RELATIVELY SPEAKING;
AN ACCOUNT OF THE RELATIONSHIP BETWEEN
LANGUAGE AND THOUGHT IN THE COLOR DOMAIN.

Debi Roberson & J Richard Hanley

Department of Psychology,

University of Essex

Is the 'striped' appearance of the rainbow a construct of our higher mental processes, or is it determined at an early perceptual level by the organization of human color vision? If your language doesn't have separate terms for 'blue' and 'green' (and many languages, including Welsh, do not) would you perceive these shades as more similar than a speaker of English? Although the nature of the relationship between natural language and our mental representation of the experienced world has been probed by philosophers, psychologists, linguists and anthropologists in many areas (e.g. number: Gumperz and Levinson, 1997, Gordon, 2004; spatial relations: Majid, Bowerman, Kita, Haun and Levinson, 2004, Choi, McDonough, Bowerman and Mandler, 1999; time: Boroditsky, 2001; shape: Lucy, 1992, Roberson, Davidoff and Shapiro, 2002), the color domain has been a principle focus of investigation.

One reason why so much research has focused on color categorization is that the range of colors visible by humans is large (approx. 2 million Just Noticeable Differences), but the range of color terms available to describe them is generally small (languages contain between 2 and 22 basic[1] terms). Not only do some languages have gross differences in terminology (e.g. the use of a single term to refer to everything that an English speaker would call either blue, green or purple), but even those languages with similar color vocabularies have slight variations in the range of stimuli covered by a particular term (e.g. the different ranges covered by the English term 'blue' and the Italian term '*blu*').

Another reason why the color domain may be a particularly fruitful ground for

[1] The criteria set by Kay, Berlin & Merrifield (1991) for terms considered to be 'basic' are terms that are monolexemic, present in the ideolect of every observer, not subsumed within the range of another term.

investigating the relationship between language and thought is that the acquisition of color vocabulary by children is typically rather slow and error-prone compared to acquisition in other domains (Bornstein, 1985; Au & Laframboise, 1990; O'Hanlon & Roberson, 2006). By 2-years-of-age, children can learn a novel category label from a single exposure if it is a highly imageable concrete noun (Heibeck & Markman, 1987), leaving only the briefest of intervals to study category acquisition. In contrast, children may still make color-naming errors at 6-years-of-age (Roberson, Davidoff, Davies & Shapiro, 2004) so there is an extended period in which the process of color term acquisition can be studied.

This chapter is divided into six sections. The first sets out the background to the debate about the relationship between language and cognition in the color domain. The second explains how recent studies of color recognition employing visual search tasks have clarified this relationship. This section also argues that these studies point to the existence of two separate systems that influence perception and categorization of color, one of which is linguistically based and one of which is not affected by language at all. The third section critically evaluates recent claims by Kay, Regier and Cook (2005) that there are similarities between color terms in the world's languages that point to the existence of color universals. In the fourth section, we examine children's color term acquisition in an attempt to trace the mechanisms by which color categories are acquired. It also discusses whether infants have an innate pre-partitioned organization of color categories that is over-ridden during the learning process. In the two final sections, we outline some outstanding questions, note some methodological constraints on the conclusions that can be drawn from the accumulated evidence and argue that much more empirical investigation is still needed in this field.

Background to the debate.

Historically, the debate concerning color categorization has been sharply divided between what have come to be known as the universalist and the relativist positions. At one extreme, the view is that color categories are based on perceptual primitives, given either by the visual environment (Shepard, 1992) or by the properties of the human visual system (Kay & McDaniel, 1978), and are therefore universal. By contrast, the relativity hypothesis emphasizes the role of cultural needs in shaping both language and cognition: "the essential idea of linguistic relativity (is) the idea that culture, *through* language, affects the way we think, especially perhaps our classification of the experienced world" (Gumperz & Levinson, 1997, p 612). According to relativists, therefore, color categories are learned and are likely to vary as a function of cultural and linguistic differences (Ratner, 1989).

The wide variability of color naming that exists in different cultures is clearly consistent with the relativist proposal that language can influence which categorical distinctions a child comes to develop (Lantz & Stefflre, 1964; Ratner, 1989). However, as we explain later in this section, advocates of the universalist view have contested this conclusion by claiming that linguistic color categories and the mental organization of the perceived color space can be completely independent of one another (Rosch Heider, 1972a, 1973; Bornstein, 1985).

The relativist[2] view of a close link between linguistic categorization and the cognitive organization of color was based on the findings of Brown and Lenneberg (1954). They showed that colors that were easier to name in English were easier for (English speaking) participants to remember across short retention intervals and easier to communicate to others. Brown and Lenneberg, in common with many subsequent investigators, believed that memory performance provided insights into the way in which color categories are structured in the human cognitive system (we return later to the issue of whether this claim is justified). Cross-cultural support for the relativist view came from studies that demonstrated a close relationship between memory and naming in other languages also (Lantz & Stefflre, 1964; Stefflre, Castillo Vales & Morley, 1966). It was concluded that ease-of-naming would be a generally good predictor of memory accuracy in all languages across a wide range of color stimuli.

Berlin & Kay's (1969) survey of color terms in different languages led them to a radically different view. They suggested instead that some ways of organizing colors into categories were better than others. They proposed a common evolutionary trajectory for color vocabularies with the optimal arrangement as the endpoint. In that case, the number of color terms available in different languages reflected the point along that trajectory that a particular language/culture had reached. Western languages, like English, had reached the endpoint of that evolutionary trajectory and used an optimal set of color terms and categories. Pressure for color vocabularies to evolve towards an optimal set would arise because all humans shared underlying cognitive representations of the optimal organization (black, white, grey, red, yellow, green, blue, pink, purple, orange

---

[2] Often reported as a Whorfian view, following the writings of Benjamin Lee Whorf (1956)

and brown), even if they did not express those categories in their language. This organization at a language-independent cognitive level was proposed to be innate and hardwired into the organization of color vision pathways (Kay & McDaniel, 1978; Bornstein, 1985). Saunders & van Brakel (1997) review these proposals in detail.

Berlin and Kay's (1969) hypothesis was supported by Rosch Heider's studies of a small remote branch of a hunter-gatherer tribe, the Dugum Dani (Heider & Olivier, 1972; Rosch Heider, 1972a, 1973, 1975). The Dani used only two basic terms for the whole range of visible colors (although Rosch Heider 1972b did in fact suggest that many of her Dani participants used additional color terms). Her results indicated that Dani memory patterns were not well predicted by their color naming performance. Instead, they showed similar patterns of memory for colors to speakers of American English. For example, the Dani remembered colors that were good examples of English color categories (e.g. red, blue, green etc.) better than colors that were hard to name for English speakers. These findings supported the proposal that a particular set of color categories might be panhuman cognitive universals that could transcend terminological differences. Under this view there could be such large differences between the "structure of the color space in memory" and the structure of the lexical categories used to describe them (Heider & Olivier, 1972, p. 351) that the two sets (one in language, the other in thought) could be effectively orthogonal.

A complete disconnection of this nature between thought and language would have widespread implications for theories of cognition generally. Under such a universalist view, no learning or transmission of cognitive color categories would be required (since everyone would have the same set from birth). However, speakers of languages that express a different (non-optimal) number of categories would still at some point have to learn the appropriate reference set of exemplars for the linguistic terms that were used in their culture (Bornstein, 1985). Nevertheless, it remains unclear how a detailed category structure of this nature could be innately specified or how these differences between linguistic and cognitive categorization come to exist in certain cultures. It is also unclear why even English-speaking children appear to find color terms very difficult to learn (Bornstein, 1985) given that their language codes all members of the proposed universal categories.

Moreover, there are empirical as well as conceptual difficulties for this version of the universalist position. It is now clear that there are methodological problems with Rosch's experiments with the Dani that make her results difficult to interpret. Lucy and Schweder (1979) noted that the two sets of colors (best examples vs. poor examples of English categories) used by Rosch were not equally discriminable because the arrays were ordered by hue and brightness. It turned out that the best examples had fewer close competitors surrounding them. Garro (1986) repeated Rosch's experiment with English speakers using a randomized array and found that best examples of English color categories were still better remembered than poor examples. However these findings left unanswered the question of how the Dani would have performed with a randomized array.

In order to investigate this issue further, a new series of investigations was conducted that involved adult speakers of two other languages with a small number of color terms (Davidoff, Davies & Roberson, 1999; Roberson et al., 2000; Roberson, Davidoff, Davies & Shapiro, 2005). Both the Berinmo language, which is spoken in

Papua New Guinea, and the Himba language, which is spoken in Northern Namibia, contain only five basic color terms compared to the eleven present in English (see figure 1). When recognition memory for color was examined in both these cultures, Rosch's results were replicated so long as the arrays were ordered by hue and brightness. However, when the array was randomized, and the number of close competitors equated for the best and poor examples, the Himba and Berinmo no longer showed any memory advantage for English best examples. Instead, speakers of each language recognized good examples of *their own* linguistic color categories better than poor examples, regardless of these items' status in English color categories (Roberson et al., 2005). Paired-associate learning of colors to pictures of familiar objects also failed to show any advantage for the proposed universal prototypical examples in either Berinmo or Himba speakers. A similar lack of pre-eminence for this particular set, either in naming or categorization, was reported by Jameson and Alvarado (2003) in Vietnamese speakers. These findings show that there is no single set of prototypical colors that are universally cognitively privileged. Rather, those stimuli that are best examples of an individual's own named categories are remembered more easily than those that are not.

(Figure 1 about here)

In a further series of experiments, Roberson et al. (2000, 2005) investigated categorical perception (CP) of color in speakers of English, Berinmo and Himba. CP refers to the sharp peak in the relative discriminability of colors that cross a category boundary compared to discriminability within a color category (Harnad, 1987) so that continuous quantitative differences along a continuum come to be perceived as discrete qualitative changes at category boundaries. For English speakers, it is claimed that pairs of colors that cross the boundary (e.g. between blue and green) are discriminated faster and more accurately than pairs of colors with equal physical separation that are both good examples of green or blue.

The issue that Roberson et al. investigated was whether speakers of Berinmo and Himba would show also CP at the boundaries of the English categories green and blue. Moreover, would the Berinmo and Himba show CP at category boundaries within their own language that do not exist for English speakers. Participants were shown a colored target and had to decide which of two stimuli, presented five seconds later, was identical to the target. For each language tested, performance was facilitated when the target and distractor stimuli had different color names (e.g. in English, a blue target with a purple distractor) relative to the same name (e.g. in English, two different shades of blue). The results indicated that all three groups of participants showed CP, but only at color boundaries that were explicitly marked in their own language. Crucially, there was no effect of the proposed universal boundary between green and blue for speakers of Himba and Berinmo whose languages do not make this distinction.

One criticism that has been made of Roberson et al's (2000) findings with the Bernimo has been that the speakers of this language live close to the equator and that as a consequence their eyes may have been prematurely damaged by strong sunlight (UVB) (Lindsey & Brown, 2002). As the lens ages, it becomes denser and less clear, a process known as 'brunescence' which may particularly affect the ability to discriminate colours in the blue-green range. Lindsey and Brown gave young adults in the USA observation conditions (colored lenses) that simulated lens brunescence in the elderly and found that their naming classification of colors in the critical region changed. They suggested that individuals born with an innate boundary between green and blue, such as the Berinmo, could fail to distinguish them linguistically because by adulthood they would have lost discriminative ability in that region. However, such an argument cannot explain why Roberson et al. (2000, 2005) found no evidence of any deficits when they tested the color vision of their Himba and Berinmo participants. Hardy, Frederick, Kay and Werner (2005) provided the most direct test of Lindsey and Brown's hypothesis by examining older adults in the USA who were known to suffer from lens brunescence. Hardy et al. found that color naming for stimuli that were nominally green, blue-green or blue was virtually identical for observers with and without lens brunescence when viewing the same (unfiltered) stimuli. Thus it seems that the effects of lens brunescence cannot explain the differences that Roberson et al. (2000) observed between the performance of Berinmo and English speakers.

Recent studies of color categorization in adults.

Much of the research that we have discussed so far has employed memory tasks to investigate the relationship between language and color categorization. In these experiments, a strong link between naming and recognition might have emerged because individuals chose to rely on verbal coding to retain information about color during the retention interval. Even in Roberson et al.'s categorical perception experiments, a color had to be held in memory for 5 seconds before the target and distractor items were presented. Consequently, these studies leave open the question of whether linguistic coding affects perception or the ability to retain a color in memory (Munich & Landau, 2003).

Gilbert et al. (2006) devised a matching-to-sample visual-search task that appeared to make little or no demands on memory to investigate CP for colored targets. Participants were told to fixate on a cross in the centre of the computer screen. They were then asked to report the location of an 'oddball' colored target appearing amongst an array of distractors that were identically colored. Participants showed clear evidence of CP on this task. They were faster to detect a difference between the target and background when the target and background colors came from different categories (e.g. blue target, green background) than when both target and background came from the same category (e.g. different shades of blue) even when the amount of physical separation between target and background were held constant.

A critical question is whether CP on this visual search task occurs only at boundaries between colors in the putative universal set or whether it also occurs at boundaries that are not marked in English. This issue has been recently investigated with speakers of Russian by Winawer et al. (2007), and with speakers of Korean by Roberson, Pak and Hanley (2008). Winawer et al's Russian participants showed CP at the boundary between *siniy* (dark blue) and *goluboy* (light blue), which are distinct "basic" color terms for speakers of Russian. English speakers, who would call all these stimuli "blue", did not show the same cross-category advantage. Roberson et al.'s Korean participants showed CP at the boundary between *yeondu* (yellow-green) and *chorok* (green), which are distinct "basic" color terms for speakers of Korean but not for speakers of English. No evidence of CP was shown by native English speakers at this boundary.

Because the experimental tasks were not tests of memory, these two studies provide a clear demonstration that categorical perception of colors is constrained by culture and language. Consequently, these two studies provide overwhelming evidence that superior discrimination of stimuli that cross a category boundary (such as that found for English speakers at the boundary between blue and green) does not provide evidence for a set of universal color categories that are hard-wired in the human visual system.

CP effects were once thought to reflect low-level visual processing (Bornstein & Korda, 1984). Is it therefore the case that these results mean that linguistic processing affects early stages of color processing? Two sets of additional findings have provided information about the precise point at which verbal codes influence color categorization in this experimental paradigm. First, Gilbert et al. (2006) and Winawer et al (2007) showed that CP was not observed in perceptual tasks when participants carried out a concurrent verbal task. Under verbal suppression, all equally-spaced separations of color were equally easy to discriminate. Second, Gilbert et al (2006) and Roberson et al. (2008) reported that the CP effect was only found for colors that were presented in the right visual field, presumed to preferentially access language-processing areas in the left hemisphere. No difference between within- and across-category pairings of targets and distractors was observed for colors presented in the left visual field, which gains preferential access to the right hemisphere. Gilbert et al. (2006) also showed CP was found only in the left hemisphere of a patient in whom the corpus callosum, which connects the two hemispheres of the brain, had been surgically severed.

Further evidence that left hemisphere brain regions associated with language processing are actively associated with post-perceptual processing of color has been provided by a recent fMRI study (Tan et al., 2008). Easy-to-name colors evoked stronger activation in areas associated with language than hard-to-name colors. The authors suggest that these results support the rapid automatic activation of verbal color codes during perceptual decisions about color.

It is not hard to produce an explanation of the way in which left-hemisphere language processes might produce categorical perception. Let us assume that decisions about whether a target and background are the same can be made on the basis of either a right hemisphere perceptual code or a left hemisphere verbal code, and that when the two codes conflict, accuracy and speed will be reduced. Automatic activation of color names should therefore impair judgments about whether, for example, two different shades of blue are from the same category because the linguistic information that they are the same will conflict with the perceptual information that they are different. Decisions for items from different categories (e.g. a blue and a green) will be faster and more accurate because both linguistic and perceptual codes indicate that target and background are different. When the left-hemisphere language system is suppressed by verbal interference, or is not accessed because information is presented directly to the right hemisphere, the verbal code is not generated and there is never any source of conflict with the perceptual code. Hence there is no advantage for comparisons that fall across linguistic category boundaries.

If this account of categorical perception is true, then it follows that the ability to decide whether two colors are different probably depends on right hemisphere processing systems and does not require any form of verbal mediation. Color categories and color categorization, however, are entirely the product of the left hemisphere language system. Because verbal and perceptual codes for color are automatically activated relatively rapidly, there can be errors if the codes conflict. But where these two sources of information yield congruent information, memory for colored stimuli is good and decisions about color can be made accurately and rapidly. It therefore appears that categorical perception of color, contrary to what has often been claimed, does not in fact reflect superior discrimination of colors when they cross a category boundary. Instead it appears to reflect the fact that decisions about color are hampered when perceptual codes and verbal codes are in conflict. This conflict occurs when a task requires that two different shades of a primary color must be treated as different even though they share the same label.

Are there universal 'tendencies' in color naming?

Recent formulations of the universality hypothesis have acknowledged that differences in color category boundaries between languages influence memory for color and that linguistic boundaries determine the points at which categorical perception for color is observed (Kay et al. 2005). Thus there is now some common ground between the universalist and relatativist positions.

At the same time, however, Kay et al. (2005) still maintain that in different languages, there are strong universal tendencies both in color naming and in selection of the best examples of categories, which are held to cluster near the prototypes for English white, black, red, blue, green and yellow (Kay et al., 2005). However, in place of the eleven originally proposed universal categories (Berlin & Kay, 1969) recent formulations suggest that there are instead universal tendencies in color naming. For instance, they propose that the naming systems of different languages that all use 5 color terms are more similar to each other than would be expected by chance. Regier, Kay & Khetarpal (2007) suggested that instead of a single optimal system with 11 color terms, there might be optimal ways of dividing color space into 3, 4, 5, 6, etc. categories (see also Kay & Regier, 2007). They allow that these proposed 'optimal' partitions could be based on some properties of perceptual color space (Jameson, 2005), on some properties of the visual environment (Yendrikhovskij, 2001, Shepard, 1992), or on socio-linguistic negotiation among speakers (Steels & Belpaeme, 2005). They also acknowledge that even though the number of the languages surveyed to date are only a tiny fraction of the 6000+ languages still extant, many of them fail to fit the proposed optimal pattern.

There are, however, a number of methodological factors that may have inflated the apparent similarity of different naming systems in some investigations (Hickerson, 1971; Wierzbicka, 1999; Saunders & van Brakel, 2001). First, the range of stimuli used to collect cross-cultural naming data for the World Color Survey (Kay, Berlin & Merrifield, 1991) contains 320 highly colorful (saturated) stimuli (from the outer 'skin' of the Munsell sphere) that constrain the possible pattern of color labeling to a fixed set. Most traditional communities, lacking printing and dying facilities, may see such colorful stimuli for the first time when naming them for an experimenter. In some cultures speakers willingly extend their color terms to all these colors (e.g. Berinmo speakers). In other cultures, informants are more reticent, and many stimuli are left unnamed (e.g. Himba speakers; many South American informants interviewed by MacLaury, 1997). In addition, the possible extension of their color terms to less saturated stimuli (from the inner portion of the Munsell sphere) is rarely explored. Restrictions of the stimulus set may have led to an underestimation of the degree to which linguistic categories in one language might differ from those in another. In our own studies, the similarity between Himba and Berinmo naming patterns for the most colorful stimuli (.61 inter-language agreement) does not extend to less saturated stimuli (.27 inter-language naming agreement). Himba and English speakers use a large number of secondary terms to label less colorful stimuli (e.g. maroon, dun, olive, khaki) while Berinmo speakers readily extend basic terms to such stimuli.

In conclusion, reliance on the naming of only the most colorful stimuli may have led, in the past, to overestimation of the similarity of divergent languages' color term systems (Lucy, 1992; Lucy & Schweder, 1979; Saunders & van Brakel, 1997; Levinson, 2001). A similar observation of the limitations imposed by conducting cross-cultural naming studies with a restricted stimulus set is made in regard to cross-cultural studies of object naming by Malt, Sloman and Gennari (2003).

Second, for simplicity, the naming maps commonly produced contain only the modal names given, and only those modal terms deemed by the experimenter to be 'basic'. Consequently, much individual variation in naming is lost both within and across languages. Reporting only modal names reduces the size of cross-linguistic discrepancies between areas where name agreement is low or many participants fail to provide any name for stimuli (Jameson and Alvarado, 2003). It may also result in routine 'regularization' of large and complex data sets in order to decide which terms should be counted as 'basic' and included and which should be counted as 'non-basic' and excluded (see Saunders & van Brakel, 2001; Lucy, 1997). Indeed, some languages have been reported to have no 'basic' color terms at all (Kuschel & Monberg, 1974).

Even if there are genuine similarities between certain color systems, do we really need to invoke color universals to explain them? There are obvious cultural factors that could explain at least some of these similarities. First, similar cultural needs, such as evolutionary pressure for successful frugivory (Sumner & Mollon, 2000; Komarova, Jameson & Narens, 2007) could cause some category divisions to be more likely than others. The existence of dyes means that certain shades of color can be artificially generated much more easily than others, and may be labeled in many languages as a consequence. Second, it is clear that cultural contact between speakers of different languages has increased the similarity of the color categorization systems that these languages employ. For example, the introduction of the term 'burou' for colors in the blue-green range into Herero and subsequently into Himba came directly from the German word *blau* during the time that Namibia was a German colony.

In conclusion, therefore, we believe that methodological problems with the way in which the data have been collected renders unsafe any claim that there are universal tendencies that lead different cultures and different linguistic systems to divide up the color space in similar ways. Moreover, if there are genuine similarities between color categories in different cultures, it seems quite possible that they can be explained as easily by shared culture as by universal properties of the human visual system.

Children's color term acquisition

A number of recent computational models of color category learning have highlighted the importance of communication between agents in establishing an optimal set of categories for a perceptual continuum of color (Steels & Belpaeme, 2005; Belpaeme & Blys, 2005; Jameson, 2005). These studies have suggested that language plays an important role in initially establishing shared categories within a community, since simulations of category acquisition without communication among agents fail to speedily establish an optimal set.

Sometime between infancy and adulthood, children acquire a set of color terms and this linguistic categorization appears subsequently to influence their judgments of color. Given that color terms vary so widely across languages and that this variation has such profound behavioral consequences, the question arises as to when and how the differences come about. One possibility is that human infants are born with a pre-partitioned set of cognitive color categories that are universal and innately specified (e.g. blue, green, yellow, red and possibly also pink, purple, brown and orange), but that these are over-written during development by those categories in current use within the infant's culture and language (Bornstein, 1985). If so, given the evidence we have reviewed from color categorization in adults, language learning would appear to completely eliminate that original set.

Bornstein (1985) predicted that acquiring color terms would be more difficult for children learning a language in which an innate, hue-based, universal set must be over-written by a new set, even if the new set contained fewer terms in total. Bowerman and Choi (2001) suggested that language acquisition would have to overcome great resistance in order to re-structure mental life, where *any* robust and pre-potent organization of the perceived world exists pre-linguistically. Thus, the acquisition of a new set of named categories whose divisions cut across a proposed universal set should show a radically different developmental pattern to that of English-speaking children who would only have to learn to map appropriate labels to a set of already present cognitive categories.

Roberson et al. (2004) tested these claims by comparing the color naming and memory of young children in the UK who were learning English and children from Namibia who were learning Himba. Naming and comprehension were studied systematically over a three-year period in order to establish a reliable measure of children's color term acquisition, in speakers of different languages. They tested 28 English 3-year-olds before they entered pre-school and, subsequently, through three years of formal education. They also tested 63 Himba 3-year-olds, few of whom received any formal education during the period of the study. Children's color term knowledge and memory for colors were tested at six-month intervals over three years. The children completed a color term listing task ("tell me all the colors that you know"), color naming ("what color is this?"), color term comprehension ("can you find a red one?") and a recognition memory task in each of the six testing sessions.

Despite the considerable environmental, linguistic and educational differences between the two groups, the process of color-term learning appeared to be remarkably similar in the two groups. There was no predictable order of category acquisition across either group, consistent with other recent studies (Macario 1991; Mervis, Bertrand & Pani, 1995; Pitchford & Mullen 2001; Shatz et al. 1996). Within each group individual children displayed almost every possible order of acquisition and, at the end of the study, there were still some children from both language groups who could not correctly use all their color terms (even though the English children had had three years of specific instruction).

In recognition memory, from an initial reliance on perceptual similarity, an influence of language categories became evident as soon as children acquired color terms. Himba and English children who knew no color terms, showed similar patterns of memory errors and, critically, both patterns appeared to be based on perceptual distance rather than a particular set of predetermined categories. Of those children knowing one or more color terms at the first time of testing, Himba children showed better memory for the items that are good examples of Himba, but not of English categories, while English children showed the reverse pattern. Such rapid divergence in recognition patterns for the two groups, from the time that the first terms are learnt, suggests that color categories in both languages are learned using similar mechanisms. These data, like those for adult Himba and Berinmo speakers, argue for a pivotal role of language in shaping color categorization. Considering the trajectory of color term acquisition in the two cultures, the longitudinal results suggested that children continue to refine the range of referents for each of their color terms for some years after they first show evidence of term knowledge for the best exemplars of categories.

For both populations, once color terms were acquired, memory performance was determined by the number of terms known. Children identified more examples of terms that they knew than of terms that they did not know, regardless of the absolute number of terms known. Knowledge of even one color term changed patterns of color recognition, and from this point on there were language-dependent differences between the two groups. In addition, the type of recognition errors made by each group of children diverged over time, so that more errors were made to best examples of the appropriate language categories than to other alternatives, even though some items that were best examples of a category in one language, were poor exemplars in the other.

Overall, children from the two cultures seemed to acquire their color terms in the same gradual fashion, and knowledge of the appropriate terms influenced memory accuracy and memory errors so that the patterns of performance, from a common beginning, diverged increasingly over time. There was no evidence that English-speaking children learnt their color terms more easily than Himba children, even though the English terms map directly onto a proposed innate set. Nor was there evidence that either group of children appeared to have a pre-partitioned representation of color at 3-years-of-age, before they learn color terms (but see also Franklin et al., 2005).

Non-human primates have similar wavelength discrimination to humans (Sandell, Gross & Bornstein, 1979; Matsuzawa, 1985). So, if there is a set of universal color categories, one might also expect to find evidence for their existence in these animals. However, a recent study of color discrimination in baboons (Fagot, Goldstein, Davidoff & Pickering, 2006) failed to support the hypothesis that color categories are explicitly instantiated in the primate color vision system. In a match-to-sample paradigm, in which human participants showed a sharp category boundary between blue and green, none of the baboons showed any inclination to partition the range of blue and green stimuli on which they were trained into two categories despite good color discrimination.

Current evidence thus suggests that if there is an innate set of cognitive categories that are present in young infants, then a) they are species specific and thus do not result from some property of the visual system that is shared with other primates; and b) they are not retained once adult linguistic categorization is in place. An alternative possibility might be a scenario for color vision similar to that observed for auditory stimuli. In the auditory case, infants up to the age of 6 months appear to be sensitive to a wide range of categorical differences, including some that are not marked in their native language, such as the phonemic distinction between '*l*' and '*r*' for Japanese speakers (Werker & Tees, 1984). Their auditory system becomes selectively tuned to the appropriate categories for their native language sometime in the second six months of life. After 12 months of age, infants lose the ability to make some distinctions that younger infants make successfully. Such a possibility has not yet been investigated in infants with regard to color categorization.

Finally, a number of studies show that children achieve competent use of color terms relatively late compared to their acquisition of terms for other dimensions (Andrick & Tager-Flusberg, 1986; Mervis et al., 1995; Braisby & Dockrell, 1999; Sandhofer & Smith, 2001). Nevertheless, children appear to understand that color terms form an independent lexical semantic category by 2 years of age (Backsheider & Shatz, 1993; Sandhofer and Smith, 1999). At this stage, knowledge of color terms seems unrelated to the ability to use them. For example, when asked "tell me all the colors that you know", 3-year-old English and Himba children were as likely to list terms that they were unable to use correctly as terms that they could use correctly (Roberson et al., 2004). Learning color terms may promote selective attention to color, so children only achieve a comprehensive conceptual representation of the color domain after acquiring a sizeable color vocabulary (Sandhofer and Smith, 1999). A difficulty in learning the referents for novel color terms might arise because the ability to abstract any object properties (color, size, form and motion) all develop slowly (Pitchford and Mullen, 2001) or because children are predisposed to attend to object shape when interpreting novel object labels (Smith, Jones & Landau, 1992; O'Hanlon & Roberson, 2006). Whatever the explanation, it is clear that learning the appropriate set of referents for color terms is a more difficult task for young children than one might have expected if they were simply learning appropriate labels for innately specified universal color categories.

Outstanding questions.

Although it appears that categorical effects of color in adults depend on access to the language system, some studies have reported evidence of categorization of color in 4-months-old infants. Bornstein, Kessen, and Weiskopf (1976), Catherwood, Crassini, and Freilberg (1990), and Franklin and Davies (2004) have claimed that young infants do make apparently categorical distinctions of color continua. However, a study by Davidoff, Roberson, de Haan and Davies (2005) found no differences in novelty preference for changes in color that were either within or across adult category boundaries. Much more research is needed to establish why minor changes in experimental paradigm produced such different results. If this evidence of pre-partitioning proves robust, more research is also needed to examine how and why it is lost when language is learnt.

What drives so many different cultures to arrive at even coarsely similar solutions to the problem of categorizing the continuum of visible colors? We should be wary of assuming that those similarities that do exist provide evidence for color universals before alternative explanations have been fully explored. If an eventual set of eleven basic categories were in some way optimal, why would some cultures maintain a small set of linguistic categories in their own language while surrounded by other languages that have larger sets? Why would Russians and Koreans develop additional basic color terms beyond those used in English? Whatever the origin of the observed differences between the color terminologies of different societies, any comprehensive model of color categorization needs to explain both the observed similarities and the differences between color naming systems. The origins of linguistic color categories in different societies might be constrained by either cultural or environmental needs, or both, and both may change, over time, in different ways in different communities. For discussion of these issues in domains other than color, see Nisbett, Peng, Choi, & Norenzayan (2001), Sera et al. (2002) and Wierzbicka (1999, 2005).

There remain several other outstanding questions that are both fundamental to the debate and beyond the scope of empirical investigations to date. Is the development of adult color categorization a unique case? If not, to what extent does it follow a similar pattern to other modalities that come to be perceived categorically? In studies of object classification, evidence from cross-linguistic studies best fits a hybrid model in which some broad, shared, non-linguistic understanding of a domain combines with varying cultural pressures to differentiate particular aspects of a dimension at particular times in their history (Malt et al., 2003). Recent computational models of color category instantiation support the view that a combination of shared domain structure (in terms of both available learning mechanisms and the range of visible colors in the environment) and language is needed to explain shared color category structure. Such a combination of factors might lead to differences between linguistic categorization systems that also vary depending on the degree (and nature) of interactions between linguistic communities.

Conclusion

Early research in the field led to the conclusion that there are separate levels of categorical representation of color, one cognitive and impervious to language and another more superficial linguistic level (Rosch Heider, 1972a; Heider & Olivier, 1972). The evidence that we have reviewed in this chapter also points to the existence of distinct linguistic and non-linguistic color systems. However, there is no evidence that the linguistic system is in any way superficial. Linguistic categorization in different languages and cultures partitions the same range of visible colors in different ways and these differences affect decisions about color even on visual search tasks. Evidence suggests that categorical effects in color perception and memory occur as a result of access to lexical codes for color in adults. Moreover, children appear to acquire adult-like patterns of discrimination and memory for color as soon as they learn the appropriate color terminology for their language and culture. This argues against the view that linguistic categorization of color is superficially overlaid on some more important cognitive structure.

When linguistic categorization is prevented (in adults), or is not yet in place (in the case of young children) participants behave as if they perceive an undifferentiated continuum of just-noticeable-differences. There does, therefore, appear to be a separate non-linguistic system (possibly in the right hemisphere) that can make extremely fine discriminations between colors and decide whether two colors are identical or not. There is no evidence that language learning has any effect on the way that this system processes color. We do not believe, however, that this system ‘knows’ precise information concerning similarities and differences between two shades of color (e.g. that one is brighter than another, that one is more saturated than another, or that two different shades may share the same name). We do not believe, therefore, that this is the system that makes us see the rainbow as comprising seven distinct colors. Categorical knowledge of this kind is only available to the left hemisphere language-based color system, and people with different linguistic categories may well see a smaller or greater number of colors in the rainbow as a consequence.

Theorists who have supported the universalist position in the past now accept that linguistic differences between speakers of different languages influence color categorization (e.g. Kay et al., 2005). The relativist position has also been modified because theorists who support the relativist position acknowledge the existence of a separate color processing system that is completely independent of linguistic influence. Nevertheless, some important differences still remain between the relativist position that has been put forward in this chapter and the universalist position put forward by Regier and colleagues in the current volume. We believe that further investigation of the remaining controversial issues is needed. It is also important that dual process theories of color perception of the kind advocated in this chapter are subjected to rigorous empirical scrutiny. The domain of color has been and remains a fruitful ground for examining the relationship between language and thought. Now is not the time to discontinue the debate or the investigation.

Author information

Prof Debi Roberson, Dept. of Psychology, University of Essex, Wivenhoe Park, Colchester, Essex, UK, CO4 3SQ. Tel: 44-1206 873710, Fax: 44-1206 873590, email: robedd@essex.ac.uk Prof J Richard Hanley, Dept. of Psychology, University of Essex, Wivenhoe Park, Colchester, Essex, UK, CO4 3SQ. Tel: 44-1206 874331, Fax: 44-1206 873590, email: rhanley@essex.ac.uk

Acknowledgments

We are grateful to Barbara Malt and Phillip Wolff for helpful comments on an earlier draft of this chapter.

References

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

Au, T.K., & Laframboise, D.E. (1990). Acquiring color names via linguistic contrast: The influence of contrasting terms. *Child Development, 61*, 1808-1823.

Backsheider, A.G., & Shatz, M. (1993). Children's acquisition of lexical domains. Paper presented at the twenty-ninth regional meeting of the Chicago Linguistic Society, Chicago.

Belpaeme, T. & Bleys, J. (2005) Explaining universal color categories through a constrained acquisition process. *Adaptive Behavior, 13*, 293-310.

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

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. J*ournal of Experimental Psychology: Human Perception and Performance 2*, 115-129.

Bornstein, M.H. & Korda, N.O. (1984) Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research, 46*, 207-222.

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

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* (pp. 475–511). Cambridge, England: Cambridge University Press.

Braisby, N, & Dockrell, J. (1999) Why is color naming difficult? Journal of Child Language, 26, 23-47.

Brown, R. & Lenneberg, E. (1954) A study in language and cognition. *Journal of Abnormal and Social Psychology, 49*, 454-462.

Catherwood, D., Crassini, B. & Freiberg, K. (1990) Infant response to stimuli of similar hue and dissimilar shape: Tracing the origins of the categorization of objects by hue. *Child Development 60*, 752-762.

Choi, S., McDonough, L., Bowerman, M. & Mandler, J.M. (1999) Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development, 14,* 241-268.

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

Davidoff, J., Roberson, D., De Haan, M. & Davies, I.R.L. (2007) Colour categorization in infants. Report on ESRC grant R000223894. http://www.esrcsocietytoday.ac.uk/

Fagot, J., Goldstein, J., Davidoff, J. & Pickering, A. (2006) Cross-species differences in color categorization. *Psychonomic Bulletin & Review, 13*, 275-280.

Franklin, A. & Davies, I.R.L., (2004) New evidence for infant colour categories. *British Journal of Developmental Psychology, 22*, 349-377.

Franklin, A., Clifford, A. Williamson, E., Davies, I.R.L. (2005) Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental*

Child Psychology, 90, 114-141.

Garro, L.C. (1986) Language, memory and focality: A reexamination. *American Anthropologist, 88*, 128-136.

Gilbert, A.L., Regier, T., Kay, P. & Ivry, R.B. (2006) Whorf is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103*, 489-494.

Gordon, P. (2004) Numerical cognition without words. Evidence from Amazonia. *Science, 306*, 496-499.

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

Hardy, J.L., Frederick, C.M., Kay, P. & Werner, J.S. (2005) Color naming, lens aging and grue – What the optics of the aging eye can teach us about color language. *Psychological Science, 16*, 321-327.

Harnad, S. (1987) Psychophysical and cognitive aspects of categorical perception: A critical overview. In: S. Harnad (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press.

Heibeck, T.H., & Markman, E.M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*, 1021-1034.

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.

Hickerson, N.P. (1971) Review of Berlin & Kay (1969) *International Journal of American Linguistics, 37*, 257-270.

Jameson, K.A. (2005) Why GRUE? An interpoint-distance model analysis of composite color categories. Cross-cultural research, 39, 159-204.

Jameson, K.A. & Alvarado, N. (2003) Differences in Color Naming and Color Salience in Vietnamese and English. *COLOR Research & Application, 29*, 128-134.

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

Kay, P. & McDaniel, C.K. (1978) The linguistic significance of the meanings of basic colour terms. *Language* ***54*** (3), 610-646.

Kay P. & Regier, T. (2007) Color naming universals: the case of Berinmo. *Cognition.*

Kay, P., Regier, T. & Cook, R. S. (2005) Focal colors are universal after all. *Proceedings of the National Academy of Science, 102*, 8386-8391.

Komorova, N.L., Jameson, K.A. & Narens, L. (2007) Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology, 51*, 359-382.

Kuschel, R. & Monberg, T. (1974) “We don’t talk much about colour here”: A study of colour semantics on Bellona Island. *Man, 9*, 213-242.

Lantz, D. & Stefflre, V. (1964) Language and cognition revisited. *Journal of Abnormal and Social Psychology* ***69*****,** 472-481.

Levinson, S.C. (2001) Yélî Dnye and the theory of basic color terms. *Journal of Linguistic Anthropology, 10*, 3-55.

Lindsey, D.T. & Brown, A.M. (2002) Color naming and the photoxic effects of sunlight on the eye. *Psychological Science, 13,* 506-512.

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

Lucy, J.A. (1997) The linguistics of 'color'. In: *Color categories in thought and language*. C.L. Hardin & L. Maffi (eds), 320-346. New York: Cambridge University Press.

Lucy, J.A. & Shweder, R.A. (1979) Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist, 81*, 581-605.

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

MacLaury, R.E., (1997) *Color and cognition in mesoAmerica: Construing categories as vantages. A*ustin, Tx: University of Texas Press.

Majid, A., Bowerman, M., Kita, S., Haun, D.B.M. & Levinson, S.C. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Science, 8*, 108-114. Mahwah, NJ: Lawrence Erlbaum Associates.

Malt, B.C., Sloman, S.A. & Gennari, S.P. (2003) Universality and language specificity in object naming. *Journal of Memory and Language, 49*, 20-42.

Matsuzawa, T. (1985) Color naming and classification in a chimpanzee (Pan-Troglodytes) *Journal of Human Evolution, 14*, 283-291.

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.

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

MA: MIT Press.)

Nisbett, R.E., Peng, K.P., Choi, I. & Norenzayan, A. (2001) Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*, 291-310.

O'Hanlon, C.G. & Roberson, D. (2006) Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology, 94*, 275-300.

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

Ratner, C. (1989) A sociohistorical critique of naturalistic theories of color perception. *Journal of Mind and Behavior 10*, 361-373.

Regier, T., Kay, P. & Khetarpal, N. (2007) Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences, 104*, 1436-1441.

Roberson, D., Davidoff, J. & Shapiro, L. (2002) Squaring the circle: The cultural relativity of good shape. *Journal of Cognition and Culture, 2*, 29-52.

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

Roberson, D., Davidoff, J., Davies, I.R.L. & Shapiro, L. R. (2004) The Development of Color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General, 133*, 554-571.

Roberson, D., Davidoff, J., Davies, I.R.L. & Shapiro, L. (2005) Colour categories in Himba: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50,* 378-411.

Roberson, D., Pak, H. and Hanley, J.R. (2008) Categorical perception of colour in the left and right hemisphere is verbally mediated: Evidence from Korean. *Cognition, 107*, 752-762.

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

Rosch, E.H. (1972b) Probabilities, sampling and ethnographic method: The case of Dani colour names. *Man, 7*, 448-466.

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

Rosch, E. (1975) The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance* ***1*** (4), 303-322.

Sandell, J.H., Gross, C.G. & Bornstein, M.H. (1979) Color Categories in Macaques. *Journal of Comparative and Physiological Psychology, 93,* 626-635.

Sandhofer, C.M., & Smith, L.B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology, 35*, 668-679.

Sandhofer, C.M. & Smith, L.B. (2001) Why children learn color and size words so differently: Evidence from adults' learning of artificial terms. *Journal of Experimental Psychology: General, 130,* 600-620.

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

Saunders, B.A.C. & van Brakel, J. (2001) Rewriting color. *Philosophy of the Social Sciences, 31*, 538-556.

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.

Shepard, R.N. (1992) The perceptual organisation of colors: An adaptation to regularities of the terrestrial world? In: J.H. Barkow, L. Cosimedes & J. Tooby (eds.) *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Smith, L.B., Jones, S.S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *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.

Steels, L. & Belpaeme, T. (2005) Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28*, 469-489.

Stefflre, V., Castillo Vales, V. & Morley, L. (1966) Language and cognition in Yucatan: A cross-cultural replication. J*ournal of Personality and Social Psychology, 4*, 112-115.

Sumner, P & Mollon, J.D. (2000) Catarrhine photopigments are optimized for detecting targets against a foliage background. *Journal of Experimental Biology, 2003*, 1963-1986.

Tan, L.H., Chan, A.H.D., Kay, P., Khong, P-L., Yip, L.K.C. & Luke, K-K. (2008) Language affects patterns of brain activation associated with perceptual decision.

Proceedings of the National Academy of Sciences, 105, 4004-4009.

Werker, J.F. & Tees, R.C. (1984) Cross-language speech-perception – Evidence for perceptual reorganization during the 1st year of life. *Infant Behavior & Development, 7*, 49-63.

Whorf, B.L. (1956) The relation of habitual thought and behavior to language. In *Language, Thought and Reality: Essays by B.L. Whorf*. J.B. Carroll (ed.) Cambridge: M.I.T. Press.

Wierzbicka, A. (1999) "Universals of colour" from a linguistic point of view. *Behavioral & Brain Sciences, 22*, 724-725.

Wierzbicka, A. (2005) In defense of 'culture'. *Theory & Psychology, 15*, 575-597.

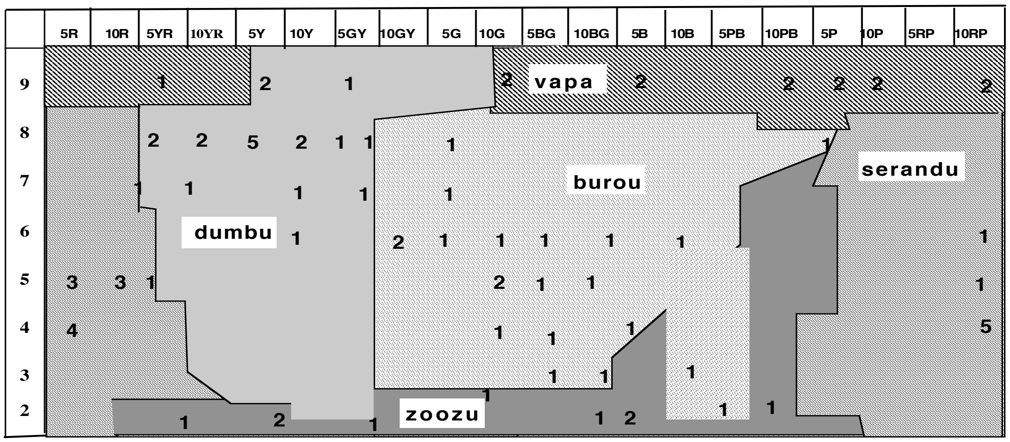
Winawer, J., Witthoft, N., Frank, M.C., Wu, L. and Boroditsky, L. (2007) Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*, 7780-7785.

Yendrikhovskij, S.N. (2001) Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology, 45*, 409-417.

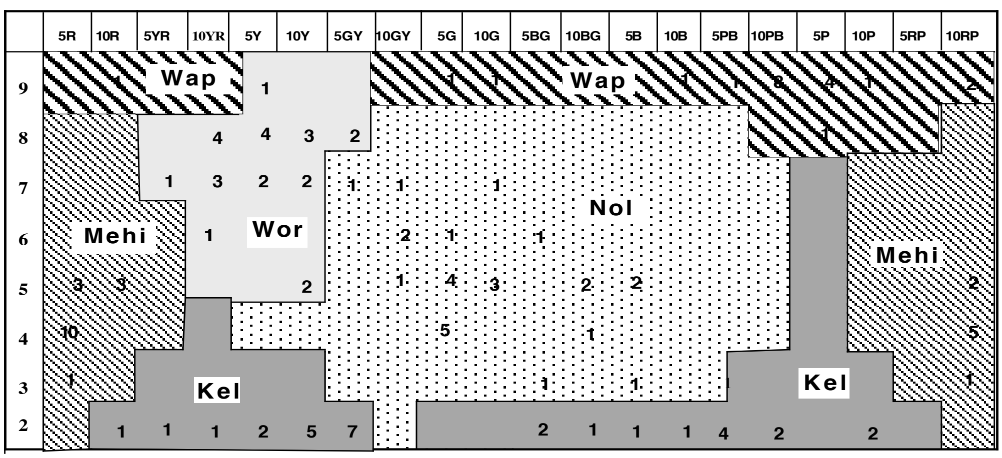
Figure 1. Distribution of Himba named categories and choices of best exemplar for the 160 chip saturated array (for 31 observers) compared to those of English and Berinmo speakers for the same array. Numbers represent number of individuals choosing an exemplar as best example of the category. Dots on English graph represent the locations of best examples for English speakers.

Figure 1

Himba naming distribution



Berinmo naming distribution



English naming distribution