Color naming universals: the case of Berinmo

Paul Kay
International Computer Science Institute
University of California, Berkeley
paulkay@berkeley.edu

Terry Regier
University of Chicago
regier@uchicago.edu

Word count: 2985.

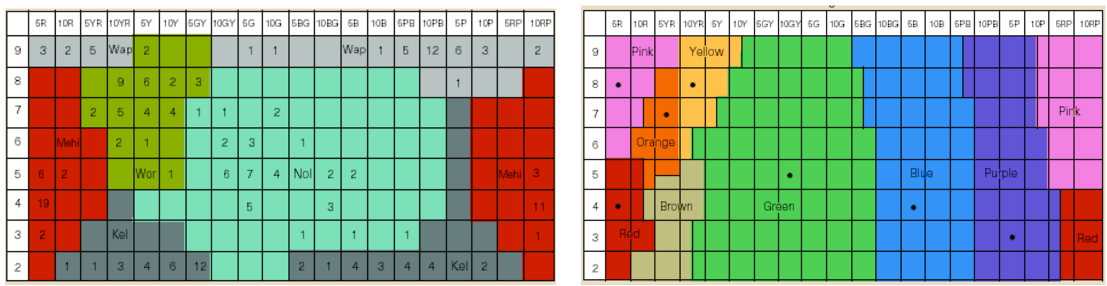
Abstract

Proponents of a self-identified 'relativist' view of cross-language color naming have confounded two questions: (1) Is color naming largely subject to local linguistic convention? and (2) Are cross-language color naming differences reflected in comparable differences in color cognition by their speakers? The 'relativist' position holds that the correct answer to both questions is Yes, based on data from the Berinmo language of Papua New Guinea. It is shown here that the Berinmo facts instead support a more complex view – that cross-language color naming follows non-trivial universal tendencies, while cross-language color-naming differences do indeed correlate with differences in color cognition. The rhetoric of 'relativity' versus 'universalism' impedes understanding of cross-language color naming and cognition.

The 'linguistic relativity' versus 'linguistic universals' debate in the color domain has revolved around two distinct questions, often insufficiently distinguished. (1) Do the languages of the world lexically carve up the color space largely arbitrarily? (2) Where color-naming differences among languages occur, do they correlate with corresponding differences in memory, learning and discrimination of colors? A committed relativist wants the answers to be Yes and Yes; a committed universalist wants the answers to be No and No. In our view, currently available evidence points strongly toward the answers No and Yes, providing aid and comfort to neither extreme position. There are non-trivial universal tendencies in cross-language color naming (Berlin and Kay 1969, Kay and McDaniel 1978, Boynton and Olson 1987, Uchikawa and Boynton 1987, MacLaury 1997, Kay and Maffi 1999, Lindsey and Brown 2002, Kay and Regier 2003, Regier and Kay 2004, Regier, Kay and Cook 2005) but at the same time color-naming differences occur and do correlate with color memory, learning and discrimination (Kay and Kempton 1984,Uchikawa and Shinoda 1996, Roberson and Davidoff 2000, Roberson, Davies and Davidoff 2000, Özgen and Davies 2002, Witthoft et al. 2003, Roberson et al. 2004). It appears that both the universalist and relativist dogmas obscure an interestingly complex situation.

The Berinmo language (Sepik-Ramu family, Papua New Guinea)[1] has been involved in the confusion just noted. Berinmo has a color naming system that clearly differs from that of English (Roberson et al. 2000), as can be seen in Figure 1.[2]

[1] Berinmo is spoken in the Bitara and Kagiru villages (possibly amongst others) located near the Sepik River in northeast Papua New Guinea. (Roberson p.c.). The Ethnologue identifies the language spoken in these villages as Berinomo. We have retained Roberson's spelling but wished to point to the identical reference of "Berinmo" and "Berinomo."



Berinmo English

Fig. 1. Color categories in Berinmo and English (Roberson et al., 2000).

With regard to question (2), Debi Roberson and colleagues have shown that the differences between Berinmo and English color category boundaries correlate with, and presumably cause, differences in memory, learning and discrimination of colors in speakers of the two languages (Davidoff and Roberson 1999, Roberson, Davies and Davidoff 2000, Roberson, et al. 2004). They also failed to replicate Rosch's well-known finding of better color memory for proposed universal focal colors (Heider 1972, Heider and Olivier 1972). These focal colors have been understood to be the cognitive underpinning for cross-language naming universals; to challenge their existence or effectiveness is implicitly to challenge one of the bases for universals of color naming. Perhaps for this reason, Roberson and associates take their results to be relevant to question (1) as well, proposing an uncompromisingly anti-universalist account of color naming. Although they do occasionally acknowledge that the Berinmo color naming system is similar to that of some other languages (e.g., Roberson et al. 2005: 402), the broad conclusions they draw have quite the opposite flavor:

> [C]olor categories [are] a function of cultural experience and only, at most, loosely constrained by the default neural organization (Roberson, Davies and Davidoff 2000)

---

[2] The Berinmo data were gathered using a reduced, 160-chip, version of the 320 chromatic chip Munsell palette used by Lenneberg and Roberts (1956), Berlin and Kay (1969), The World Color Survey [WCS] (Kay and Regier 2003, Regier and Kay 2005), and the Mesoamerican Color Survey (MacLaury 1997). The reduced version was originally used by Eleanor Rosch in her study of the Dani (Trans-New Guinea family, Papua New Guinea, Heider 1972, Heider and Olivier 1972), with which the Berinmo data were compared in detail (Davidoff and Roberson 1999, Roberson, Davies and Davidoff 2000). The 160-chip array was created from the 320-chip array by removing every other (hue) column. In the comparisons we make below between the Berinmo data and WCS data, we use the 320 chip format, treating each "missing" Berinmo hue column as if every chip in that column was named like the chip to its right. Diagrams like those in Figures 1 and 2 (among others) represent for each stimulus chip in the palette the name most frequently given to that chip. We refer to such diagrams as "mode maps". A reasonable approximation to the colors seen in the full palette may be seen at http://www.icsi.berkeley.edu/wcs/study.html.

> [W]e propose that color *categories are formed from boundary demarcation based predominantly on language*. Thus, in a substantial way we present evidence in favor of *linguistic relativity* (Roberson, Davies and Davidoff 2000. Italics added)

They explicitly mention only one constraint on color naming across languages, 'grouping by similarity', an idea they emphasize repeatedly:

> The most important [non-linguistic] constraint [on color terminologies] would be that similar items (as defined by perceptual discrimination) are universally grouped together. Thus, no language would exhibit categories that include two areas of color space but excludes [sic] an area between them. (Roberson, Davies and Davidoff 2000)
>
> Whilst Berinmo speakers, like those of all other languages hitherto investigated, appear to group contiguous areas of the colour space together, no evidence was found for these sections to correspond to a limited set of universal basic color categories (Roberson et al. 2002)
>
> No language has ever been reported to have a category that includes two areas of color space (e.g. yellow and blue) but excludes an area between them (green). There is no associative chain of similarity that could connect yellow to blue without passing through green. Grouping always follows principles of similarity (as defined by perceptual discrimination) and *the only free parameter appears to be the placement of boundaries between categories* (Roberson 2005, Italics added).

By implication, so long as similar colors are grouped together, anything goes in cross-language color naming. In particular, on this view, the actual *location* of categories in color space is apparently not constrained: the placement of boundaries is considered a 'free parameter' under control of local linguistic convention. Here, we test this view of color naming against the language that suggested it in the first place – Berinmo. We hope to show, contra Roberson and associates, that Berinmo color naming fits a pattern that is both narrowly specified and widely distributed. While linguistic category boundaries do affect color discrimination and memory in speakers of Berinmo and other languages, the placement of those boundaries is constrained by universal forces.

Grouping by similarity

If color categories are constrained primarily by the 'grouping by similarity' principle, and boundaries are demarcated by local linguistic convention, there should be nothing privileged about the locations of color category boundaries in Berinmo – the boundaries could just as easily have been drawn elsewhere. To pursue this idea, we considered the actual Berinmo data, and 19 hypothetical variants of it, obtained by rotating the original data by 2, 4, 6, etc. Munsell hue columns (the 320 chip Munsell palette contains 40 columns, nominally of psychologically equal hue steps). Figure 2 illustrates unrotated Berinmo, and two of the rotated hypothetical variants.[3]

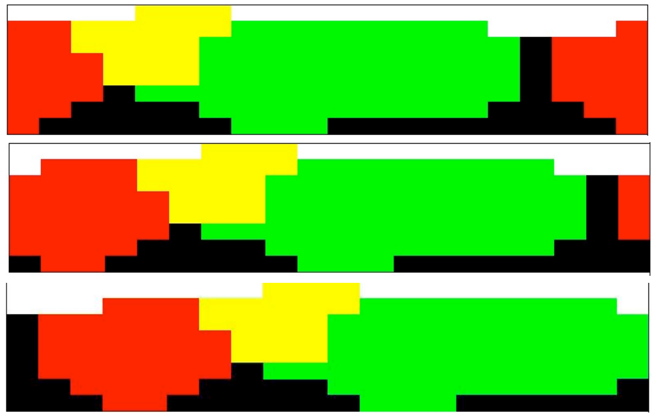


Figure 2. Berinmo (a) unrotated, and rotated (b) four and (c) eight columns.

This yields a set of hypothetical Berinmo-based languages that all obey the ‘grouping by similarity’ principle equally well, that also all maintain the shape of the categories and their positions relative to each other, and that vary only in *where* in color space these categories are located. If the locations of Berinmo boundaries are constrained by universal forces, then boundaries in other (real) languages should align more closely with boundaries in the unrotated (actual) version of Berinmo than with those in the hypothetical rotated versions. However, if ‘grouping by similarity’ is the only substantial constraining force in color naming, we would not expect the unrotated version of Berinmo to be privileged in this manner.

We compared color category boundaries in the real and hypothetical versions of Berinmo with those in the 110 languages of the World Color Survey (Cook, Kay & Regier in press). To do this, we first constructed, for Berinmo and all rotated variants, and for each of the WCS languages, a ‘boundary map’ indicating where category boundaries fall in that language. This was done by scoring each chip in the mode map as a ‘boundary chip’ if any of its four neighbors (directly above, below, to the left, or to the right) was given a different name.[4] Figure 3 illustrates how boundaries are compared across two languages: first the mode map for each language is converted to a boundary map; here, boundary chips are shown in white and non-boundary chips in black. Then

---

[3] Roberson et al.’s (2000) figures of the Berinmo data appear less elongated than ours because they report data for only every other column of the stimulus array, while we display all 40 columns, filling in “missing” columns as described above.
[4] For purposes of determining neighbors, all chips in row B were taken to be directly below chip A0; analogously, all chips in row I were taken to to be directly above chip J0. No other chromatic chips were taken to be neighbors of any of the neutral (non-colored) chips. Roberson et al. (2000) do not report Berinmo names for the neutral chips in the array, which are not displayed here; however, for comparison with other languages which do have values for these chips, we have assumed that the Berinmo “black” term (*kel*) would extend from J0 through F0, and the “white” term (*wap*) from A0 down to E0. This assumption is a matter of convenience and not critical to the test we pursue here.

the two boundary maps are compared by simply counting the proportion of chips with the same scoring (black or white). We refer to the degree of alignment as the 'boundary match' between the two languages.

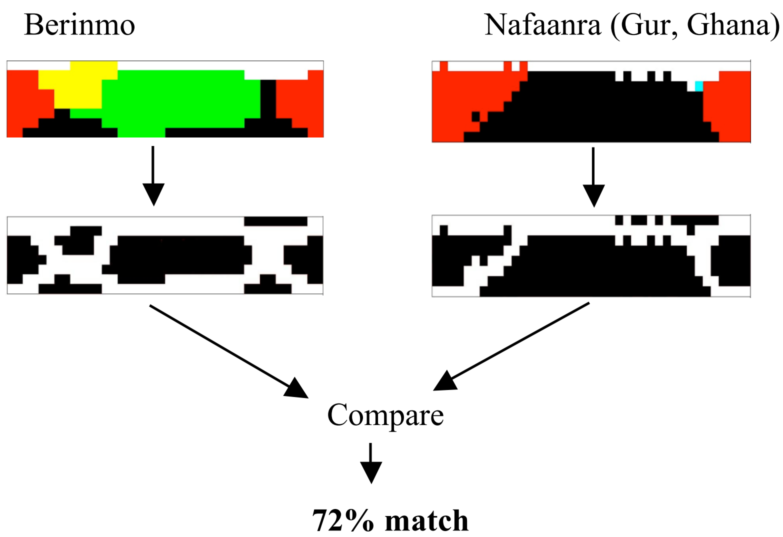


Figure 3. Comparison of Berinmo and Nafaanra boundaries.

For Berinmo and each of its rotated variants, we calculated the boundary match to each WCS language and averaged these matches across all WCS languages, yielding a measure, for each variant of Berinmo, of how well that version of Berinmo matches boundaries in the WCS overall. The results are shown in Figure 4.

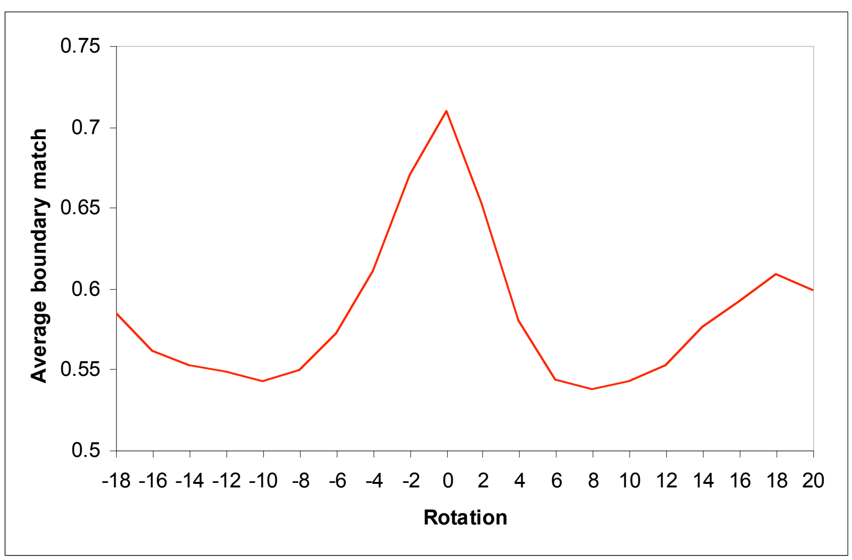


Figure 4. Average boundary match with WCS of various rotated versions of Berinmo, including no rotation.

Actual Berinmo boundaries (zero rotation) are more typical of WCS boundaries in general than are those of any hypothetical five-term language with the same shape as Berinmo but rotated to any degree in the Munsell hue plane. Moreover, in the neighborhood of true Berinmo, the greater the degree of rotation the less correspondence with WCS boundaries.[5] This suggests that the locations of Berinmo category boundaries reflect universal constraints stronger than 'grouping by similarity'.

Taking a closer look

The above demonstration shows that Berinmo is more similar to other languages when unrotated than when rotated – suggesting that Berinmo obeys universal constraints. But just how great is this similarity? Are there languages with boundaries quite similar to Berinmo? Or is Berinmo in fact quite dissimilar from other languages, despite the fact that rotated Berinmo is even more dissimilar? Let us look first at Roberson and associates' comparative data. Roberson and colleagues have studied color naming and cognition in two languages with five basic color terms: Berinmo and Himba, the latter identified by Roberson et al. (2004) as a historically isolated dialect of Herero, a Bantu language of Namibia. '… the Himba are semi-nomadic tribesmen inhabiting an arid region: their visual diet of open desert, scrubland and mountain is radically different to that of Berinmo speakers' deeply shaded and lush forest territory' (Roberson et al. 2004). Yet, if we look at the color naming systems of Berinmo and Himba, as recorded by Roberson and associates, we see they are quite similar. Figure 5 is reproduced directly from Roberson et al. (2004).

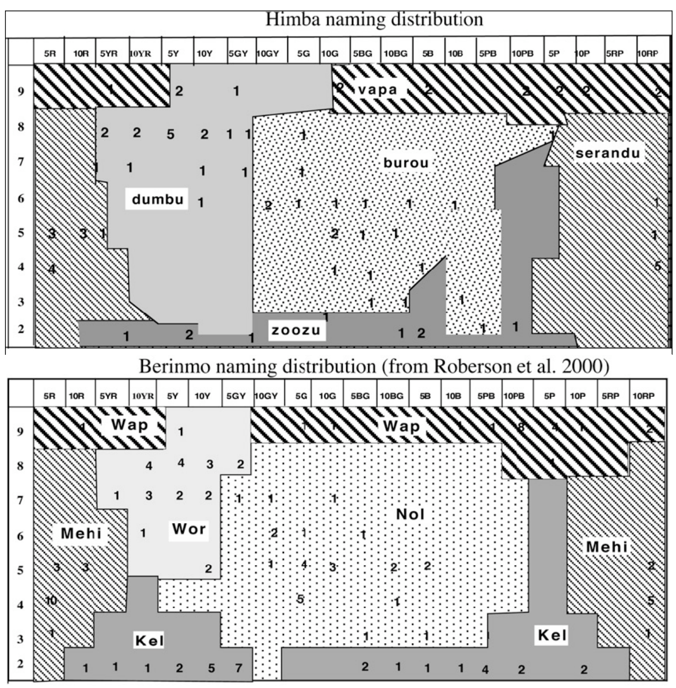


Fig. 5. Himba (above) and Berinmo (below) mode maps (source: Roberson et al. 2004).

[5] The secondary maximum near 180 degree rotation is apparently caused by the fact that such a rotation brings the boundary of Berinmo grue into rough coincidence with parts of the boundaries of general WCS red and yellow and vice versa. In effect, 180 degree rotation causes Berinmo cool (grue) and WCS average warm boundaries (red or yellow) to roughly coincide.

Each language has a black term, *zoozu, kel*, which extends well into purples; a white term, *vapa, wap*, which extends over all the lightest hues (Munsell Value 9) except those named by the yellow term; a red term, *serandu, mehi*, that extends to pink and light purples; an extended yellow term, *dumbu, wor*, that also covers orange, yellowish greens, light browns and olives; and a grue term, *burou, nol*, that lacks some yellowish greens and extends somewhat into purple. The major area of disagreement in these plots of modal naming judgments is the brown area, which is mostly covered by downward extension of the yellow term in Himba and upward extension of the black term in Berinmo. This area is singled out by Roberson et al. (2004) as one of the two with the lowest consensus in Himba. 'A few areas have very low agreement on naming… one corresponds roughly to English *brown*, the other to English *purple*.'

To extend the comparison, it may be instructive to compare Berinmo visually to some other five-term systems, these from the WCS. Figure 6 presents a visual comparison of modal responses on the color-naming task of Berinmo with eight languages from the WCS.

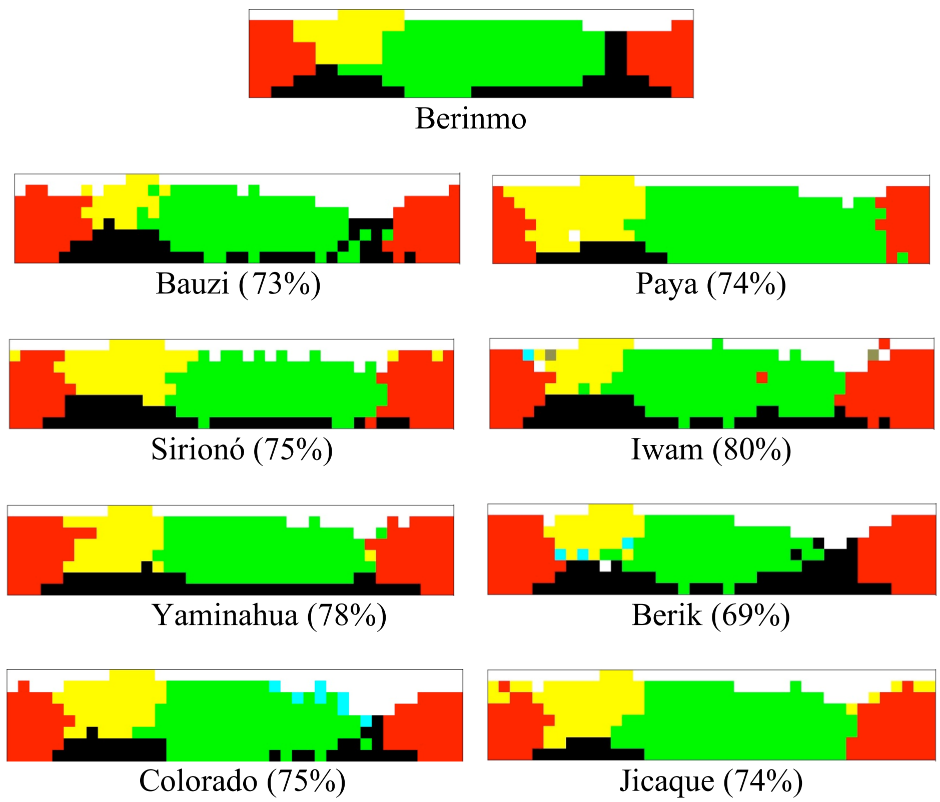


Figure 6. Modal naming responses of Berinmo speakers and those of eight languages from the WCS data, each with five basic color terms. The percentage following each WCS language name is the boundary match with Berinmo. Raw data on which these mode maps are based is available at http://www.icsi.berkeley.edu/wcs/data.html

The eight comparison languages shown in Figure 6 represent seven different language families and one language isolate: Bauzi, Geelvink Bay family, Irian Jaya; Paya, Chibchan family, Honduras; Sirionó, Tupi family, Bolivia; Iwam, Sepik-Ramu family, Papua New Guinea; Yaminahua, Panoan family, Peru; Berik, Trans New Guinea family, Irian Jaya; Colorado, Paezan family, Ecuador; Jicaque, Isolate, Honduras. Qualitatively, Berinmo color naming appears to be quite similar to that of other five-term languages from a range of genetically and geographically separated language families, all of which show clear similarities to each other.

Conclusion

As we have noted above (and previously, Regier, Kay and Cook 2005) there is ample evidence that differences in color category boundaries between languages may influence color memory, learning or discrimination (Heider 1972, Heider and Olivier 1972, Kay and Kempton 1984,Uchikawa and Shinoda 1996, Roberson and Davidoff 2000, Roberson, Davies and Davidoff 2000, Özgen and Davies 2002, Witthoft et al. 2003, Roberson et al. 2004). These results have for the most part been established by comparing a behavioral color response between speakers of English and one of a handful of languages, all differing from English in the placement of some lexical color category boundary. The general pattern of these studies is that the nominally non-linguistic behavioral response function exhibits an inflection point or discontinuity at the speakers' lexical boundary, distinguishing the two subject-groups' non-linguistic behavior in parallel with the lexical difference.

Berinmo has been perhaps the most intensively studied of this small group of languages. Independently of this fact, Roberson et al. (2000) were unable to replicate in Berinmo Rosch's Dani results regarding the apparent salience of the focal points of English color terms with respect to memorability in a language with a different color lexicon (Heider 1972, Heider and Olivier 1972). Since Rosch had made universal focal colors a cornerstone of her explanation of universal color naming, Roberson and her associates were apparently led to conclude from their inability to replicate Rosch's Dani experiments with the Berinmo that (1) the defining features of color categories are boundaries rather than foci and (2) there are no universal constraints on color term boundaries, other than grouping similar colors together.

We do not agree that these conclusions follow from the observations on which they are based. This study has been devoted to assessing the empirical status of the second conclusion. To this end, we have accepted for the purposes of comparison the equation of color categories with their boundaries. Using this criterion, we have compared Berinmo color categories to the 110 WCS languages as a whole, to eight selected five-term WCS languages and to Himba, the other five-term language studied by Roberson and associates. In each case, Berinmo color category boundaries appear to be typical of the comparison class. Furthermore, best example choices in Berinmo similarly appear to align closely with those of WCS languages (Regier, Kay, & Cook, 2005). There is no evidence in Berinmo color naming to challenge the findings of universal constraints on color naming.

More broadly, we argue that the separate questions of (1) the existence of universal constraints on color naming and (2) the influence of color-naming differences on differences in color cognition should not be confounded under a rhetoric of 'relativism' versus 'universalism.' Current evidence supports *both* the existence of universal constraints on color naming *and* the influence of color-naming difference on color memory and discrimination.

References

Berlin, Brent & Paul Kay (1969) *Basic Color Terms: Their universality and evolution*. Berkeley and Los Angeles: University of California Press.

Boynton, Robert M. & C.X. Olson (1987) Locating basic colors in the OSA space. *Color Research and Application* 12, 94-105.

Cook, Richard S., Paul Kay & Terry Regier (in press) The World Color Survey Database: History and Use. Cohen, Henri and Claire Lefebvre (eds.) *Handbook of Categorisation in the Cognitive Sciences*. Amsterdam and London: Elsevier.

Davidoff, Jules, I.R.L. Davies & Debi Roberson (1999) Is color catgorisation universal? New evidence from a stone-age culture. *Nature* 398, 203-204.

Heider, Eleanor A. (1972) Probabilities, sampling and ethnographic method: The case of Dani colour names. *Man* 7, 448-466.

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

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

Kay, Paul & Chad K. McDaniel (1978) The linguistic significance of the meanings of basic color terms. *Language* 54, 610-646.

Kay, Paul & Luisa Maffi (1999) Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist* 101, 743-760.

Kay, Paul & Terry Regier (2003) Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences* 100, 9085-9089.

MacLaury, Robert E. (1997) *Color and cognition in Mesoamerica : constructing categories as vantage*s. Austin: University of Texas Press.

Lenneberg, Eric H. & John M. Roberts (1956) The Language of Experience: A study in methodology. *International Journal of American Linguistics*. Memoir 13.

Lindsey, D.T. & Angela M. Brown. (2002) Color naming nd the phototoxic effects of sunlight on the eye. *Psychological Sciences* 13, 506-512.

Özgen, Emre & Ian R. L. Davies (2002) Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General* 131, 477-493.

Regier, Kay & Cook (2005) Focal colors are universal after all. *Proceedings of the National Academy of Sciences* 102:8386-8391

Regier, Terry & Paul Kay (2004) Color naming and sunlight: Commentary on Lindsey and Brown (2002). *Psychological Science* 15(4), 289-290.

Roberson, Debi (2005) Color categories are culturally diverse in cognition as well as in language. *Cross-Cultural Research* 39, 56-71.

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

Roberson, Debi, Jules Davidoff, Ian R.L. Davies & Laura R. Shapiro (2004) Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology* 50, 378-411.

Roberson, Debi., I. R.L. Davies & Jules Davidoff (2002) Colour categories are not universal: Replications & new evidence. In B. Saunders & J. van Brakel (Eds.) *Theories, technologies, instrumentalities of color: Anthropological & historical perspectives*. Lanham, Maryland: University Press of America, Inc.

Roberson, Debi, Ian R.L. Davies & Jules Davidoff (2000) Color categories are not universal: Replications & new evidence from a Stone-age culture. *Journal of Experimental Psychology: General*, 129, 369-398.

Uchikawa K. & Robert M. Boynton (1987) Categorical color perception of Japanese observers: comparison with that of Americans. *Vision Research* 27, 1825-1833.

Uchikawa, K& H. Shinoda (1996) Influence of basic color categories on color memory discrimination. *Color Research and Application* 21,430-439.

Witthoft, Nathan, Jonathan Winawer, Lisa Wu, Michael Frank, Alex Wade and Lera Boroditsky (2003) Effect of language on color discriminability. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society,* R. Alterman, D. Kirsh, (Eds.) Mahwah, NJ: Lawrence Erlbaum, pp. 1247-1252.