a permanent base to which they return in the short rainy season, they move around a series of temporary 'bush camps' during the dry season to find grass and water for their cattle. The area has few natural resources and their homes, clothes, tools and artifacts are made from cattle products. Contact with the outside world is sporadic and, for a number of the youngest children tested, the experimenter was the first white person they had encountered. Few of the children received any formal education during the period of the study. For those that did, school attendance was restricted to brief periods and color terms were not included in the curriculum. Himba has five BCTs according to the criteria of Kay et al. (1991), although one, *burou*, is a recently borrowed term from Herero. They are monolexemic, not subsumed under the meaning of other terms, not restricted to a narrow class of objects and understood by all observers. The five terms are: *serandu* (broadly *red* with *orange* and *pink*), *dumbu* (broadly *beige* with *yellow* and some light *green*), *zoozu* (broadly all dark colors and *black*), *vapa* (broadly all light colors and *white*) and *burou* (broadly *green* with *blue* and *purple*).

The present study seeks to answer questions concerning color category acquisition, primarily from analyses of the child's color term knowledge and memory errors. There are different predictions for errors made during early color term acquisition depending on their origin. If there is a predetermined, universal set of 11 pre-linguistic cognitive categories then children who have yet to learn any of the appropriate labels for their language should show similar patterns of memory errors, inasmuch as these reflect an inevitable organization of color. Within-category confusions should be more common than across-category confusions even with equal difference steps between stimuli (measured in a perceptually uniform metric such as the L*a*b* dimensions of C.I.E. space; Wyszecki & Stiles, 1982). Crucially, these patterns would resemble the organization of color by adult English speakers into eleven basic categories. So, if children start with a predetermined set of perceptual color categories, English children learning color terms need only to learn to map a set of labels onto their existing color categories. As they also receive extensive training in a cultural environment where these are highly salient, English children should learn the universal set of color categories before Himba children learn their linguistic set. Himba children would not only lack explicit instruction, but also be disadvantaged by having to override the universal set in order to learn the set appropriate to their own language.

An alternative account suggests that errors derive from the color terms in the speaker's language. Children might acquire some familiarity with the set of categories appropriate to their language and culture before they learn to apply any terms correctly. In this case, one might expect to see a reflection of the different adult patterns of naming in the cognitive organization of color even for those children who have yet to succeed in naming color stimuli. At later stages, the pattern of errors should be clearly dependent on the color terms in the speaker's language. In addition, young children's error-prone performance in color naming may itself influence their cognitive representation of color. If, as Slobin (1996, 2003) suggests, "thinking for speaking" maintains some patterns of association at a high level of activation, this may contribute to the difficulty children have in overcoming initial naming errors or overextensions. However, this account would require a different explanation for memory errors made by children who definitely know no color terms. Without a predetermined organization, initial errors must depend on perceptual similarity. Confusions would be made to nearest perceptual neighbors regardless of category. Hence both groups of children who know no color terms in their own language would be expected to make similar errors. As children acquire color terms the effects of categorization (i.e., greater within than between-category similarity) would become more important than perceptual similarity.

Previous studies with English speaking children (Andrick & Tager-Flusberg, 1986; Sandhoffer & Smith, 1999) found considerable discrepancy between children's productive naming of color stimuli and their comprehension of color terms. It could be that productive naming is a more difficult task, since the number of available responses is potentially infinite. Moreover, Zelazo and colleagues (Zelazo & Reznick, 1991; Zelazo, Frye & Rapus, 1996) found an asynchronous pattern of knowledge development in young children, depending on the modality of response required. Where verbal and manual responding were contrasted, children often pointed to a correct choice, but said the incorrect one. Such behavior might be particularly evident in a population of children, like the Himba, if they are learning to overwrite universal cognitive categories with a new set of linguistic ones. The present study therefore, in addition to eliciting the lexical color terms that children knew, examined both productive and comprehension measures of children's color term knowledge. We took a composite measure of term attainment first proposed by Soja (1994) that includes both correct naming of stimuli and correct indication of appropriate stimuli when asked to point to stimuli that would be called by a particular name (e.g., '*red*').

In summary, the study had two principal aims. The first aim was to compare universal and relativist hypotheses of color categorization by comparing children's acquisition of color categories over time in two different languages and cultures in order to establish whether the pattern of acquisition relates to the appropriate set of adult categories. Important data would come, in particular, from children's memory confusions at the first time of testing before they appropriately label colors. The second aim, in light of the widely varying reports of children's acquisition of color concepts, was to systematically compare the similarities and differences of the learning trajectories of children from two different cultures as they acquired their appropriate set of color categories, by tracking naming and comprehension of color terms and cognitive organization of color longitudinally over a three-year period.

Method

Participants.

The English children were: 32 three-year-olds (12 males, 20 females), mean age 37.9 months; 36 four-year-olds (14 males, 22 females), mean age 47.5 months. The Himba children were: 42 three-year-olds (25 males, 17 females) estimated[1] mean age 40 months and 27 four-year-olds (13 males, 14 females) estimated mean age 50 months. Of the sixty-nine Himba children, 63 became the longitudinal sample, completing all 6 tests (30 males and 33 females). These children were tested on each occasion in their home villages in northern Namibia. All had received no schooling at the first time of testing. At the 6th time of testing, 34 children had attended a newly introduced mobile school for a period of between 6 months and 1 year. The remaining children had never attended school. Of the thirty-two three-year-old English children, twenty-eight became the longitudinal sample, completing all 6 tests (11 males and 17 females). These children were tested in their own home on the first occasion and subsequently in pre-school and then at primary school in Witham, UK. Six Himba children and eight English children moved home during the course of the study and became unavailable for further testing. All children, in both populations, had normal color vision as assessed by the City Color Vision Test (Fletcher, 1980). In addition, 24 adult English volunteers and 24 adult Himba speakers (paid for their participation) with normal color vision named the 22 stimuli.

All Himba children were rewarded for their participation with gifts of maize flour for their families and small toys. English children received stickers.

Materials.

The stimulus set consisted of two identical sets of 22 Color Aid matt surface colored squares measuring 2 inches square and backed with stiff card. Eleven of the stimuli were the best examples (focal colors) of the English basic categories *black, white, gray, red, green, blue, yellow, pink, orange, purple and brown.* The other eleven were chosen to be intermediate between each of the chromatic categories (e.g., half way between *pink* and *orange*), given the arrangement of color in a spherical three-dimensional space. Appendix A gives the designations and C.I.E. L*a*b* co-ordinates for each color.

Figure 1 shows the distribution of Himba color terms used by adults to name a range of 160 Munsell stimuli varying in hue (horizontal axis) and lightness (vertical axis) compared to the terms used by English adults. Together these terms were used to name 86.2% of the stimuli.

In addition a number of secondary terms, particular to the Himba dialect and normally used specifically to describe the color of animal hides (cattle, goats etc.) were also used by a number of observers (*vinde, vahe, kuze, honi*). These represented 8.6% of total names given.

For Himba speakers, those tiles most central to each of the five categories were calculated on the basis of adult naming agreement. For three categories (*serandu*, *zoozu* and *dumbu*), more than one tile was considered to be a best example on the basis of adult naming agreement (3 for *serandu*, 2 each for *zoozu* and *burou*). The nine tiles designated as best examples are hereafter referred to as 'focals' since this terminology has been consistently used in past research (e.g., Rosch, 1972, 1973).

However, the criterion used here (consensus of adult naming agreement) does not necessarily imply that these are the only good examples of adult Himba categories. Adult speakers of Himba, like the Berinmo participants tested by Roberson et al. (2000) and the Dugum Dani participants tested by Rosch (1972) showed little agreement when asked to select the best examples of their basic categories from a much larger set of colors (160). The categories of these languages may not have a single focal point for all speakers. Indeed, even for English speakers there is variability in choices of best example.

Procedure.

English children were tested by the Experimenter in their own home, in the first instance and subsequently in a quiet corner of their pre-school / primary school. Children were tested seated at a table near a window in conditions of natural daylight. Himba children were tested by the Experimenter with the help of an interpreter, in their home villages, seated at a table in shaded natural daylight. All responses were recorded by the Experimenter in the order in which they were given.

The children were asked to complete four tasks on all test sessions. The first was a listing task in which children were asked to "tell me all the colors that you know". Children who did not respond were encouraged "do you know the names of any colors?" Color naming and comprehension were tested with two tasks, one productive, in which children were shown each tile, one at a time, and asked - "what color is this?" followed by a comprehension task in which all 22 tiles were laid, in random order, in front of the child and they were asked - "show me a *red* / *serandu* one". When the child had made a selection, this was noted by the experimenter and the child was asked "is there another *red* / *serandu* one?" This procedure continued until the child responded negatively. Children were classified as knowing a term when they both correctly named the focal and pointed correctly to it when asked in the comprehension task, with the proviso that they not point to it in response to more than one other term (based on Soja, 1994 and Pitchford & Mullen, 2001). The lenient criterion of accepting one wrong selection was chosen since the continued questioning by an adult might encourage children to choose tiles more than once. All children also completed a memory task in which the complete array of 22 tiles was laid out in random order, in front of the child and then covered up. The child was shown a tile (from an identical set) for 5 seconds and told that the same colored tile was 'hiding' under the cloth. The tile was then removed and the array of 22 exposed without delay. The child was asked to find the color that they had just seen. Two practice trials were given, with feedback, using tiles not included in the set of 22, to establish that children understood the task. Reaction times were not formally measured on the recognition memory task, as stimuli were hand-presented. Informal measurement confirmed, however, that responses were typically made within 15 seconds of uncovering the test array, by children in both populations. The list task was always completed first and the naming task was always completed before the comprehension task. Order of tasks for the memory and naming/comprehension tasks was counterbalanced across participants and across test sessions.

Results

A) First testing

a) Children who know no color terms.

Children were considered as knowing a color term if they passed both the naming and comprehension tests. Table 1 shows the distribution of term knowledge at the first test. Our first analyses are restricted to those children who knew no color terms, since it is in this group of children that the predictions of universal BCTs can be most clearly tested. We will consider children who knew one or more color terms in the following section. Nine English children (6 girls, 3 boys) and twenty-six Himba children (11 girls, 15 boys) failed to pass the combined naming and comprehension criteria for any of their color terms at first testing. Of the nine English children, only one failed both tasks for all eleven focal items. The remaining eight passed either the naming or the comprehension task for a mean of 1.78 focal items (range = 1 – 7). Similarly, for the Himba children, only one failed both tasks for all nine focal items. The remaining twenty-five children passed either the naming or the comprehension task for a mean of 1.89 focal items (range = 1 – 8) and may have some partial knowledge of the appropriate set of categories for their language and culture. Thus, our strict criteria for allocating children as knowing no color terms are conservative for the prediction that their memory error patterns will be based on perceptual distance.

i) Memory confusions.

The strong proposal that color categories are universal, innate and independent of language would predict that both Himba and English children who knew no color terms would share the same set of cognitive categories (those that correspond to the English basic categories). Such organization of color categories should be evidenced by similar confusions in memory between the two populations, because colors belonging to the same category should appear more alike than those from different categories. However, it is also the case that items within the same category (two *reds*) are closer to each other in perceptual terms than, say, a *red* and a *green*. A pattern of memory confusions based solely on perceptual distance (rather than categorization) would also find more memory errors for perceptually closer items but there should be differences. For example, naming of poor examples of categories (e.g., *dark orange* /

dark blue) overrides the greater perceptual similarity to an exemplar of another name category (*red* / *black*) in favor of the categorical name (*orange* / *blue*). Therefore, we examined memory confusions with respect to both perceptual distance and proposed predetermined categories to see how far each factor would explain the data.

Separate 22 x 22 dissimilarity matrices were constructed for the memory confusions of English and Himba children. SPSS Replicated Multidimensional Scaling (RMDS, Young & Harris, 1994) was applied to the matrices. In a multi-dimensional solution of a matrix, were 2 stimuli always confused with each other, they would occupy the same point in space. Were they never confused, they would be placed as far apart as possible (given their relationships to all the other stimuli). In comparing two solutions, RMDS measures the departure from goodness of fit using S stress (squared stress values) that represent the distance that each point in the solution would need to be moved for the two arrays to fit perfectly over each other. Kruskal's (1964) primary approach was used for tied ranks.

The memory confusion matrices for those children who knew no color terms in each population were compared to the confusion matrix that results from judgments based on log perceptual distance (ΔE: calculated as the Euclidean distance between each pair of stimuli in the perceptually uniform C.I.E. L*a*b* space, Wysecki & Styles, 1982) and to the matrix derived from name similarity in English (based on adult naming of the 22 stimuli). For the full set of twenty-two stimuli, S stress values for each comparison are given in Figure 2.

The Mantel test was used to compare the relative strength of the relationships between pairs of matrices in (Mantel, 1967; Legrende, 2000). There was a significant relationship between Himba memory and log perceptual distance [$r = .343$, $p < .001$], English memory and log perceptual distance [$r = .494$, $p < .001$] and also between

English and Himba memory [$r = .551$, $p < .001$]. A Fisher's r' test revealed no significant difference between the strength of these relationships ($z < 1$). Thus for both groups of children who know no color terms, very similar errors were made and these appear to be closely associated with perceptual distance. In contrast, a comparison of the pattern that would be produced by English naming revealed no significant relationship to Himba memory [$r = .079$, $p > .2$], English memory [$r = .061$, $p > .2$] or log perceptual distance [$r = .108$, $p > .1$].

In summary, memory errors for children knowing no color terms were strongly influenced by perceptual distance. Importantly, the error patterns of both groups of children were clearly *not* influenced by the eleven basic color categories of English. Errors appeared to be based on perceptual distance rather than a particular set of predetermined categories. Navy *blue*, for example, was more often confused with the more perceptually similar *purple* or *black* than with focal *blue*. This point is addressed further in the longitudinal analyses.

ii) Memory for focal colors.

Another prediction of a universalist hypothesis comes from the fact that focal items are presumed to have a natural salience regardless of a particular language's terminology (Rosch, 1973). On this view, the same focal points will exist for languages with fewer terms, even if those terms reflect conjunctive categories, for example: *blue*-with-*green* or *yellow*-with-*green* (Berlin & Kay, 1969). Therefore, stimuli central to the eleven English categories (focal stimuli) should be remembered better than other items, even by individuals who cannot name them. Our stimuli were chosen to contrast naming and memory for the focal points of English color categories to stimuli that lie between these categories. While some of the eleven focal stimuli are also central to Himba categories, some are not; indeed, some colors that are focal for

Himba are peripheral to English categories. There were thus four different types of focal status amongst the stimuli: tiles that were focal for English only (*gray, brown, orange, purple, green and pink*), tiles that were focal for Himba only (dark navy *blue*, dark *orange*, *orangey-pinky-red*, *greeny-blue*), tiles that were focal for both languages (*black, white, red, yellow, blue*) or tiles that were focal for neither language (the seven remaining stimuli).

Table 2 shows the mean proportion of memory accuracy for both groups for each type of stimulus. A two (Language: English vs. Himba) by four (target type: English focal vs. Himba focal vs. focal in both languages x focal in neither language) ANOVA examined memory accuracy across the four different types of stimulus. There was a significant main effect of Language (English children were more accurate than Himba children): $F(1,34) = 14.47$, $\eta^2 = .426$, $p = .001$, and a significant effect of Target type: $F(3,102) = 5.05$, $\eta^2 = .149$, $p = .003$, but no significant interaction: $F(3,102) < 1$. To further investigate the effect of target type, data from both language groups were combined and a further one way (Target type: English focal vs. Himba focal vs. focal in both languages x focal in neither language) ANOVA conducted. There was a significant effect of Target type: $F(3,105) = 4.21$, $\eta^2 = .120$, $p = .008$. Post-hoc Newman-Keuls analysis revealed that only those targets focal in neither language were recognized significantly less often than those focal in both languages ($p > .01$). There were no other significant differences.

iii) List task.

Although children only offered color terms, there was no consistency about the terms listed. Of the nine English children who knew no color terms, four also listed none. Of the other five, two listed three, one listed five and two listed six terms. Of the twenty-six Himba children who knew no color terms, twenty-three listed none. One child listed two terms, one listed three and one listed four.

Discussion

For children with no color term knowledge, both English and Himba children make similar memory errors and, crucially, neither error pattern resembles that derived from the eleven basic terms of English. Furthermore, for these children, there was no specific advantage for colors that were focal to either language. However, colors not focal in either language were recognized significantly less successfully than those focal in either English, or Himba or both languages. Closer examination of the four sets, however, revealed that perceptual distance was not equal across the sets. In particular, the distance of each item from its nearest neighbors (ΔE) in the set that were focal in neither language was significantly smaller than those of the other three sets (all ps $< .05$). 67.5% of errors within the set focal in neither language were for the items with the smallest (ΔE^2) differences. Thus, the high rate of errors for this set derives from perceptual confusion and is unrelated to focality. For both groups of children, there was no significant difference in memory accuracy for the other three groups of stimuli despite the fact that some Himba focals are non-focal for English speakers and vice-versa.

Thus, before a child had acquired any color terms there was no evidence of any innate categorical organization. Memory confusions appeared to be based on perceptual distance rather than a particular set of predetermined categories. It would seem that the innate neurophysiological basis for perceptual discrimination is not itself sufficient to provide color categories (see also Gellatly, 1995). Indeed, the argument is stronger in light of the conservative criteria for defining color term knowledge. The majority of children who knew no color terms showed partial knowledge of their own set of color terms and volunteered only color terms in the list task, although these bore no relation to their knowledge. Such partial knowledge might have been expected to drive apart the cognitive organization of color for the two groups and make the memory patterns more closely resemble that of their community's adult naming. Such partial knowledge, however, does not appear to be enough to override perceptual similarity when memory confusions are made. We return to this issue in considering the longitudinal data.

b) Children who know 1 or more color terms

i) Order of term acquisition.

Universalist theories (e.g., Kay & McDaniel,1978) predicted, based on fuzzy set theory, that, for English speakers, primary terms (red, yellow, green, blue) would be learnt before secondary (orange, brown, purple, pink), because the primary terms would correspond to neural opponent process responses, while the secondary terms would correspond to combinations of responses. While the physiological basis for this argument has been shown to be false, many researchers continue to infer the case for heightened salience of primary over secondary terms on these grounds (see Saunders & van Brakel, 1997 and Jameson & Alvarado, 2003b for discussions). Indeed, previous studies of English speaking children have not upheld the prediction concerning primary terms (Mervis, Catlin & Rosch, 1975; Shatz et al., 1996).

The English data reported here show a generally similar pattern to that found in Pitchford and Mullen (2001) and overall, children knew more primary (mean = 2.37) than secondary (mean = 1.35) colors, [$t(58) = 7.45$, $p < .001$]. However, when the data were considered for individuals, there was little consistency in which terms were learnt first.

For the English sample, of the four children who knew one term, two knew *blue*, one knew *green* and one knew *pink*. Of the two children who knew 2 terms, one knew *yellow* and *red*, the other knew *yellow* and *green*. Three term combinations included: *brown, pink, black; brown, green, black; brown, blue red; yellow, white, green*. None of the eight children who knew 4 terms knew the same four and only two of the twelve children who knew 5 terms knew the same 5. To estimate the degree to which these patterns match the probability that primary terms are generally learned before secondary ones, the data were subject to a Guttman analysis (Hammond, 1990). Only 51 of the 2048 possible random combinations of known terms were observed. When matched to a model in which primary terms appear before secondary, only 3 children's profiles matched the model, while 48 did not, producing a coefficient of reproducibility of only .415 (upwards of .7 is generally considered useful).

Despite the smaller potential variation, a similar pattern was observed among Himba children. Of the sixteen children who knew only one term, seven knew *vapa*, four knew *serandu*, two knew *dumbu*, two knew *zoozu* and one knew *burou*. Of the ten children who knew two terms, five knew *serandu* and *vapa*, two knew *zoozu* and *vapa*, two knew *serandu* and *dumbu* and one knew *dumbu* and *vapa*. Every possible combination was observed for three and four term knowledge and no child knew all five terms.

Discussion

For children who knew at least one color term, there was little evidence, in either language, of a predictable order of acquisition. Item dependence was previously noted by Mervis et al., (1975) who reported that more children knew red and orange than knew purple and brown. In the present study, fewer English children knew brown and gray than other terms, supporting the findings of Pitchford and Mullen (2001) but three of the English children who knew just three color terms did know brown. Pitchford and Mullen argued that, in addition to restricted lightness levels, these colors have low perceptual salience, relative to other basic colors. Gray and brown, they suggested, are such common colors that they would give little predictive information for object identification and so are harder to learn. It might be that salience could be boosted for some individuals, for instance, by owning a brown dog or a gray cat. However, there is no evidence that children learn color terms from their reliable associations to particular objects. Indeed, it is unlikely given the neuropsychological evidence for a double dissociation in performance on the two tasks (Luzzatti & Davidoff. 1994; Miceli et al, 2001; Davidoff & Roberson, 2004). This issue is considered again in the examination of the longitudinal data.

Another similarity in performance between the two populations concerned the equivalence of naming and comprehension. In their attainment of color terms, we did not observe the predicted asynchrony of knowledge development (comprehension better than naming) found in other areas by Zelazo & Reznick (1991). In other tasks, comprehension may be easier because the name is given to the child and the range of items that can be matched to it is limited to the stimulus set. The naming task may be harder because the child must generate a label from a, potentially, infinite set. However, there are two possible reasons why an asynchronous pattern of knowledge development might not occur here. First, children from both cultures only volunteered color words in response to the question "what color is this", so knowledge of the appropriate lexical category restricted the potentially available set of responses. Second, children may have overextended their choices in color comprehension tasks as well as in naming tasks (a result found by Backsheider & Shatz, 1993). The consistent lack of asynchrony in both populations is an indication that children's difficulties in establishing color categories are at a conceptual level (i.e., where to place boundaries along a continuous dimension) rather than at a lexical level (i.e., selection of the appropriate label for a known group of stimuli).

ii) Memory for focal colors.

Memory for focal colors was tested using the same procedures as with children who knew no color terms. Table 3 shows the mean proportions of memory accuracy for English and Himba children for targets focal only in English, focal only in Himba, focal in both languages or focal in neither language. Accuracy for each different type of target was examined with a 2 (Language: English vs. Himba) by 4 (Target type: English focal vs. Himba focal vs. focal in both languages vs focal in neither language) ANOVA. There were significant main effects of Language [$F(1, 99) = 22.65$, $\eta^2 = .26$, $p < .001$], of Target type [$F(3, 297) = 34.79$, $\eta^2 = .35$, $p < .001$] as well as a significant interaction between Language and Target type [$F(3, 297) = 5.52$, $\eta^2 = .056$, $p = .001$]. Newman-Keuls pairwise comparisons of the interaction revealed that English children recognized significantly more targets than Himba children of those that were focal in both languages, focal in English and those focal in neither language (all ps < .05). English children recognized significantly more targets that were focal in both languages than any other type (all ps < .01) and more targets focal in English than those focal in Himba or in neither language ($p < .05$, $p < .01$ respectively). Himba children recognized significantly more of those targets focal in both languages than focal in English or in neither language (both ps < .01) and more targets focal in Himba than those focal English or in neither language ($p < .05$, $p < .01$ respectively). Thus, both groups of children recognized more of the stimuli that were focal than those that were not-focal in their own language.

The previous analysis took no account of the effects of color term knowledge on memory for focal colors. For a child who knew only one term, there was only one chance to correctly identify the focal tile of the term that they knew, but ten chances to correctly identify the focal tiles of the terms they did not know. The opposite was true for a child knowing ten terms. To examine the effect of term knowledge on memory, we therefore calculated the proportion of correct memory choices for terms known and not known for each child. Figure 3 illustrates memory accuracy for both groups by terms known.

For English children who knew at least one, but not all terms, the proportion of correct identifications for the focal chips was compared in a two (Age: 3 vs. 4) by two (Knowledge: known vs. not known) ANCOVA (with number of terms known as co-variant). There was only one reliable effect: a significant effect of knowledge [$F(1,55) = 18.45$, $\eta^2 = .335$ $p < .001$]. There were no significant effects of age [$F(1,55) = 1.943$, $\eta^2 = .035$, $p = .169$] number of terms known [$F(1,55) = 1.663$, $\eta^2 = .027$, $p = .133$] or any interactions [all Fs < 1]. For Himba children who knew at least one, and not all terms, the identical analysis again revealed a very similar pattern of results. There was a significant effect of knowledge [$F(1, 48) = 36.41$, $\eta^2 = .758$, $p < .001$], but no significant effect of age [$F(1,48) = 2.387$, $\eta^2 = .050$, $p = .129$], no significant effect of number of terms known [$F(1,48) < 1$], and no significant interactions [all Fs < 1]. Thus, the analyses showed, for both populations, that there was a stable relationship between memory accuracy and term knowledge that was unrelated to age. Children in both populations and at both age groups, at all levels of color knowledge, correctly remembered more stimuli for terms that they knew than for those that they did not know.

A calculation of the mean perceptual distance of memory errors also showed an effect of knowledge. Mean error distances are shown in Table 4.

An analysis by item, for children who made errors both on tiles for which they did and on tiles for which they did not know the term, revealed that, for English children, errors were made to more distant distracters when the terms were not known than when they were known: $t(10) = 3.89$, $p < .01$. The same was true for Himba children: $t(8) = 2.42$, $p < .05$.

The data were further investigated to establish whether passing both tests (our criterion for 'knowing' a term) was more beneficial than passing either test alone. These analyses considered separately, for each group, only the items that were focal in their own language, since it was not possible to judge children's naming as either correct or incorrect for items on which adult naming agreement was very low. Table 1 shows the number of children in each group at each stage of knowledge for each color term. Figure 4 illustrates the effect of knowledge status on memory. For the English focal items, a one way (Level: fail both tests vs. pass naming vs. pass comprehension vs. pass both tests) ANOVA on recognition accuracy showed a significant effect of knowledge status: $F(3,30) = 7.70$, $\eta^2 = .69$ $p < .001$. Newman Keuls pairwise comparisons showed that those targets for which both tests were failed were recognized significantly less often than those in all other conditions (all ps < .05). There were no other significant differences. For the Himba focal items, a similar one way (level) ANOVA again showed a significant effect of knowledge status: $F(3,27) = 3.08$, $\eta^2 = .35$, $p < .05$. Newman Keuls pairwise comparisons showed that those targets for which both tests were failed were recognized significantly less often than those for which both tests were passed ($p < .05$). There were no other significant differences.

Discussion

Not surprisingly, English children performed significantly better than Himba children at the recognition memory task; this was the case even for children at the youngest age group who had not yet started pre-school. However, and more important, effects of focality only appear as the child acquires color terms. English children remembered significantly more of the items that are focal in both languages and more of the items focal only in English. Himba children remembered significantly more of the items focal in both languages and of those focal only in Himba. Thus, focality is language dependent rather than universal. Knowledge of the appropriate term also reduces the distance of errors, when they are made and even partial knowledge of a term increases the probability of correct identification in the memory task.

iii) List task:

No child from either culture offered anything other than color terms (rather than other words) when asked to list the color words they knew. However, while English children listed only basic color terms, four Himba children included at least one secondary term. English 3-year-old children listed a mean of 1.26 terms (range 0:6). English 4-year-olds listed a mean of 2.73 (range 0:7). Himba 3-year-old children listed a mean of .85 terms (range 0:4), while 4-year-olds listed a mean of 1.40 terms (range 0:5). Children from both cultures often offered color terms that they could not identify. Of those terms listed for which a child did not know the term, 30% were to targets for which both tests were failed, 42% to targets that were only named correctly, and 28% to targets that were only comprehended.

Discussion

The list task was the only one for which color knowledge did not appear important. Like the children who knew no color terms, children frequently listed a term they did not know. Thus, children from both cultures seem to acquire the concept of color, and a set of terms that describe colors, before they can correctly apply those terms. A similar result has been found previously for English speaking children (Andrick & Tager-Flusberg, 1986; Davidoff & Mitchell, 1993; Mervis et al., 1995; Gottfried & Tonks, 1996). The lack of relationship between the listing of color terms and category knowledge is further evidence against the idea that learning color terms requires only the mapping of linguistic terms onto existing categories. The argument is stronger for English speaking children because they listed only basic terms.

B) Longitudinal data

By the sixth time of testing, all English children had completed two and a half years of education. At the sixth test, 29 Himba children had never attended school, while 34 children had attended one of the newly established mobile schools for between six months and one year. Attendance at school was determined solely by geographical location and not by ability. Where schooling was available, all children in a village attended. Instruction at mobile schools was given in Herero, and color terms were not included in the curriculum. However, as Herero has additional color terms (*pinge* [pink] and *grine* [green]) that are not used in Himba, in the subsequent analyses, findings are considered with and without school attenders.

i) Order of term acquisition.

The numbers of the 28 English children in the longitudinal study who passed the knowledge criteria for each term on each occasion are shown in Table 5. For the 63

Himba children in the longitudinal study, numbers of children passing knowledge criteria for each term on each occasion are shown in Table 6. Passing the criterion for three of the five Himba terms involved correctly naming and comprehending more than one tile. As this might be a harder criterion to reach than that required for English children (1 tile per term), numbers in parentheses in Table 6 show the number of Himba children passing criterion for each term when just one tile is considered for each color term. For the twenty-nine Himba children who had never attended school, the mean number of terms known at test 6 was 3.21. For the thirty-four children who had attended mobile schools the mean number of terms known was 2.74. There was no significant difference in term knowledge between these groups [$t(61) = 1.62$, $p > .05$].

At the 6th test for English children, one child knew five terms, one knew seven, two knew eight, six knew nine, seven knew ten and twelve knew all eleven terms. At the 6th test for Himba children, two knew no color terms, seven knew one, fourteen knew two, sixteen knew three, twenty-three knew four terms and only one knew all five terms. In the interval between each testing time (six months), most English children and some Himba children acquired more than one new color term. It thus was not possible to determine the order in which terms were acquired in a systematic way. However, two additional measures were used in an attempt to address the question of the order of term acquisition. In the first analysis, each basic color term was assigned an ordinal position for each child, (allowing for ties and based on the first time that the child passed the knowledge criteria for that term). The rank order of terms for English children across all six tests was: *yellow, red, orange, blue, green, black, pink, white, purple, brown, gray*. In an alternative analysis, the average age at which both tests were passed was calculated for each color term. Using this criterion, the order of terms for English children was: *yellow, red, black, blue, orange, white, pink, green, purple, brown, gray* (the average age of acquisition ranging from 44.13 months for *yellow* to 52.9 months for *gray*). Both measures are imprecise, given the time elapsed between tests and the fact that some children acquired up to five new terms between two testing sessions, but neither reveals a division strictly based on primary and secondary terms. Previous investigations of children's color term acquisition have assessed the age at which children 'know' color terms either on a single test (Pitchford & Mullen, 2001; Rosch, 1971; Macario, 1991) or over a short acquisition period (Andrick & Tager-Flusberg, 1986; Sandhoffer & Smith, 1999; Mervis et al., 1995). A longitudinal study allows us to verify that such estimates reflect permanent category acquisition rather than a transitory effect of intensive training.

For English children, early failures on the naming task exclusively involved incorrect use of a basic color term (e.g., white called 'pink'; orange called 'yellow'; brown called 'black'; red called 'orange') later failures largely involved incorrect use of secondary terms (e.g., red called 'lilac'; pink called 'skin colored'; purple called 'multi-colored'). Only one error at test 6 involved the wrong application of a basic term (pink called 'purple'). For Himba children, both early and late failures on the naming task frequently involved incorrect use of both basic and secondary terms (e.g., *zoozu* called '*dumbu*' or '*vinde*'; *serandu* called '*dumbu*' or '*honi*'; *vapa* called '*dumbu*' or '*vambi*').

For both populations, some children who passed the criteria for knowing a term on one occasion failed it on another. For the English children, the four terms considered 'primary' receive intense adult input and feedback in the first year of pre-school. Sixty-six per-cent of English children knew all of red, blue, green and yellow by the end of their first year of pre-school (test 3), compared to 42% who knew all the other basic terms but this figure rises to 55% with the exclusion of brown and gray. The difference in number of subsequent errors for terms that had been passed on a previous occasion for ‘primary’ terms (mean = 1.71) and ‘secondary’ terms (mean = 2.18) was not significant [$t(27) = 1.61$, $p > .1$]. Across all 6 tests, the average proportion of times on which children passed criterion for a term on one occasion, but subsequently failed was 8% (range 0 – 18%). For the Himba children, across all 6 tests, the average proportion of times on which they passed criterion for a term on one occasion, but subsequently failed on a following was also 8% (range 0 – 24%).

In summary, the rate of subsequent failure of one or other criterion, after having passed both on a previous test is small but very similar across the two populations, despite the English children receiving more intensive reinforcement in the intervening 6 months between tests. The present data would suggest that previous studies using only a single test give slight underestimates for the age of color term acquisition.

ii) Memory for focal colors.

The effect of focal status on memory over time was examined in each language, in similar analyses to those used for the first testing, with a 4 (Target type: English focal vs. Himba focal vs. focal in both languages vs. focal in neither language) by 6 (Time of test: 1 – 6) ANOVA. These results are illustrated in Figure 5. For English children, there were significant main effects of target type [$F(3, 486) = 62.42$, $\eta^2 = .39$, $p < .001$] and time of test [$F(5, 162) = 18.38$, $\eta^2 = .57$ $p < .001$] but no significant interaction [$F(15, 486) = 1.60$, $\eta^2 = .05$ $p = .069$]. Performance on all types of target improved significantly over time. To examine the effect of target type, data were collapsed across the six tests and a one-way ANOVA comparison conducted (Target type: English focal vs. Himba focal vs. focal in both languages vs.

focal in neither language). The main effect was significant [F (3, 333) = 30.6, η^2 = .27 p < .001]. Examination of the simple main effects revealed that targets focal in both languages were recognized significantly better than any other target type (all ps < .01) and targets focal in English were recognized significantly more often than those focal in Himba or in neither language (both ps <.01).

For Himba children, in the initial analysis of Target type by Time of test, there were significant main effects of time of test [F (5, 372) = 22.51, η^2 = p < .001] and of target type [F (2, 744) = 29.21, η^2 = p < .001] but no significant interaction [F (10,744) < 1]. Performance on all types of target improved significantly over time. To examine the effect of Target type data were collapsed across the six tests and a similar one-way ANOVA conducted (Target type: English focal vs. Himba focal vs. focal in both languages vs. focal in neither language). The main effect of Target type was significant [F (3, 753) = 34.38, η^2 = .21 p < .001]. Examination of the simple main effects revealed that targets focal in both languages and focal in Himba were recognized significantly more often than those focal in English or in neither language (all ps <.01).

The difference between the two languages, and the superior recognition of those items focal to the speaker's language persisted over time in spite of the overall improvement in recognition memory.

Overall English children's recognition performance was better across all tests. One possible explanation is that color terms are taught in English schools. However, in part, the increased recognition accuracy may be accounted for by the acquisition of more, smaller categories because the broader categories of the Himba language allow more within-category errors. By the final test, the number of tiles selected for each term by English children was close to one. Thus, knowledge of the appropriate term was a better memory aid for English children, for whom there is only one best example of each basic term in the set, than for the Himba, for whom there are several equally good examples of each category.

Memory performance across all 6 tests for focal items from categories for which children knew terms was compared to memory for focal items from categories for which they did not. Table 7 shows mean proportion of terms known across tests. For both groups, known items were identified significantly more often than unknown items in the first tests (all ts > 2.09, all ps < .07). For both groups, by test six, there was no significant difference in recognition accuracy between known and unknown items. These results are illustrated in Figure 6.

iii) Perceptual vs. categorical errors.

We compared the distance (measured in C.I.E. space) of errors to the focal items of known and not known terms over time. Figure 7 illustrates these results. In addition to a reduction in the number of errors over time, for both groups the mean perceptual distance of erroneous choices also decreased across tests. For English children, across all tests, there was a trend for smaller error distances when a term was known than when it was not [t (5) = 2.337, p = .06]. For Himba children, across all tests, error distances were significantly smaller when a term was known [t (5) = 4.332, p < .01].

In most cases, making more within-category errors would also reduce the perceptual distance of confusions. We therefore examined two cases in which perceptual and categorical errors could be directly contrasted. Of particular interest was children's performance on the navy blue tile, which lies perceptually between English focal blue and black. For English speakers, this tile is in the same category as the focal blue tile. For Himba speakers, however, it is in the same category as the black tile (and both are equally focal). Within the set of 22 tiles, there is a closer perceptual alternative than either of these; this is the English focal purple tile. If choices were only influenced by perceptual similarity the purple tile should be a more frequent erroneous choice than either the lighter blue or the black tile for both populations.

We examined across tests, the number of times that those children who made a recognition error for the navy-blue tile chose the focal blue, the black tile or the focal purple tile. Figure 8 illustrates these results. A 2 (Language: English vs. Himba) by 3 (Error type: perceptual vs. English category vs. Himba category) chi-square comparison of English and Himba children's errors at the first two tests (counting only the first error per child) revealed no significant difference between groups [$\chi^2_{(2)} = .68$, $p > .7$]. However, a 2 (Language: English vs. Himba) by 3 (Error type: perceptual vs. English category vs. Himba category) chi-square comparison of English and Himba children's errors at the final two tests (counting only the first error per child) revealed a significant difference between groups [$\chi^2_{(2)} = 9.81$, $p < .01$]. An examination of the residuals revealed significant differences between the two groups in number of choices of the blue tile [$z = 2.2$, $p < .05$], the black tile [$z = - 2.0$, $p < .05$] but no significant difference in choices of the purple tile [$z = .2$, $p > .4$]. Thus, initially both groups made errors of all types but by the final two tests significantly more children from each group make categorical errors appropriate to their own language. To eliminate the possibility that this result was skewed by the attendance at school of some of the Himba children during the final year of testing, the data were re-analyzed, removing the errors of all children who had attended school. The pattern was not affected; a chi-square comparison of the remaining data still revealed a significant difference between groups [$\chi^2_{(2)} = 12.27$, $p < .01$].

A similar comparison was carried out for children's performance on the dark orange tile, which lies perceptually between English focal red and focal orange. Figure 9 illustrates these results. For English speakers, this tile is in the orange category. For Himba speakers, however, it is in the same category as the red tile (and also focal). The red tile is also the closest perceptual alternative within the set. If choices were only influenced by perceptual similarity the most frequent erroneous choice for both populations would be the red tile.

We examined across tests, the number of times that those children who made a recognition error for the dark-orange tile chose either the focal red, or the focal orange tile. A 2 (Language: English vs. Himba) by 2 (Error type: English category vs. Himba category) chi-square comparison of English and Himba children's errors at the first two tests (counting only the first error per child) revealed no significant difference between groups [$\chi^2_{(1)} = 1.16$, $p > .2$]. A 2 (Language: English vs. Himba) by 2 (Error type: English category vs. Himba category) chi-square comparison of English and Himba children's errors at the 5th and 6th test revealed a significant difference between groups [$\chi^2_{(1)} = 11.98$, $p < .001$]. Again, by the final test, the two groups make categorical errors consonant with their own language. To eliminate the possibility that this result was skewed by the attendance at school of some of the Himba children during the final year of testing, the data were re-analyzed, removing the errors of all children who had attended school. Again, an analysis of the remaining data revealed a significant difference between groups [$\chi^2_{(1)} = 12.16$, $p < .001$].

iv) List task.

The probability of a child listing a term that they knew, rather than one that they did not know, was investigated over time. Figure 10 illustrates these results. For English children a one-way ANOVA (Time of test: 1 – 6) showed a significant effect of time [$F(5,145) = 11.02$, $\eta^2 = .28$, $p < .001$] with a significant linear trend [$F(1,150) = 53.44$, $\eta^2 = .27$, $p < .001$]. Over time children were increasingly likely to list terms that they knew than those that they did not know. For Himba children, a similar one-way ANOVA (Time of test: 1 – 6) showed a significant effect of time [$F(5,372) = 12.72$, $\eta^2 = .15$, $p < .001$] with a significant linear trend [$F(1,377) = 55.43$, $\eta^2 = .13$, $p < .001$]. Again, across tests, children were increasingly likely to list terms that they knew than those that they did not know.

Discussion

The longitudinal data show that the pattern of term acquisition is slow for both languages. Several of the English children could not correctly name and point to named examples of all eleven basic color terms at age six, even after three years of instruction. In the context of a gradually improving memory performance, the longitudinal data reinforce the first testing data showing that children gained a particular advantage in memory for the colors focal to their own language.

Our data point strongly to effects based on color term knowledge. Across all tests, children from both groups recognized a greater proportion of target chips for terms they knew than for terms they did not know; however, this is an effect that necessarily decreased over time. For a child who knows no color terms, it may be necessary to consider all 22 alternatives in the recognition phase. For a child who knows one term (e.g., pink) and who is shown a blue tile, the probability of being accurate should increase, because none of the 4 alternatives that could be called pink need be considered in the recognition phase. In the extreme case, once a child knows ten terms, the range of possible alternatives at recognition is dramatically reduced, since none of the known alternatives need be considered for a tile they do not know.

Thus, recognition for tiles for which the term is unknown should approach ceiling as a child acquires a full set of color terms.

Our longitudinal sample also provided data for the contrast between perceptual and categorical mechanisms underlying of color term acquisition. As the number of errors in the memory task decreased steadily over time, so did the perceptual distance of those errors from the target. For Himba children, in particular, the perceptual distance of errors to targets for which the term was known was significantly less than that to targets for which the term was not known. However, these data need to be considered in light of the increasing influence of naming and categorically derived performance.

The influence is highlighted by the case of a non-focal tile that falls into different categories for adult speakers of the two languages; there is a widening difference between the types of errors that children make over time. Speakers of each language showed an increasing tendency to make within-category rather than cross-category errors. The name-based errors eventually overcome the tendency to make errors to the closest perceptual stimulus.

General Discussion

The study set out to investigate the origin and development of color term knowledge and cognitive color categories in two very different languages and cultures. Himba children live in a sparsely populated arid environment; their homes, clothes, tools and artifacts are made from cattle products and their contact with the outside world is sporadic. The English children tested live in a town in a temperate climate, attend school, watch television and are surrounded by brightly colored man-made objects. Despite these differences and the difference in the number of terms eventually acquired, there were remarkable similarities between the two populations.

The first aim of the study was to compare the universalist and relativist explanations for the development of color categories. Most Himba children receive no formal education and for those that do attend school, knowledge of color terms is not a curriculum requirement. Despite such considerable environmental, linguistic and educational differences, universalist theories of color categorization would predict that both groups of children would have the same initial cognitive organization of the color space (11 categories) and that this organization would remain in place for both groups of children, despite the super-imposition of differing sets of linguistic categories (Bornstein, 1985; Rosch, 1973). We presented three main arguments against the universalist position from our new data.

The first argument concerns the errors made in recognition. Previous research with adults also used memory errors as the main data to argue for linguistic relativity (Roberson et al., 2000); however, that procedure has been recently criticized by Munnich and Landau (2003). They argued that language specific errors might merely derive from the participant using a verbal code to remember colors. In other tasks, children as young as 16 months attend to labels as well as to perceptual properties when encountering novel objects and often overlook perceptual similarity in favor of judging same-label items as most similar (Welder & Graham, 2001). So, language-based memory errors might not necessarily contradict the universal account of color categories. However, these arguments cannot apply to children who know no color terms.

For children with no color term knowledge, the pattern of both English and Himba memory errors appeared very similar and, crucially, neither pattern resembled that derived from the eleven basic categories of English. Both appeared to be based on perceptual distance rather than a particular set of predetermined categories. Thus, the present data argue against an innate origin for the 11 basic color terms in English. However, a large proportion of the world's major languages have the same number of color categories and one may ask why. It is possible that the 11-color organization yields the optimal combination of discriminability and cognitive economy for recognition and representation of large numbers of colors. If so, languages with fewer terms would gain by introducing / borrowing new terms, when increasing technological advances or contact with other cultures introduce a greater need to communicate more precisely about color. Nevertheless, even if the 11-term organization were found to be optimal, and eventually adopted by all cultures, it need not be innate. The present results suggest that children do not have a universal set of pre-determined categories, but rather gradually acquire the organization of categories that are appropriate to their own language and culture.

The second argument, from the present data, against the universalist position concerns the question of whether there is a predictable order in which color terms are acquired. Although the order of acquisition observed over time differed according to the measure used, no measure showed the pattern, predicted by universalist theory, in which primary colors (*red, blue, green and yellow*) were learned first. Over the course of the longitudinal study, neither population showed a predictable order of acquisition, although English children generally acquired the terms *brown* and *gray* later than other terms, supporting similar results found in previous studies of English children (e.g., Macario, 1991; Pitchford & Mullen, 2001; Mervis et al., 1975; Shatz et al. 1996). The present study supported previous findings of the lack of a predictable order of term acquisition in both languages.

The third argument against the universalist position concerns the role of a particular set of focal colors. For both Himba and English children who knew no color terms, there was no advantage in memory for terms focal in either language. However, an advantage for focal colors became evident as soon as children acquired color terms. Of those children knowing at least one color term at the first time of testing, English children, show superior memory performance for the items that are focal to English and to both sets of categories. Himba children show superior recognition for those items that are focal in Himba and in both Himba and English categories. Such rapid divergence in the cognitive organization of color for the two groups, from the time that the first terms are learnt, suggests that cognitive color categories are learned rather than innate.

The advantage present for focal colors at first testing increased throughout the longitudinal study. Thus, the importance that Rosch gave to focality in establishing categories seems justified from the present data; nevertheless, it is important to stress that the focality is not universal but, as shown both at first testing and longitudinally, language dependent. For English children, this effect is unsurprising since these are just the colors that are taught from the earliest age and most readily available in their playthings. For Himba children, focality was determined on the basis of adult naming agreement. Those targets deemed focal were those for which over 90% of adults agreed on the name. Other targets received little adult naming agreement. Himba children do not encounter constant presentation, through printing, dying and film of best example, highly saturated colors. In their environment only muted, natural colors are encountered, for which adult naming may often disagree. Children should then be faster to learn those colors that adults reliably call by the same name and hence the superiority for “focal” colors.

The second principal aim of the study was to examine the trajectory of color term acquisition in both cultures. Previous studies have found conflicting evidence about the age at which children reliably produce and use color terms appropriately (Backscheider & Schatz, 1993; Rice, 1980; Andrick & Tager-Flusberg, 1986; Sandhoffer & Smith, 1999; Macario, 1991) with estimates of the age of acquisition varying between two and six years. The present results suggest that children continue to refine their conceptual color categories for some years after they first show evidence of term knowledge for 'focal' colors. Children know that a set of terms refer to 'color' and can select color as a property on which to match objects as early as two years of age (Soja, 1994). In the present study, three-year-olds in both cultures listed only color terms when asked. However, some children, from both language groups, in this study could not correctly apply all their basic color terms (even though the English children had had three years of specific instruction). Although the expected discrepancy between naming and comprehension (Zelazo & Reznick, 1991; Zelazo et al., 1996) was not observed in either group, the acquisition of color terms did appear to be genuinely slow, effortful and error-prone in both cultures.

Not surprisingly, English children acquired their first color words earlier than the Himba. Greater exposure to colored objects and the increased cultural salience of color in Western society may contribute to an earlier conceptual understanding of color as a separable dimension. However, from then on the differences between the groups are less than marked than the similarities, which are clearly seen in their performance on the recognition memory tasks. We have already noted the similar importance that language-dependent focal colors have for both groups; even more important are the similar effects of color term knowledge. For both populations, once color terms are acquired, memory performance is determined by the number of terms known. Children make more correct identifications of focal items for terms that they know than for terms that they do not, regardless of the absolute number of terms known. Thus, the effect of term knowledge on memory cannot be an artifact of superior memory and language skills of children with higher general intelligence; children who know more terms get the same proportion of the items they know correct as those who know few. Knowledge of even one color term appears to change the cognitive organization of color, and from this point on there are language-dependent differences between the two groups. Once knowledge is acquired, it appears to restructure the cognitive organization of color in a reliable way and this restructuring relates to term acquisition per se, not to maturation or educational input.

Acquisition of term knowledge causes a reduction of memory errors and these change in nature over time. The effects of naming are particularly evident in the case of two items that are called by the same name in one language and by different names in another, such as navy blue, or dark orange. By the time children are six years old, the few errors that are made to these tiles are to within- rather than cross-category items, regardless of perceptual distance. It is not simply the case that improving memory allows children to make fewer and less distant errors.

In summary, the present study set out to examine when and how children acquire a set of color categories appropriate to their own language and culture (Roberson et al., 2000). The present results suggest that what is universal about the acquisition of color vocabularies is a gradual progression from an uncategorized organization of color based on perceptual similarity (where dimensions are viewed as continua) to a structured organization of categories that varies across languages and cultures. The increase in the influence of linguistic categorization on memory for colors is progressive and cumulative in both groups. Moreover, without intensive adult input, color category acquisition is universally slow and effortful.

References:

Andrick, G.R. & Tager-Flusberg, H. (1986) The acquisition of colour terms. *Journal of Child Language 13*, 119-134.

Backsheider, A. G. and Shatz, M. (1993) *Children's acquisition of the lexical domain of colour*. Paper presented at the 29th annual proceedings of the Chicago Linguistic Society, Chicago.

Berlin, B. & Kay, P. (1969) *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Boroditsky, L. (2001) Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*, 1-22.

Bornstein, M.H. (1985) On the development of color naming in young children: Data and theory. *Brain and Language 26*, 72-93.

Bornstein, M.H., Kessen, W. & Weiskopf, S. (1976) Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance 2*, 115-129.

Bowerman, M. & Choi, S. (2001) Shaping meanings for language: universal and Language specific in the acquisition of spatial semantic categories. In: M. Bowerman & S. Levinson (Eds.*), Language acquisition and conceptual development*. Cambridge: Cambridge University Press.

Boynton, R.M. (1997) Insights gained from naming the OSA colors. In C.L. Hardin (Ed.) *Color categories in thought and language*. Cambridge, UK: Cambridge University Press.

Carey, S. (1978) The child as word learner. In J. Bresnan, G. Miller and M. Halle (eds.), *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press, 264-293.

Cruse, D.A. (1977) A note on the learning of colour names. *Journal of Child Language, 4*, 305-311.

Davidoff, J., Davies, I. & Roberson, D. (1999) Colour categories of a stone-age tribe. *Nature 398*, 203-204 & 402, 604.

Davidoff, J. & Mitchell, P. (1993) The colour cognition of children. *Cognition, 48*, 121-137.

Davidoff, J. & Roberson, D. (2004) Preserved thematic and impaired taxonomic categorisation: A case study. *Language & Cognitive Processes*, *19,* 137-174

Fletcher, R. (1980). *The City University Colour Vision Test.* Windsor, Berks: Keeler Ltd.

Gellatly, A. (1995) Colourful Whorfian ideas: Linguistic and cultural influences on the perception and cognition of colour, and on the investigation of them. *Mind & Language, 10*, 119-225.

Gennari, S.P., Sloman, S, Malt, B. & Fitch, W.T. (2000) Motion events in language and cognition. *Cognition, 83*, 49-79,

Gottfried, G.M. & Tonks, J.M. (1996) Specifying the relation between novel and known: Input affects the acquisition of novel color terms. *Child Development, 67*, 850-866.

Guest, S. & Van Laar, D.L., (2000) The structure of colour naming space. *Vision Research, 40,* 723-734.

Guest, S. & Van Laar, D.L., (2002) The effect of name category and discriminability on the search characteristics of colour sets. *Perception, 31*, 445-461.

Gumperz, J.J. & Levinson, S.C. (1997) Rethinking linguistic relativity. *Current Anthropology, 32*, 613-623.

Hammond, B.J., (1990) Testing time: Advances in logic analysis. *IEE Review, 36*, 103-106.

Heider, E.R. & Olivier, D.C. (1972) The structure of the color space in naming and memory for two languages. *Cognitive Psychology 3*, 337-354.

Jameson, K.A. & Alvarado, N. (2003a) Differences in color naming and color salience in Vietnamese and English. *Color Research & Application, 28*, 113-138.

Jameson, K.A. & Alvarado, N. (2003b) The relational correspondence between category exemplars and naming. *Philosophical Psychology, 16*, 26-49.

Kay, P., Berlin, B. & Merrifield, W. (1991) Biocultural implications of systems of color naming: *Journal of Linguistic Anthropology, 1*, 12-25.

Kay, P. & Kempton, W. (1984) What is the Sapir-Whorf hypothesis? *American Anthropologist 86*, 65-78.

Kay, P. & McDaniel, C.K. (1978) The linguistic significance of the meanings of basic colour terms. *Language, 54*, 610-646.

Kruskal, J.B. (1964) Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. *Psychometrika, 29*, 1-27.

Legrende, P. (2000) Comparison of permutation methods for the partial correlation and partial Mantel tests. *Journal of Statistics and Computer Simulation, 67*, 37-73.

Levinson, S.C. (1996) Language and space. *Annual Review of Anthropology, 25*, 353-382.

Lin, H., Luo, M.R., MacDonald, L.W. & Tarrant, A.W.S. (2001) A cross-cultural colour-naming study. Part I: Using an unconstrained method. Color Research and Application 26, 40-60.

Lucy, J.A. (1992) *Language and diversity of thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Luzzatti, C. and Davidoff, J. (1994) Impaired retrieval of object-colour knowledge with preserved colour naming. *Neuropsychologia, 32*, 933-950.

MacLaury, R.E. (1987) *Color and cognition in Mesoamerica: Constructing categories as vantages*. Austin: University of Texas Press.

Macario, J.F. (1991) Young children's use of colour in classification: Foods and canonically coloured objects. *Cognitive Development, 6*, 17-46.

Malt, B.C. & Johnson, E.C. (1998) Artifact category membership and the intentional-historical theory *Cognition, 66*, 79-85

Malt, B.C., Sloman, S.A. & Gennari, S. (2003) Speaking versus thinking about objects and actions. In: D. Gentner & S. Goldin-Meadow (Eds.) *Language in mind: Advances in the study of language and thought*. Cambridge, Mass.: MIT Press.

Mantel, B.(1967) The detection of disease clustering and a generalized regression approach. *Cancer Research 27*, 209-220.

Mervis, C.B., Catlin, J. & Rosch, E. (1975) Development of the structure of color categories. *Developmental Psychology 11* , 54-60.

Mervis, C.B., Bertrand, J. & Pani, J.R. (1995) Transaction of cognitive-linguistic abilities and adult input: A case-study of the acquisition of color terms and color-based subordinate object categories. *British Journal of Developmental Psychology, 13*, 285-302.

Miceli, G., Fouch, E., Capasso, R., Shelton, J.R., Tomaiuolo, F. & Caramazza, A. (2001) The dissociation of color from form and function knowledge. *Nature Neuroscience, 4*, 662-667.

Mollon, J.D. (1999) Color vision: Opsins and options. Proceedings of the National Academy of Science, USA, 96, 4743-4745.

Munnich, E. & Landau, B. (2003) The effects of spatial language on spatial representation: Setting some boundaries. In: D. Gentner and S. Goldin-Meadow (Eds.) *Language in mind: Advances in the study of language and thought.* Cambridge: Mass., MIT Press.

Pilling, M., Wiggett, A., Özgen, E. & Davies, I.R.L. (2003) Is color categorical perception really perceptual? *Memory & Cognition*, 31, 538-551.

Pitchford, N.J. & Mullen, K.T. (2001) Conceptualization of perceptual attributes: A special case for color? *Journal of Experimental Child Psychology, 80*, 289-314.

Rice, M. (1980) *Cognition to Language: Categories, word meanings, and training*. Baltimore: University Park Press.

Roberson, D. & Davidoff, J. (2000) The categorical perception of colours and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*, 977-986.

Roberson, D., Davidoff, J. & Shapiro, L. (2002) Squaring the circle: The cultural relativity of good shape. *Journal of Cognition and Culture, 2*, 29-53.

Roberson, D., Davies I. & Davidoff, J. (2000) Colour categories are not universal: Replications and new evidence from a Stone-age culture. *Journal of Experimental Psychology: General , 129*, 369-398.

Rosch, E.H. (1971) "Focal" color areas and the development of color names. *Developmental Psychology, 4*, 447-455.

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

Rosch, E.H. (1973) Natural categories. *Cognitive Psychology 4*, 328-350.

Sandhoffer, C.M. & Smith, L.B. (1999) Learning color words involves a system of mappings. *Developmental Psychology, 35*, 568-679.

Saunders, B.A.C. & van Brakel, J. (1997) Are there non-trivial constraints on colour categorization . *Behavioral & Brain Sciences, 20*, 167-178.

Sera, M.D., Berge, C., & del Castillo Pintado, J. (1994) Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. Cognitive Development, 9, 261-292.

Sera, M.D., Elieff, C., Forbes, J., Burch, M.C., Rodriguez, W. & Dubois, D.P. (2002) When language affects cognition and when it does not: An analysis of grammatical gender and classification. Journal of Experimental Psychology: General, 131, 377-397.

Shatz, M., Behrend, D., Gelman, S.A. & Ebeling, K.S. (1996) Colour term knowledge in two-year-olds: evidence for early competence. *Journal of Child Language, 23*, 177-199.

Slobin, D.L. (1996) From "thought and language" to "thinking for speaking" In J.J. Gumperz and S.C. Levinson (Eds.), *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

Slobin, D.L. (2003) Language and thought online: Cognitive consequences of linguistic relativity. In: D. Gentner and S. Goldin-Meadow (Eds.) *Language in mind: Advances in the study of language and thought*. Cambridge: Mass., MIT Press.

Smith, L.B. (1984) Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. *Child Development, 55*, 363-380.

Soja, N.N. (1994) Young children's concept of color and its relation to the acquisition of color words. *Child Development, 65*, 918-937.

Sturges, J. & Whitfield , T.W.A. (1997) Salient features of Munsell colour space as a function of monolexemic naming and response latencies. *Vision Research, 37,* 307-313.

Valberg, A., (2001) Unique hues: An old problem for a new generation. *Vision Research, 41*, 1645-1657.

Webster, M.A., Miyahara, E., Malkoc, G. & Raker, V.E. (2000) Variations in normal color vision. II. Unique hues. *Journal of the Optical Society of America, A – Optics Image Science and Vision, 17*, 1545-1555.

Welder, A.N. & Graham, S.A. (2001) The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72*, 1653-1673.

Wyszecki, G. & Stiles, W.S. (1982) *Color science – concepts and methods, quantitative data and formulae*. New York: John Wiley & Sons.

Young, F.W. & Harris, D.F. (1994) Multidimensional scaling. In Marija J. Norusis (ed.) *SPSS Professional Statistics 6.1*. Chicago, Ill: SPSS Inc.

Zelazo, P.D., Frye, D. & Rapus, T. (1996) An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37-63.

Zelazo, P.D. & Reznick, J.S. (1991) Age-related asynchrony of knowledge and action. *Child Development, 62*, 719-735.

Appendix A

Designations of the 22 Color Aid tiles in the C.I.E. L*a*b* metric. Stimuli were measured under D65 (6500 dg K) and viewed under daylight that varied from 5500 – 7500 dg K. Whilst naturalistic viewing conditions vary slightly over time, this is likely to have added ‘noise to the data rather than any systematic confound.

Tile	L*	a*	b*	Focal in English	Focal in Himba	Most common English name for non-focals
1	71.11	0.75	-0.61	*grey*	*x*	*navy blue*
4	100	0	0	*white*	*vapa*	*dark pink*
6	44.12	6.06	7.77	*brown*	*X*	
23	81.63	36.51	72.4	*orange*	*X*	
28	39.6	25.81	-21.39	*purple*	*X*	*Yellow*
29	36.44	6.62	-17.36	*X*	*zoozu*	*Dark orange*
33	63.42	49.8	-1.72	*X*	*X*	
42	34.71	2.68	2.27	*black*	*zoozu*	
45	90.42	8.97	83.5	*yellow*	*dumbu*	*turquoise*
46	91.57	23.78	85.43	*X*	*X*	
47	75.21	61.19	59	*X*	*serandu*	
49	58.82	61.92	43.04	*red*	*serandu*	*Orange/red*
50	49.31	0.92	-48.11	*blue*	*burou*	*Yellow/green*
53	50.97	-31.61	-13.57	*X*	*burou*	*Peach/beige*
54	57.32	-42.37	-1.97	*green*	*X*	*Pink/peach*
59	88.65	35.45	15.16	*pink*	*X*	*Peach/mauve*
77	69.6	51.72	35.63	*X*	*serandu*	*tan*
78	74.21	-32.03	48.23	*X*	*X*	
79	70.05	53.54	22.62	*X*	*X*	
80	87.12	26.95	57.93	*X*	*X*	
81	83.23	39.45	53.08	*X*	*X*	
82	56.7	60.68	12.55	*X*	*X*	

Table 1. Number of English and Himba children at each level of knowledge for the focal items of each basic color term at the first test.

English children

Level	red	blue	green	yellow	pink	purple	orange	brown	black	white	gray
fail both tasks	12	7	8	10	17	25	20	23	12	14	51
name correctly	6	18	13	11	11	16	17	8	8	6	15
comprehend	12	4	6	4	8	4	8	23	10	10	0
pass both tasks	38	39	41	43	32	23	23	14	38	38	2

Himba children

Level	Serandu	Zoozu	Dumbu	Vapa	Burou
fail both tasks	15	33	21	14	47
name correctly	11	7	14	9	7
comprehend	12	15	13	7	10
pass both tasks	31	14	21	39	5

Table 2. Proportion mean accuracy for 9 English and 27 Himba children knowing no color terms for items that are focal for English only (n = 6), for Himba only (n = 4), or for both languages (n = 5). Standard errors in brackets.

	English	Himba
Focals for English only	.42 (.06)	.18 (.03)
Focals for Himba only	.43 (.08)	.19 (.04)
Focals for both languages	.50 (.12)	.23 (.04)
Focals in neither language	.25 (.08)	.13 (.02)

Table 3. Proportion mean accuracy for 59 English and 42 Himba children, knowing at least one color term, for items that are focal for English only (n = 6), for Himba only (n = 4), or for both languages (n = 5). Standard errors in brackets.

	English	Himba
Focals for English only	.48 (.32)	.27 (.03)
Focals for Himba only	.39 (.04)	.36 (.04)
Focals for both languages	.66 (.04)	.41 (.04)
Focals for neither language	.35 (.03)	.15 (.02)

Table 4. Mean error distance (measured in C.I.E. L*a*b* space) for memory confusions made by participants to focal colors for which they knew the term and those for which they did not.

	Known	Not known
English	38.70	112.64
Himba	43.91	74.18

Table 5. Number of children /28 passing both productive and comprehension tasks for the focal items of each of the 11 basic English color terms on each of 6 tests.

Time of test	Red	Green	Yellow	Blue	Pink	Purple	Orange	Brown	Black	White	Grey
1	15	11	17	12	10	4	10	4	13	11	2
2	15	17	17	18	13	14	12	8	18	15	9
3	20	20	19	19	14	14	12	5	15	18	9
4	22	25	27	24	19	17	15	15	24	25	11
5	22	24	24	22	23	23	20	18	26	26	18
6	28	28	26	25	22	27	22	22	28	28	20

Table 6. Number of children /63 passing both productive and comprehension tasks for the focal items of each of the 5 Himba colour terms on each of 6 tests. Numbers in parentheses represent the number of children passing the tests when only one focal is considered for each term.

Time of test	*Serandu*	*Vapa*	*Zoozu*	*Dumbu*	*Burou*
1	25 (31)	30	11 (13)	17	4 (10)
2	31 (42)	43	34 (38)	21	0 (13)
3	29 (47)	46	30 (32)	23	3 (18)
4	37 (55)	49	32 (34)	30	8 (14)
5	37 (55)	57	23 (27)	35	7 (21)
6	46 (58)	60	29 (29)	43	8 (15)

Table 7. Mean proportion of terms known across all 6 tests.

English children.

Test	1	2	3	4	5	6
	.354	.506	.581	.727	.831	.896

Himba children.

Test	1	2	3	4	5	6
	.276	.409	.416	.473	.505	.590

Figure 1. Distribution of Himba naming and choices of best exemplar for the 160 chip saturated array (for 31 observers) compared to those of English speakers for the same array. Numbers represent number of individuals choosing an exemplar as best example of the category. Munsell lightness values are indicated on the vertical axis; hue values are indicated on the horizontal axis.

Figure 2. Stress values for multi-dimensional scaling comparisons of memory confusions by Himba children knowing no color terms and English children knowing no color terms compared to those found for perceptual distances and for naming by adult English speakers.

Figure 3. Memory performance on focal stimuli for three- and four-year-old Himba and English children at first test, grouped according to the number of terms known.

Figure 4. Mean proportion of correct memory identifications of focal stimuli in their own language by English and Himba children at each stage of color term knowledge.

Figure 5. Mean correct memory identifications for English focals, Himba focals, focals for both languages and focals for neither language, by English and Himba children across all tests.

Figure 6. Mean number of memory errors to focal items known and not known, by all children in both groups across tests (* = $p < .05$, ** = $p < .01$).

Figure 7. Mean perceptual distance of errors to known and not known focal targets by English and Himba children across tests.

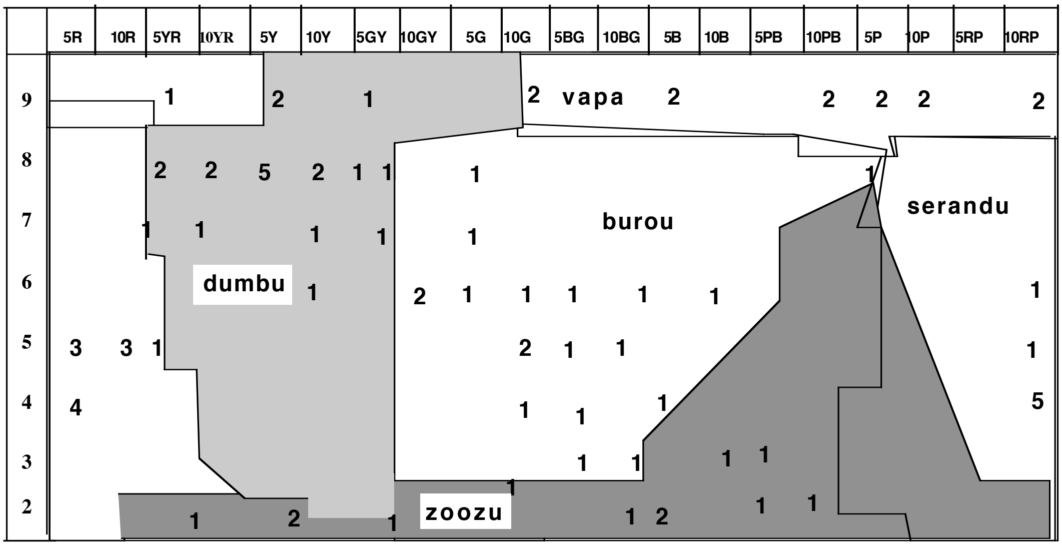
Figure 8. Mean proportion of erroneous identifications for the navy blue target tile to the black, best example blue or the closest perceptual match (purple).

Figure 9. Mean proportion of erroneous identifications for the dark orange target tile to the best example red (closest perceptual match) or best example orange tile.

Figure 10. Probability of listing known rather than unknown terms for both population across tests.

Figure 1

Himba naming distribution



English naming distribution

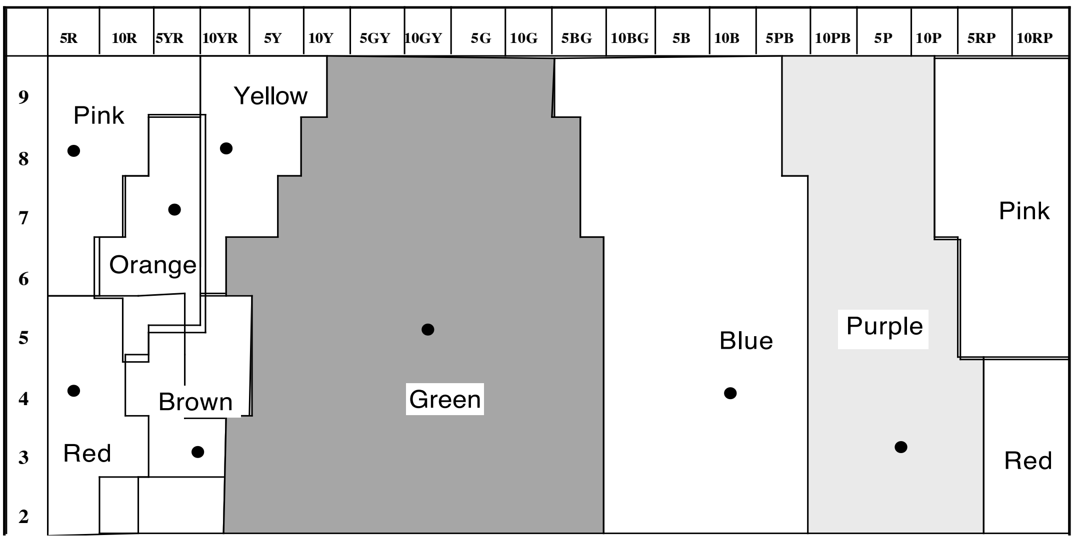


Figure 2

Note. The higher the score is, the poorer the fit.

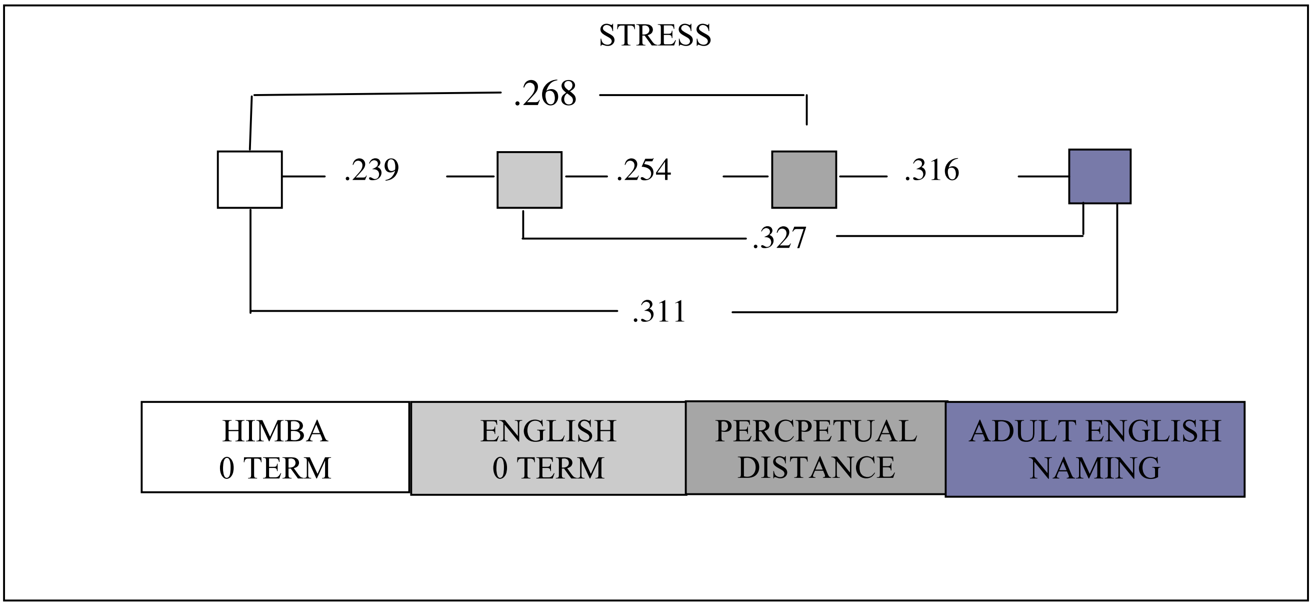


Figure 3

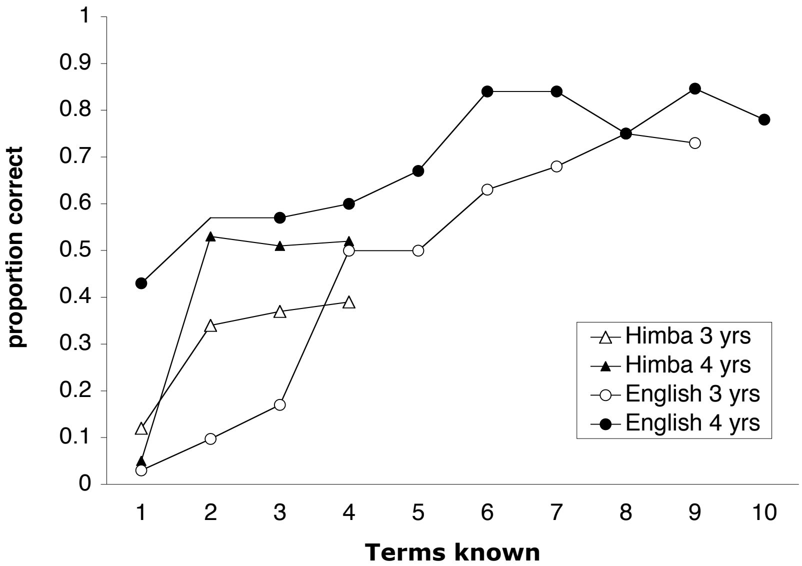

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
proportion correct
1
2
3
4
5
6
7
8
9
10
Terms known
Himba 3 yrs
Himba 4 yrs
English 3 yrs
English 4 yrs

Figure 4

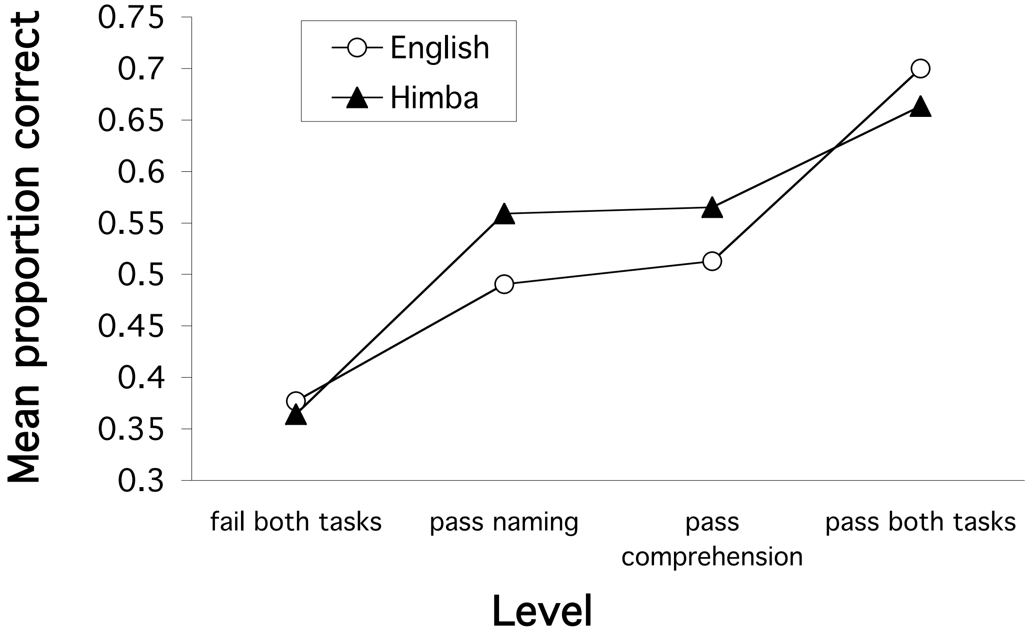
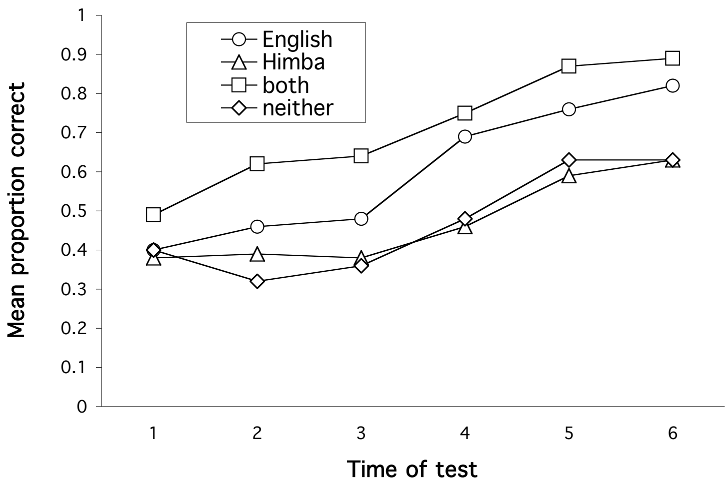

Mean proportion correct
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
English
Himba
fail both tasks
pass naming
pass comprehension
pass both tasks
Level

Figure 5

English children



Himba children

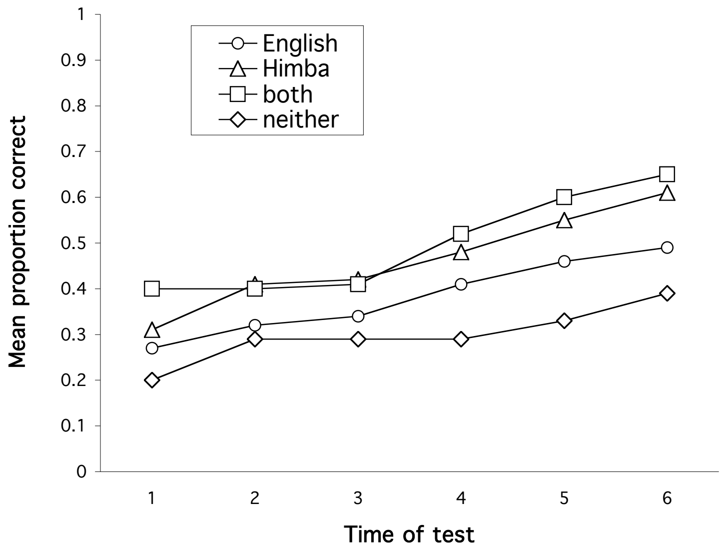
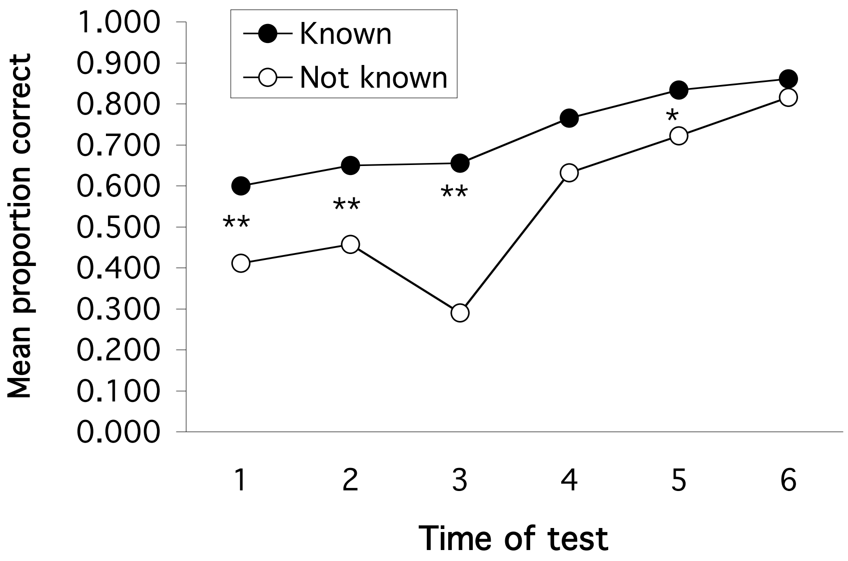


Figure 6

English children



Himba Children

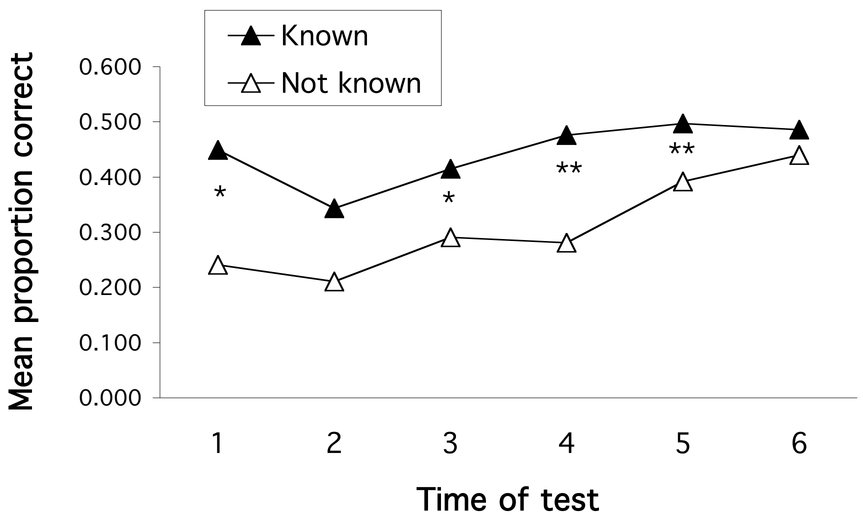


Figure 7

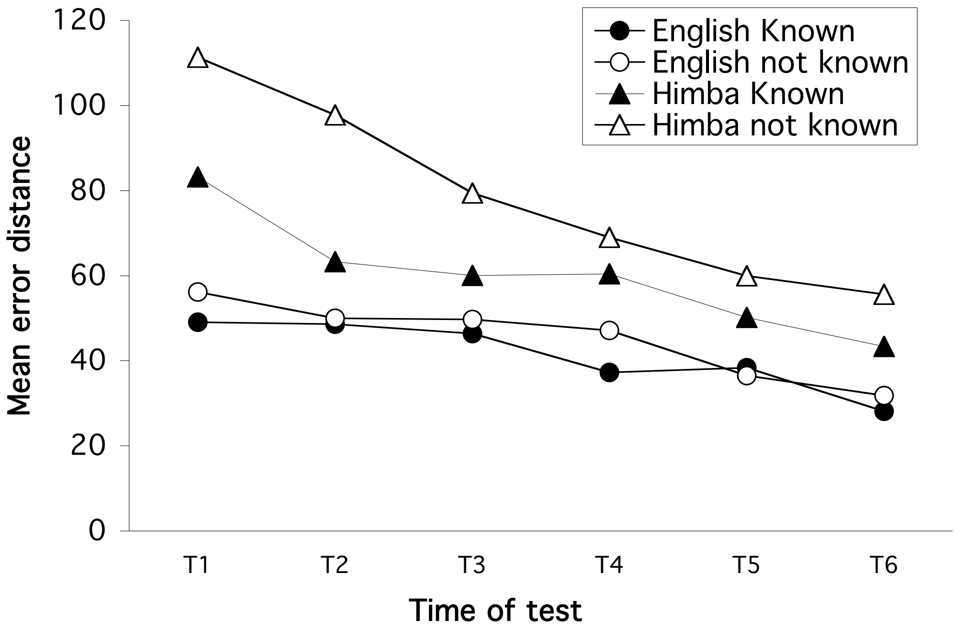
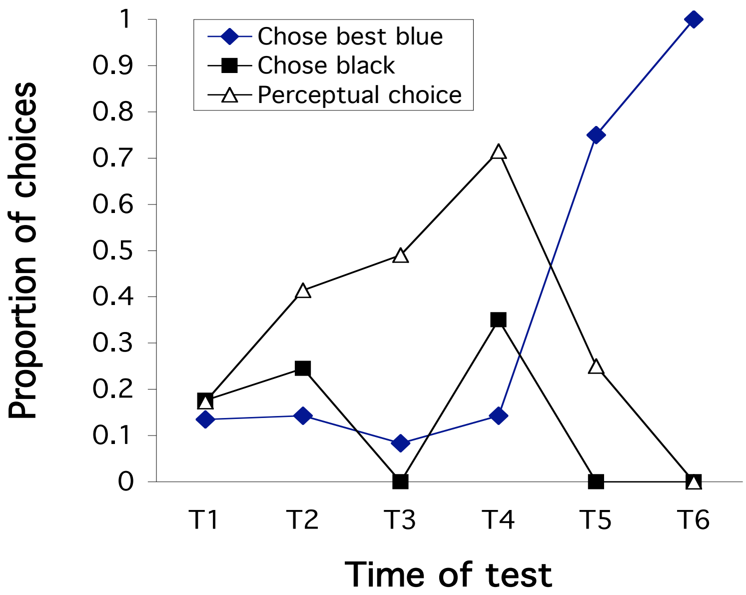

120
100
80
60
40
20
0
Mean error distance
T1
T2
T3
T4
T5
T6
Time of test
English Known
English not known
Himba Known
Himba not known

Figure 8

English children



Himba children

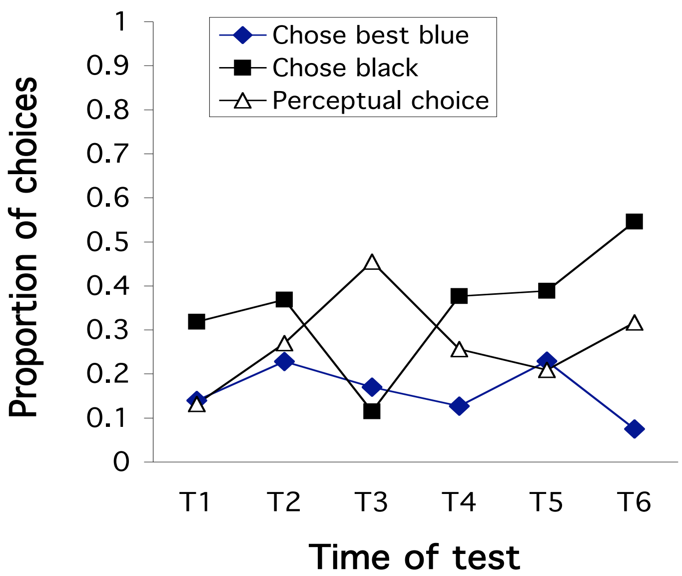
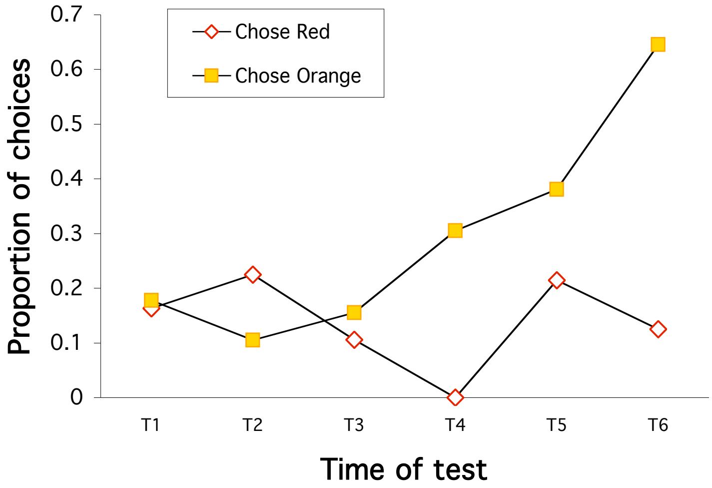


Figure 9

English children



Himba children

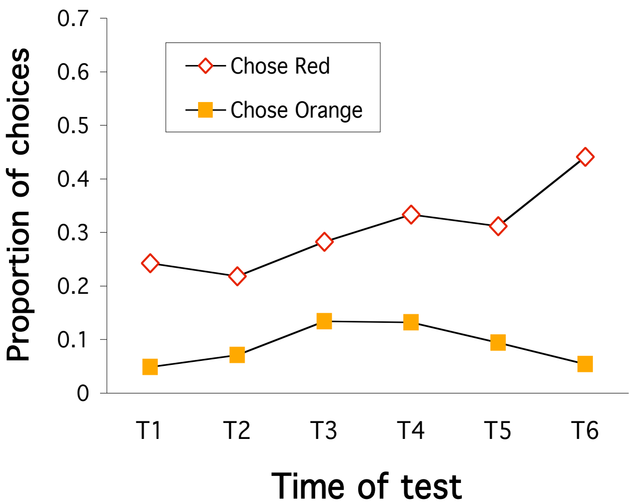
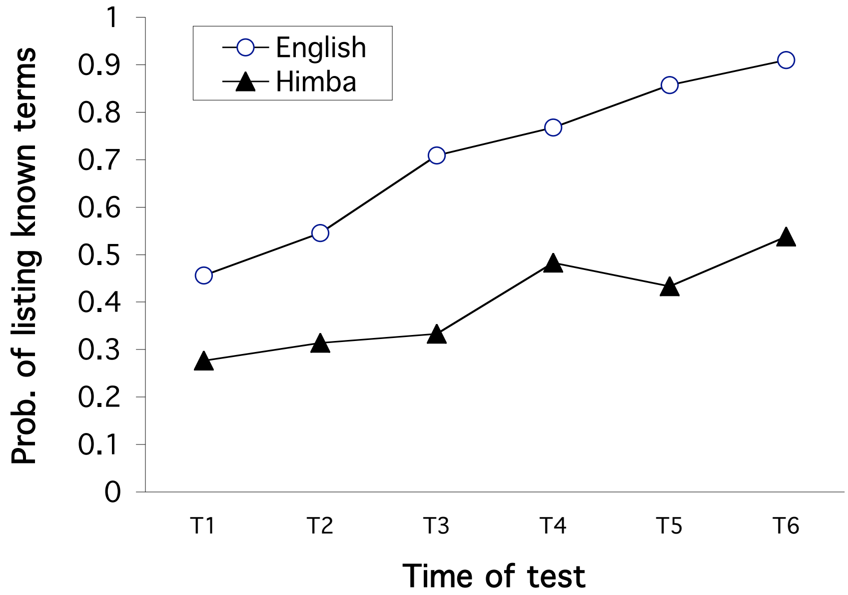


Figure 10



English
Himba
Prob. of listing known terms
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
T1
T2
T3
T4
T5
T6
Time of test

Author Note

Debi Roberson, Department of Psychology, University of Essex, UK; Jules Davidoff, Department of Psychology, Goldsmiths College, UK; Ian Davies, Department of Psychology, University of Surrey, UK; Laura Shapiro, Department of Psychology, Goldsmiths College, UK.

Laura Shapiro is now at the Department of Psychology, University of Warwick, UK.

This research is supported by Economic and Social Research Council Grant R0002383-10. We are grateful to Kemuu Jakarama who acted as interpreter and to the Himba participants in these studies. The stimuli and tasks were adapted from Boyles, 2001, PhD.

Correspondence concerning this article should be addressed to: Dr. Debi Roberson, Dept. of Psychology, University of Essex, Wivenhoe Park, Colchester, UK, CO3 4SQ. Tel:01206 873710, Fax: 01206 873590, email: robedd@essex.ac.uk.

Footnotes

1. Age estimation was carried out using a combination of measures. Taking family histories within each village established which children had been born around the time of independence (1991) and rough estimates of the order of births of other children after that time and the number of rainy seasons between the birth of siblings. Additional measures included checking that children could not touch the contra-lateral ear, with their arm raised over their head (normally attained at around 5 years), had all their milk teeth and were less than 120cms tall.