

Language, thought, and color: Whorf was half right

Terry Regier^{1*} and Paul Kay^{2*}

¹ Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, IL 60637, USA

² Department of Linguistics, University of California, Berkeley, CA 94720 and International Computer Science Institute, 1947 Center Street, Berkeley, CA 94704, USA

The Whorf hypothesis holds that we view the world filtered through the semantic categories of our native language. Over the years, consensus has oscillated between embrace and dismissal of this hypothesis. Here, we review recent findings on the naming and perception of color, and argue that in this semantic domain the Whorf hypothesis is half right, in two different ways: (1) language influences color perception primarily in half the visual field, and (2) color naming across languages is shaped by both universal and language-specific forces. To the extent that these findings generalize to other semantic domains they suggest a possible resolution of the debate over the Whorf hypothesis.

The state of the debate

A classic debate in cognitive science concerns the relation between language and perception. At one pole of this debate is the relativist stance, which holds that our perception of the world is shaped by the semantic categories of our native language, and that these categories vary across languages with little constraint - a view often associated with Benjamin Lee Whorf. At the other pole is the universalist stance, which holds instead that there is a universal repertoire of thought and perception that leaves its imprint on the languages of the world. Over the years, consensus has swung back and forth between these two poles.

The domain of color has furnished an empirical *locus classicus* of the debate for the last half-century. In a 2006 review [1] we suggested that the universalist-versus-relativist debate with respect to color often conflates 2 distinct questions:

1. Do color terms affect color perception?
2. Are color categories determined by largely arbitrary linguistic convention?

A relativist would respond ‘yes’ to both questions, whereas a universalist would respond ‘no’ to both. However, that review argued that empirically the answers to the two questions do not match. Specifically, it was argued on the basis of evidence then available that color terms do affect color perception (yes to question 1), but that there are also universal tendencies in color naming (no to question 2). It was argued that the universalist/relativist opposition

is unhelpful as a conceptual structuring device, since it does not accommodate these realities.

In the years since then, new findings have arisen that suggest a subtler view. The new evidence suggests that Whorf was partly right with respect to each of these two questions. With respect to question 1, color names do influence color perception - but primarily in the right visual field, and less so in the left. With respect to question 2, color naming across languages does reflect universal tendencies, as shown in earlier work - but also some degree of local linguistic convention. These findings suggest a way in which the recently re-opened debate over language and thought in the color domain might be resolved. And to the extent that these findings generalize to other semantic domains, they suggest a possible resolution of the Whorf debate more broadly.

Language affects perception - in half the visual field

Does language affect perception? As noted above, several studies suggest that the answer is ‘yes’, at least in connection with color. These studies have shown that there is ‘categorical perception’ (CP: faster or more accurate discrimination of stimuli that straddle a category boundary) for color, and that differences in color category boundaries between languages predict where CP will occur [2–7]. Moreover, several of these studies, and others [e.g. 8] have shown that color CP disappears with a concurrent verbal interference task, confirming that color CP is language based.

However, this straightforward Whorfian answer - and the yes-or-no framing of the very question ‘does language affect perception?’ - obscure an interesting possibility: that language might affect *half* of perception. Specifically, language might be expected to shape perception primarily in the right visual field (RVF), and much less if at all in the left visual field (LVF). This expectation follows from the observations that the left hemisphere (LH) of the brain is dominant for language, and that the visual fields project contralaterally to the brain. On this view, half of our perceptual world might be viewed through the lens of our native language, and half viewed without such a linguistic filter.

This proposal was first advanced and tested in a study by Gilbert, Regier, Kay & Ivry [8] that probed the perceptual discrimination of colors straddling the boundary between green and blue, a boundary present in English but absent in many other languages. In this study, American English speaking participants first saw a central fixation

Corresponding authors: Regier, T. (terry.regier@gmail.com); Kay, P. (paulkay@berkeley.edu)

* The two authors contributed equally to this work.

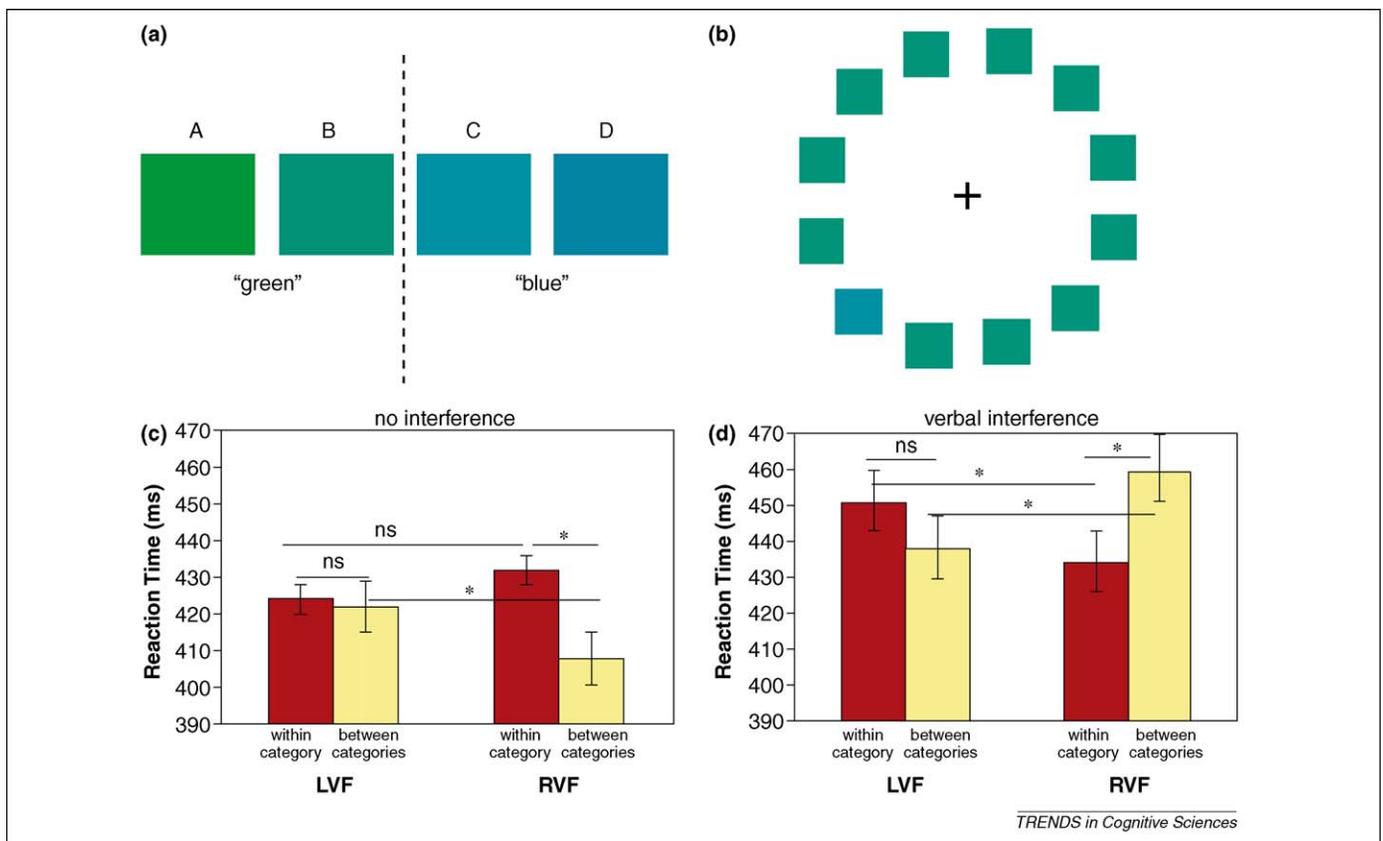


Figure 1. Lexical categories influence perception in the RVF. (a) Print-rendered versions of the four colors used. (b) Sample display for the visual search task. Participants were required to press one of two response keys, indicating the side containing the target color. (c) In the no-interference condition, category effect in the RVF only. (d) Effect reversed with verbal interference. *, $P < 0.05$, two-tailed t test; $df=10$; ns, nonsignificant. Values are mean \pm SEM. Source: [8].

cross, and then a ring of colored squares arrayed around it (Figure 1b). All of the squares were of the same color except for one (the target), and participants were required to identify whether the target appeared on the left or right half of the display by pressing a button with the corresponding hand. Critically, the target color had either the same name as the color of the other squares (e.g. a green against a background of a different green, as in stimulus pair A,B in Figure 1a), or a different name (e.g. a green against a background of a blue, as in stimulus pair B,C in Figure 1a).

One of Gilbert et al.'s experiments is depicted in Figure 1. As shown in Figure 1c, cross-category targets were identified faster than same-category targets in the RVF only. When a concurrent task requiring verbal resources (remembering an eight-digit number) was added, the RVF color CP effect disappeared, in fact was reversed, reinforcing the interpretation that when RVF color CP is found, it has a verbal basis [4,5]. In a second experiment (not shown in Figure 1) Gilbert *et al.* replicated these findings, and also found that RVF CP was not disrupted by a non-verbal interference task of difficulty comparable to the verbal interference task - again strengthening the conclusion that the lateralization pattern is linguistic in origin. Drivonikou *et al.* [9] reanalyzed by visual field color search data that had been collected earlier for a different study and found predominantly RVF color CP; they ran a new experiment similar to that of Gilbert et al. and found color CP in the RVF superior to that in the LVF. However their finding of significant,

although weaker, color CP in the LVF diverged from the Gilbert et al. result; the longer response time in the Drivonikou data suggest that this apparent RH CP might be the result of longer RTs permitting cross-callosal transfer and/or scanning (for further discussion of this possibility in these and other studies see [7,10,11]).

Gilbert et al. [8] also reported that a callosotomy ("split-brain") patient showed RVF color CP with no trace of LVF color CP. A separate study [12] reported a similar result with a second callosotomy patient in an experiment in the same paradigm as [8] but using non-color stimuli, namely silhouettes of dogs and cats. When this experiment was run with normal adult participants, CP was found in both visual fields but again more strongly in the RVF than LVF, and again this pattern was disrupted by a verbal but not a non-verbal interference task. These results confirm that linguistic CP is localized to the LH and suggest that the phenomenon is not limited to color.

Roberson, Pak & Hanley [7] examined a color boundary in Korean that does not correspond to any English color boundary in a setup very similar to that of [8]. Their eight fastest subjects showed color CP in the RVF only, whereas their eight slowest subjects showed color CP in both visual fields. Roberson et al. infer that the slower subjects' LVF CP probably reflects cross-callosal transfer. The literature so far shows a general tendency for color CP to be restricted to the RVF for short RTs, to appear in the LVF as well but to a lesser degree in medium length RTs, and to show up more or less equally in both VFs with very long RTs,

reinforcing the conclusion that LVF CP in adults reflects cross-callosal transfer, and perhaps scanning in some cases. Roberson, Pak & Hanley [7] is the only published study at this time to have tested lateralized color CP at a color category boundary other than those found in English.

So far, everything we have considered points to language-driven CP originating in the LH manifesting itself in the RVF. However, this picture is complicated by evidence of prelinguistic - and thus non-linguistic - CP in infants [e.g. 13,14,15] and the finding that color term knowledge does not affect color CP in toddlers [16]. The existence of prelinguistic CP has recently been both contested [17,18] and defended [19]. If confirmed, it raises an important question: What relation does prelinguistic CP bear to the language-driven RVF-lateralized color CP universally observed in speaking adults? Prelinguistic color CP has been observed along boundaries corresponding to some of those in familiar European (and many other) languages. Thus, a natural possibility is that prelinguistic categories serve as a starting-point for the later elaboration of linguistic categories.

Franklin et al. [20] compared infant and adult performance on a visual search task much like that of [8]. Since infants cannot reliably respond through a button press, a different response measure was used: the time it took to initiate an eye movement from central fixation to a target dot, displayed in one color against a uniform background of a different color. As in previous experiments the target and background were either of the same category (e.g. two blues) or of different categories (one green, the other blue). This method was used with adults (with less easily discriminable colors) and showed the expected RVF-dominant CP of color. In contrast, prelinguistic infants showed no CP in the RVF, and clear CP in the LVF.

Is the migration of the category effect - from the RH/LVF in infants to the LH/RVF in adults - caused by the acquisition of color terms? To pursue this question more precisely, Franklin et al. [10] tested toddlers aged two to five with the same procedure. The children's grasp of the relevant color terms *blue* and *green* was assessed and the toddler subjects sorted readily into two groups: "learners", who had not yet acquired these color terms, and "namers", who had. In the visual search task the learners patterned like infants: they displayed color CP in the LVF but not the RVF; whereas the namers patterned like adults: CP in the RVF but not the LVF. These results remain when age is added as a covariate to the analysis, as shown in Figure 2.

This study suggests strongly that it is acquisition of color terms *per se* that causes the shift of color CP from RH to LH, a conclusion that raises several questions. What is the nature of this shift? Are the RH prelinguistic categories transported to the LH, where they serve as a starting-point for the elaboration of LH linguistic categories? Or are LH linguistic categories constructed *de novo* by language, without reference to the RH prelinguistic categories? In either case, why is there no trace left of the RH prelinguistic categories once color terms have been learned - are they permanently erased [21], or merely suppressed on-line by language?

We can presently answer only one of these questions, and only tentatively. It seems likely that RH prelinguistic

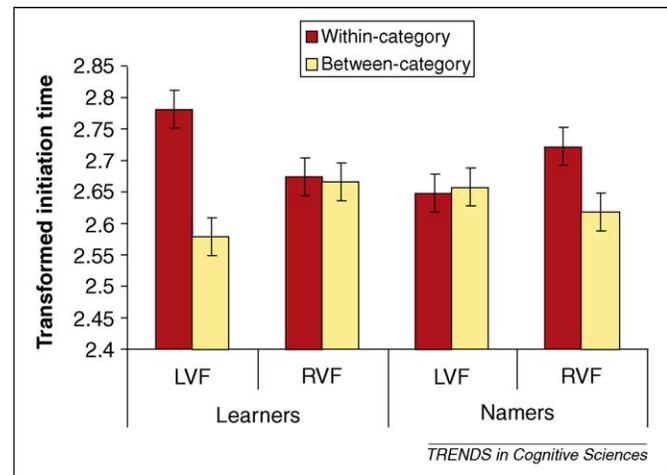


Figure 2. Category effect in the LVF for toddlers who have not yet learned the words *blue* and *green*, and in the RVF for those who have. The dependent measure is the estimated mean of the log transformation of initiation time (ms), estimated for when variance due to age is accounted for. Error bars are within-subjects 95% confidence intervals calculated by using the error term from the three-way interaction. **Source:** [10].

categories are indeed permanently erased by language, rather than merely temporarily suppressed by it. Tellingly, split-brain patients, in whom the hypothetical channel of suppression has been severed, show no evidence of LVF/RH CP [8,12]. Further research is needed to settle this question, along with the other questions opened by these hemisphere-switching findings from infants and toddlers.

So far, our discussion of the brain bases of color CP, and its lateralization, has been based on inference from behavioral experiments. Recently however a number of ERP and fMRI studies have probed this issue more directly. Three studies deal with lateralization *per se*. Liu et al. [22] performed a lateralized visual search experiment using the procedure of [8], while recording brain activity by EEG. They focused on a specific ERP (event-related potential) component that is typically evoked in visual search for a perceptually unique target among distractors. They found that this component was present in the contralateral hemisphere for both within- and cross-category targets - but only in the left hemisphere did the amplitude of this component for cross-category targets exceed that for within-category targets, providing electrophysiological confirmation of the (behaviorally established) lateralization of the categorical perception of color. In a lateralized event related fMRI study, Siok et al. [23] found that discriminating colors of different lexical categories (versus the same category) elicited faster and stronger response in the left hemisphere language regions, especially when the colors were presented in the right visual field. They found further that only for these same stimuli was activation significantly enhanced in the visual areas responsible for color perception, suggesting that the language areas might act as a top-down control source on the visual areas in color perception. Another fMRI study [24] found that speeded perceptual discrimination of colors which are easy to name, relative to colors that are harder to name, activates left hemisphere regions that mediate language processes: these same regions were independently shown to be activated when subjects named colors aloud.

Three other recent ERP studies examine neural correlates of color CP more generally, in the context of behavioral tasks requiring the visual discrimination of colors. Fonteneau & Davidoff [25] found that the peak latency in ERP for a small color difference that bridged a color term boundary (195 ms) was shorter than that for an equally-sized color difference that did not bridge a color term boundary (214 ms) - providing evidence for a neural correlate of categorical perception of color. Thierry *et al.* [26] found that speakers of Greek, which has separate basic color terms for light and dark blue, showed visual mismatch negativity, an index of automatic and preattentive change detection, for luminance modulations in the blue area, whereas English speakers did not, a difference that appeared at around 200 ms. These results imply a role of the language-specific category boundary early in color discrimination. Holmes *et al.* [27] compared English-speaking subjects' brain responses to same- versus cross-category color differences; ERPs showed shorter latencies for early components to cross-category differences than for within-category differences, mirroring RT results on a behavioral version of the task and providing evidence for an early role for categorical differences in color perception. Later components also differentiated between- and within-category differences, suggesting influence of color categories on post-perceptual processing as well.

Altogether, the recent electrophysiological and imaging studies (1) support earlier behavioral findings of linguistic category influence in color discrimination via involvement of left hemisphere language regions, (2) demonstrate that CP influence for color can be exerted both early and late in processing, and is thus likely to be partially perceptual as well as post-perceptual, and (3) show that lexical differences between languages (Greek *vs.* English) are reliably reproduced in a standard ERP indication of pre-attentive novelty of a stimulus. Also, (4) one study suggests that regions of the left hemisphere that process language act as a top-down influence on the function of visual areas in color perception.

More broadly, it now appears uncontroversial that once language is learned, its categories shape perceptual discrimination primarily in the left hemisphere/right visual field and less so if at all in the right hemisphere/left visual field - a specific and unexpected sense in which Whorf was half right.

Where do color categories come from?

If linguistic color categories do affect perception, at least in half the visual field, where do those categories come from? Why do languages have the color categories they do? An exploration of this question suggests another way in which Whorf was half right.

According to an influential universalist view, color categories across languages are organized around the universal focal colors black, white, red, green, yellow and blue [28–30]. This view gained support from the finding that color memory appeared to be privileged for these colors, even in speakers of a language with a color naming system quite different from that of English [29,30]. For many years the prevailing consensus was that these focal colors, and prelinguistic color categories organized around them, reflected in prelinguistic CP as discussed above, consti-

tuted a cognitive foundation for universals of color naming, and the debate seemed settled.

Starting in 1999, however, the debate was re-opened by a study of the Berinmo language (Sepik-Ramu family, Papua New Guinea) by Debi Roberson and colleagues [31,32]. In this study, Roberson and colleagues failed to replicate the finding of privileged memory for the proposed focal colors; in addition, they found categorical perception of color at language-defined boundaries. They concluded that there are no universal foci, that categories therefore cannot be organized around them, and that “color categories are formed from boundary demarcation based predominantly on language” [32], subject to the constraint of ‘grouping by similarity’: namely, that categories must form contiguous regions of color space. The implication is that apart from that rather loose constraint, category boundaries are determined exclusively by local linguistic convention. These authors held up Berinmo as a counterexample to universals of color naming: ‘no evidence was found for [Berinmo color categories] to correspond to a limited set of universal basic color categories’ [33] and they interpreted their results as providing ‘evidence in favor of linguistic relativity’ [32]. Thus, this study constituted a Whorfian challenge to the then-reigning universalist consensus concerning why languages have the categories they do.

Since then, considerable evidence has been advanced for universal tendencies in color naming [e.g. 34,35], as well as evidence supporting relativity [e.g. 6,17,36,37]. This evidence has often been presented as unambiguously supporting one side or the other in the debate, in articles with titles such as ‘Focal colors are universal after all’ [35] and ‘Color categories vary with language after all’ [38]. Below we briefly review recent contributions to this exchange, and then discuss subsequent proposals that might help to reconcile these two positions, and perhaps resolve the debate.

Kay and Regier [39] compared the boundaries of Berinmo color categories with those of the 110 languages of the World Color Survey (WCS) [40]. They reasoned that if the only constraint on color categories is that they must occupy contiguous regions of color space, as had been suggested, there should be nothing privileged about the location of category boundaries in Berinmo – the boundaries could just as easily have been drawn elsewhere as long as each category remained a connected region. To probe this, they created artificial variants of the Berinmo color naming system by rotating that system in the hue dimension by 2,4,... etc. hue columns, all the way around the hue circle, as illustrated in Figure 3(a) - preserving the shape and thus contiguousness of the categories, but altering their actual location in color space. For each version of Berinmo, it was determined to what extent category boundaries in that system matched those in each language of the WCS. It was found that the unrotated (attested) version of Berinmo matched boundaries in the WCS better than did any rotated (hypothetical) version of Berinmo, as shown in Figure 3(b). This finding demonstrates that color naming in Berinmo reflects universal forces that constrain the location, not just the connectedness, of categories.

Lindsey and Brown [41] tested for universal tendencies in color naming in a different way, using the same WCS

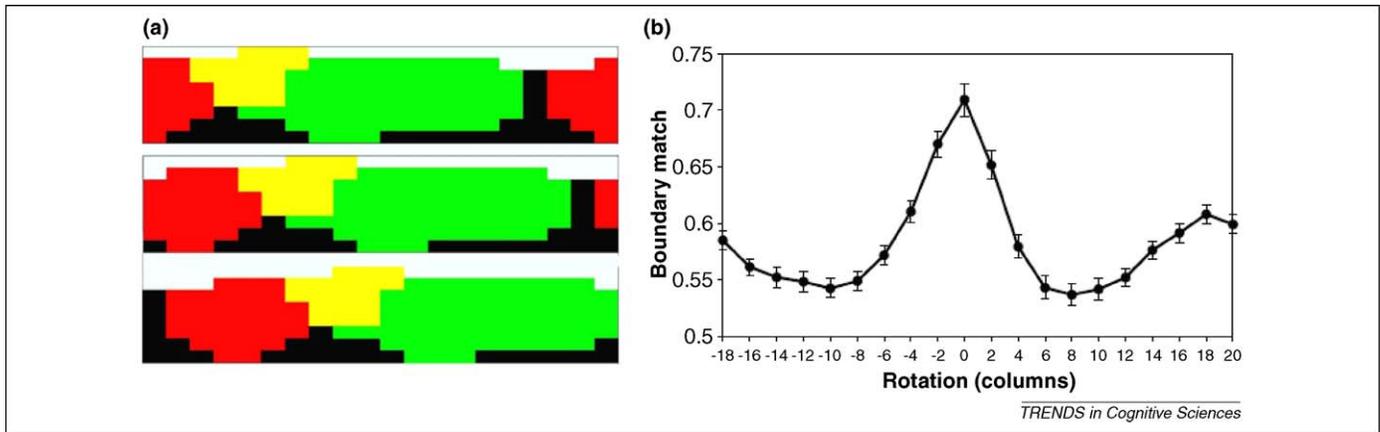


Figure 3. (a) The Berinmo color naming system (top) unrotated, and rotated four (middle) and eight (bottom) hue columns. Each colored region denotes the extension of a named color category, against a standard color naming grid. (b) Boundary match between each version of Berinmo, including the unrotated version, and WCS languages overall. For each version of Berinmo, the dot shows the mean boundary match with all WCS languages, and the bar shows the 95% confidence interval of the mean. **Source:** [39].

data. They considered the color naming systems of individuals, rather than summary measures capturing the systems of entire languages, as previous analyses had used, and performed cluster analyses based on these individual-level data. They found clusters that ‘generally correspond to readily identifiable English categories ... or their composites’. For instance, a cluster analysis with two clusters revealed categories ‘readily described as WARM (including red, yellow, orange, and pink) and COOL (including blue and green)’ [41, p. 16609]. Cluster analyses with more clusters revealed more fine-grained categories that again closely resemble those isolated on a less rigorous basis in the earlier literature [e.g. 42, 43]. Further recent studies also establish universal constraints on color naming [44, 45].

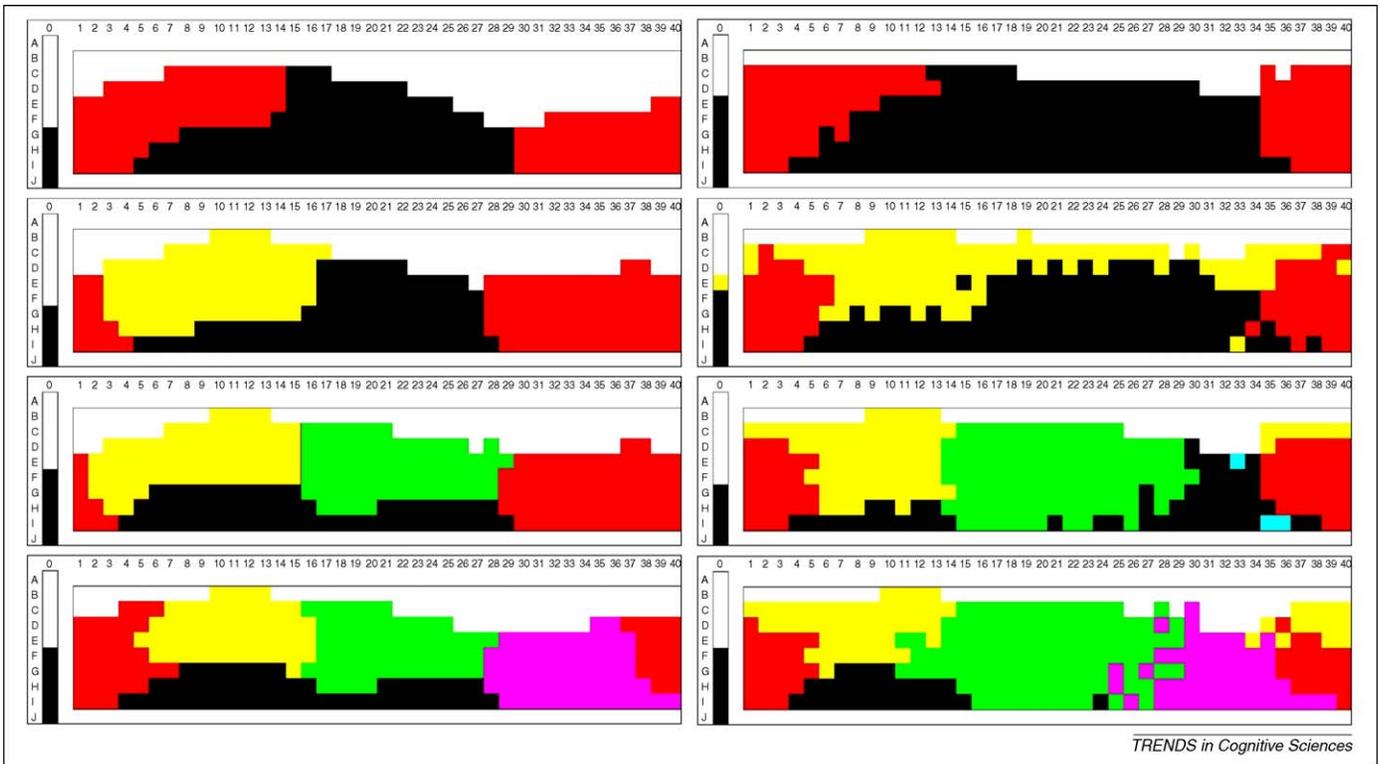
However, there is also evidence consistent with the relativist view that boundaries are determined by language. Roberson, Davidoff, Davies, and Shapiro [6] compared color naming and cognition in English, Berinmo, and in Himba, a language of Namibia. They found that Berinmo and Himba had similar but distinct systems of color naming. On a universal-foci account, Himba and Berinmo would receive the same analysis: each language has terms that appear to be focused in black, white, green, red, and yellow. Yet some of the boundaries among these categories differ noticeably between the two languages, and speakers of each language exhibit categorical perception of color at their native language’s boundaries. This finding suggests strongly that there is more to the determination of boundaries than groupings of universal foci, and that linguistic convention might indeed play some role in determining where boundaries are drawn. This view is confirmed by recent evidence from Greek–English bilinguals, in which the foci themselves shift their position somewhat, under the influence of the categories of the newly-learned language [37]. Finally, and most generally, qualitative inspection of the 110 languages of the WCS reveals some language-specific idiosyncrasies against the backdrop of broad universal tendencies [46].

On the whole, then, the universalist and relativist views regarding color naming are each partly supported and partly challenged by existing empirical data. This empirically complex picture can be accounted for in theoretically

simple terms. Jameson & D’Andrade [47] suggested that patterns of color naming could be attributed to irregularities in the shape of perceptual color space combined with general human cognitive tendencies toward constructing informationally efficient systems of concepts:

One possible explanation [for patterns of color naming in languages] is...the irregular shape of the color space. ... Hue interacts with saturation and lightness to produce several large “bumps”; one large bump is at focal yellow, and another at focal red. ... We assume that the names that get assigned to the color space...are likely to be those names which are most informative about color [47, p. 312].

This proposal can be seen as a natural generalization of the universal-foci account, in which every color is focal to some extent - the extent to which it protrudes from the irregular outer skin of color space; the originally proposed universal foci are then simply ‘more focal’ than other colors. The appeal of the proposal lies in the possibility that simple principles of categorization, operating over this irregular surface, might account for both the universal tendencies and the deviations from them that we see in the world’s languages. Regier, Kay & Khetarpal [48] formalized this idea and tested it against the WCS data. They defined a well-formedness measure that captures the extent to which a given categorical partition of the outer skin of color space maximizes perceptual similarity within color categories and minimizes it across categories [49]. Given this, they created theoretically optimal color naming systems for $n=3,4,5,6$ categories, through steepest ascent in well-formedness. The results are shown in Figure 4, compared with selected color naming systems in the WCS. This demonstrates that some languages have color naming systems that are near optimal in this sense [cf. 50]. While other languages deviated considerably from these templates, across the languages of the WCS there was a strong tendency for color naming to be shaped in part by well-formedness. Specifically, for each language in the WCS, rotated variants were created as in Figure 3(a), and for 82 of the 110 WCS languages the unrotated system was more well-formed than any rotated variant – thus



TRENDS in Cognitive Sciences

Figure 4. Theoretical optima (left) for $n=3,4,5,6$ categories, compared with color naming systems (right, top to bottom) of Ejjagam (Nigeria/Cameroon), Culina (Peru/Brazil), Iduna (Papua New Guinea), Buglere (Panama). All results mapped against a standard color grid, in which columns denote hues and rows denote lightness levels. **Source:** [48].

well-formedness provides an explanation for universal tendencies. At the same time, this account also suggests where linguistic convention might get some wiggle room. There are often several similar but different partitions that are roughly equally well-formed. Hence, local factors, including arbitrary linguistic convention, might select from among a group of highly-ranked systems that differ somewhat from each other - such as Berinmo and Himba. This recent account suggests a middle ground between nature and nurture in color naming across languages, and highlights another sense in which Whorf might have been half right.

A competing explanation holds that color naming patterns reflect the statistical distribution of colors in the world, rather than the shape of perceptual color space in our minds. Yendrikhovskij [51] showed that k -means clustering of colors obtained from images of natural scenes produced color clusters similar to those observed in the world's languages. Is color naming shaped by exogenous color diet, or by endogenous perceptual structure? The two explanations might be partially reconcilable. Shepard [52,53] argued that universal aspects of perceptual-cognitive structure might correspond to invariants in the environment, internalized over the course of evolution [cf. 54]. Thus, it is possible that color naming reflects the structure of perceptual color space in our minds - which in turn reflects the distribution of colors in the world.

There is an obvious fact that our discussion so far has not engaged: color naming systems are used for the social process of *communication* about color, and they are passed on socially from one generation to the next within a speech community. What is the role of such social forces in accounting for patterns of color naming? This question

has been increasingly addressed recently, often in agent-based simulations [55–61]. For example, Steels and Belpaeme [55] argued on the basis of such simulations that color categories might in principle reflect both perceptual structure and social coordination through language; Dowman [57] accounted for patterns of color naming in the world's languages in broadly similar terms, through cultural evolution in an idealized color space; and Komarova *et al.* [58] show that, given certain simple assumptions, a population of agents communicating about color will converge to a system of near-optimal color categories.

Of particular interest in such simulations is the question of how much explanatory force is carried by social coordination and transmission *per se*, and how much by the individual agents' biases and expectations concerning the color domain. In other words, how much of the explanation lies *between* individuals, and how much *within* individuals? A possible answer is suggested by some recent work by Griffiths and Kalish [62]. They presented a Bayesian analysis of the social transmission of language, and showed that under certain assumptions, this process converges to a distribution over languages that reflects the shared prior bias of each learner in the social chain. Thus, the eventual effect of socially transmitting language across generations is that language itself takes on the form of the learning bias in each learner's mind; in the case of color naming, this would be our prior expectations concerning the shape of color categories. And as argued above, this mental structure might in turn reflect invariants in the distribution of colors in the world. Taken as a whole, such an account might help to explain the success of some models that are based on color diet, or on the mental representation of color, with no social component at all: environmental

structure might be internalized in the mind over evolutionary time, and then externalized in the form of language through iterated social transmission, with the result that models capturing this structure either in the environment, or in the mind, can fit color naming data well. At the same time, there are insights that seem less likely to be easily capturable in these terms. For example, Jameson and Komarova [61] examined agent populations in which not all agents had the same biases, capturing various mixtures of normal and anomalous color vision. They concluded that a rather small percentage of agents with anomalous color vision could greatly affect the color category boundaries of the population - suggesting a possible explanation for some of the cross-language variation observed.

Conclusion

We have reviewed two broad recent findings here: that language affects color perception primarily in the right visual field probably via activation of language regions of the left hemisphere, and that color naming reflects both universal and local determinants. Neither of these findings is anticipated by the traditional universalist-versus-relativist framing of the debate over language and perception, and neither sits particularly comfortably with it. Instead, these findings suggest novel perspectives on the relation of language and perception. There is some evidence already available that these findings generalize to domains beyond color. The replication of the initial results on the lateralized Whorf effect in color [8,9] to dog and cat silhouettes [12] indicates that RVF-lateralized categorical perception might not be restricted to the color domain (and ongoing research suggests that optimization of well-formedness might generalize to the spatial domain). Just which of the findings reported here for color will generalize to other domains of language or to cognition more broadly, and which are restricted to color, should furnish significant research questions in the future (see also Box 1).

Box 1. Outstanding questions:

- In what other semantic domains, if any, are categories near-optimal?
- In what other semantic domains, if any, does language affect perception in half the visual field?
- What becomes of prelinguistic color categories in the right hemisphere, and by what mechanism?
- Are linguistic color categories elaborations of prelinguistic categories?
- Are non-universal aspects of color naming attributable solely to arbitrary linguistic convention, or to environmental or other non-universal forces?

Acknowledgments

This work was supported by NSF grants 0418283 (to TR) and 0418404 (to PK).

References

- 1 Kay, P. and Regier, T. (2006) Language, thought, and color: Recent developments. *Trends in Cognitive Sciences* 10 (2), 51–54
- 2 Kay, P. and Kempton, W. (1984) What is the Sapir-Whorf hypothesis? *American Anthropologist* 86, 65–79
- 3 Özgen, E. and Davies, I.R.L. (1998) Turkish color terms: Tests of Berlin and Kay's theory of color universals and linguistic relativity. *Linguistics* 36, 919–956
- 4 Roberson, D. and Davidoff, J. (2000) The categorical perception of colours and facial expressions: The effect of verbal interference. *Memory and Cognition* 28, 977–986
- 5 Winawer, J. et al. (2007) Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences* 104 (19), 7780–7785
- 6 Roberson, D. et al. (2005) Colour categories in Himba: Evidence for the cultural relativity hypothesis. *Cognitive Psychology* 50, 378–411
- 7 Roberson, D. et al. (2008) Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition* 107, 752–762
- 8 Gilbert, A.L. et al. (2006) Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences* 103, 489–494
- 9 Drivonikou, G.V. et al. (2007) Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences* 104, 1097–1102
- 10 Franklin, A. et al. (2008) Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Sciences* 105 (47), 18221–18225
- 11 Kay, P. et al. (2009) Lateralized Whorf: Language influences perceptual decision in the right visual field. In *Language, Evolution, and the Brain* (Minett, James W. and Wang, William S.-Y., eds), The City University of Hong Kong Press
- 12 Gilbert, A.L. et al. (2008) Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language* 105, 91–98
- 13 Bornstein, M.H. et al. (1976) Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance* 2 (1), 115–129
- 14 Franklin, A. and Davies, I.R.L. (2004) New evidence for infant color categories. *British Journal of Developmental Psychology* 22, 349–377
- 15 Franklin, A. et al. (2005) The nature of infant color categorization: Evidence from eye movements on a target detection task. *Journal of Experimental Child Psychology* 91, 227–248
- 16 Franklin, A. et al. (2005) Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology* 90, 114–141
- 17 Goldstein, J. et al. (2009) Knowing color terms enhances recognition: Further evidence from English and Himba. *Journal of Experimental Child Psychology* 102, 219–238
- 18 Davidoff, J. et al. (2009) Nature vs. nurture: The simple contrast. *Journal of Experimental Child Psychology* 102, 246–250
- 19 Franklin, A. et al. (2009) What can we learn from toddlers about categorical perception of color? Comments on Goldstein, Davidoff, and Roberson. *Journal of Experimental Child Psychology* 102, 239–245
- 20 Franklin, A. et al. (2008) Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences* 105, 3221–3225
- 21 Hanley, R. and Roberson, D. (2008) Do infants see colors differently? *Scientific American* May 14
- 22 Liu, Q. et al. (2009) The N2pc component in ERP and the lateralization effect of language on color perception. *Neuroscience Letters* 454, 58–61
- 23 Siok, W.-T. et al. (2009) Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences* 106, 8140–8145
- 24 Tan, L.-H. et al. (2008) Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences* 105, 4004–4009
- 25 Fontenau, E. and Davidoff, J. (2007) Neural correlates of colour categories. *NeuroReport* 18, 1323–1327
- 26 Thierry, G. et al. (2009) Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences* 106, 4567–4570
- 27 Holmes, A. et al. (2009) Neurophysiological evidence for categorical perception of color. *Brain and Cognition* 69, 426–434
- 28 Berlin, B. and Kay, P. (1969) *Basic Color Terms: Their Universality and Evolution*. University of California Press, (Berkeley)
- 29 Heider, E.R. (1972) Universals in color naming and memory. *Journal of Experimental Psychology* 93, 10–20
- 30 Heider, E.R. and Olivier, D.C. (1972) The structure of the color space in naming and memory for two languages. *Cognitive Psychology* 3, 337–354

- 31 Davidoff, J. *et al.* (1999) Colour categories in a stone-age tribe. *Nature* 398, 203–204
- 32 Roberson, D. *et al.* (2000) Colour categories are not universal: Replications and new evidence from a Stone-age culture. *Journal of Experimental Psychology: General* 129, 369–398
- 33 Roberson, D. *et al.* (2002) Colour categories are not universal: Replications and new evidence. In *Theories, technologies, instrumentalities of color: Anthropological and historical perspectives* (Saunders, B. and van Brakel, J., eds), pp. 25–35, Lanham, Maryland, University Press of America Inc
- 34 Kay, P. and Regier, T. (2003) Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences* 100, 9085–9089
- 35 Regier, T. *et al.* (2005) Focal colors are universal after all. *Proceedings of the National Academy of Sciences* 102, 8386–8391
- 36 Agrillo, C. and Roberson, D. (2009) Colour language and colour cognition: Brown and Lenneberg revisited. *Visual Cognition* 17, 412–430
- 37 Athanasopoulos, P. (2009) Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition* 12, 83–95
- 38 Roberson, D. and Hanley, R. (2007) Color vision: Color categories vary with language after all. *Current Biology* 17, R605–R607
- 39 Kay, P. and Regier, T. (2007) Color naming universals: The case of Berinmo. *Cognition* 102, 289–298
- 40 Cook, R.S. *et al.* (2005) The World Color Survey database: History and use. In *Handbook of Categorization in Cognitive Science* (Cohen, H. and Lefebvre, C., eds), pp. 223–242, Amsterdam, Elsevier
- 41 Lindsey, D.T. and Brown, A.G. (2006) Universality of color names. *Proceedings of the National Academy of Sciences* 103, 16609–16613
- 42 Kay, P. *et al.* (1997) Color naming across languages. In *Color Categories in Thought and Language* (Hardin, C.L. and Maffi, L., eds), pp. 21–56, Cambridge, Cambridge University Press
- 43 Kay, P. and Maffi, L. (1999) Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist* 101, 743–760
- 44 Kuehni, R.G. (2007) Nature and culture: An analysis of individual focal color choices in World Color Survey languages. *Journal of Cognition and Culture* 7, 151–172
- 45 Webster, M.A. and Kay, P. (2007) Individual and population differences in focal colors. In *Anthropology of Color: Interdisciplinary multilevel modeling* (MacLaury, R.E., Paramei, G.V. and Dedrick, D., eds), pp. 29–53, Amsterdam, John Benjamins
- 46 Kay, P. *et al.* (in press). *The World Color Survey*. Center for the Study of Language and Information: Stanford University. Stanford, CA
- 47 Jameson, K. and D'Andrade, R. (1997) It's not really red, green, yellow, blue: An inquiry into cognitive color space. In *Color categories in thought and language* (Hardin, C.L. and Maffi, L., eds), pp. 295–319, Cambridge, UK, Cambridge University Press
- 48 Regier, T. *et al.* (2007) Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences* 104, 1436–1441
- 49 Garner, W.R. (1974) *The Processing of Information and Structure*. Erlbaum, (Potomac, MD)
- 50 Griffin, L.D. (2006) The basic colour categories are optimal for classification. *Journal of the Royal Society: Interface* 3, 71–85
- 51 Yendrikhovskij, S. (2001) Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology* 45, 409–417
- 52 Shepard, R. (1992) The perceptual organization of colors: An adaptation to regularities of the terrestrial world? In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (Barkow, J.H., Tooby, J. and Cosmides, L., eds), pp. 495–532, New York, Oxford University Press
- 53 Shepard, R. (2001) Perceptual–cognitive universals as reflections of the world. *Behavioral and Brain Sciences* 24, 581–601
- 54 Regan, B.C. *et al.* (2001) Fruits, foliage, and the evolution of primate colour vision. *Philosophical Transactions of the Royal Society B: Biological Sciences* 356, 229–283
- 55 Steels, L. and Belpaeme, T. (2005) Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences* 28, 469–489
- 56 Belpaeme, T. and Bleys, J. (2005) Explaining universal color categories through a constrained acquisition process. *Adaptive Behavior* 13, 293–310
- 57 Dowman, M. (2007) Explaining color term typology with an evolutionary model. *Cognitive Science* 31, 99–132
- 58 Komarova, N., Jameson, K. and Narens, L. (2007) Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology* 51, 359–382
- 59 Komarova, N. and Jameson, K. (2008) Population heterogeneity and color stimulus heterogeneity in agent-based color categorization. *Journal of Theoretical Biology* 253, 680–700
- 60 Jameson, K. and Komarova, N. (2009a). Evolutionary models of color categorization. I. Population categorization systems based on normal and dichromat observers. *Journal of the Optical Society of America*, 26, 1414–1423
- 61 Jameson, K. and Komarova, N. (2009b). Evolutionary models of color categorization. II. Realistic observer models and population heterogeneity. *Journal of the Optical Society of America*, 26, 1424–1436
- 62 Griffiths, T.L. and Kalish, M.L. (2007) Language evolution by iterated learning with Bayesian agents. *Cognitive Science* 31, 441–480