Shepard & Cooper 1992). In one society where rod monochromacy commonly occurs in the population, color normal individuals share a pragmatic categorical repertoire with achromatopes who perceive a "colorless" world (Sacks 1997). In other societies, other complexities arise during processes wherein perceivers learn through social interaction to use normative linguistic codes despite perceptual differences that could undermine the code's meaning (Jameson 2005a; 2005b; Jameson et al. 2001). Thus, within populations, variation in perceptually correlated knowledge is integral to the cognitive side of learning and sharing a color repertoire, but such human variation runs counter to Assumption (B).

Addressing both (A) and (B) as suggested here would permit S&B to make useful comparisons between perceptually grounded categories shared by uniform populations and those shared by nonuniform populations.

NOTES

1. This seems to work against the suggestion that "artificial agents might end up with a quite different categorical repertoire compared to . . . human beings" (sect. 1).

2. Just as S&B demonstrate different sets of "chromatic distributions . . . do not lead to categories that are similar . . ." (sect. 5.1), so too would very different category solutions arise if initially agents were given a honey-bee observer model, and these category solutions would almost certainly bear little resemblance to the category solutions they found using their agent populations.

3. Just as dichromats are accommodated by the CIE standard observer model, but have different known metameric class relations.

Seeing and talking: Whorf wouldn't be satisfied

Boris Kotchoubey

Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, 72074 Tübingen, Germany.
boris.kotchoubey@uni-tuebingen.de
http://www.uni-tuebingen.de/medizinischepsychologie/stuff/

Abstract: Although Steeles & Belpaeme's (S&B) results may be useful for development of technical devices, their significance for behavioral sciences is very limited. This is because the question the authors asked was "Why do people use similar words in a similar way?" rather than "How can similar words stand for similar experience?" The main problem is not shared word usage, but shared references.

Polonius: What do you read, my lord?
Hamlet: Words, words, words.
—*Hamlet*, Act II, Scene II

The clarity with which the target article is written makes the critique easier. The main goal is formulated from the very beginning: To explore how colour words "may become sufficiently shared among the members of a population" (sect. 1) so that if I say "red" everybody can select a red (and not a yellow) object from a presented set. Moreover, Steels & Belpaeme (S&B) make no secret that this "goal is entirely practical . . . to design . . . robots that are able to do this task." (sect. 1) Though I am not an expert in robotics, it appears that the authors attained substantial progress in approaching their goal.

The question is, however, whether this pragmatic approach can shed light on the real mechanisms in question. I agree that the study can contribute to "designing agents that are able to develop a repertoire of . . . categories that is sufficiently shared to allow communication" (sect. 6). But I doubt that "these results are relevant to . . . an audience of cognitive scientists" (sect. 6) who are interested in the psychology of colour perception. Although the authors admit that "the artificial agents might end up with a quite different categorical repertoire compared to . . . human beings," (sect. 1) they miss a much worse peril, that their agents come to categories very similar to human categories (thereby creating the illusion of relevance), but using processing means that have nothing in common with those used by human brains.

S&B suggest that their data support the Sapir–Whorf thesis on the dependence of colour perception on language. This thesis has been formulated in rather ambitious terms, for instance, by Sapir: "We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation" (cited by Whorf 1962, p. 134), or by Whorf's commentator S. Chase: "Speakers of different languages see the Cosmos differently" (ibid, p. x). Particularly, Whorf emphasised the importance, not only of verbal categories, but rather of the syntax of different languages (e.g., tenses, subject–predicate structure, use of plurals and singulars, etc.), in organisation of our basic mechanisms of perceiving and conceiving of the world.

This expected relationship to the very structure of colour experience is lacking in the target article. Not sharing perception (e.g., the fact that you see red where I also see it) but sharing word usage is the problem the entire study is pivoted around. By the way, colour may not be the best case for study interaction between sensory and cognitive factors because the sensory information can only be obtained with central vision (there are no cones on the periphery) and high luminance (cones do not work in twilight), hence one may state that we see most objects grey most of the time. But the main point is that mere agreement in verbal behavior does not prove the agents' similarity in their "segmentation of the face of nature" (Whorf 1962, p. 241).

Of course, we cannot really know another person's sensory qualia (e.g., the qualium of redness), but we can approach this knowledge by using a broad range of methods, beyond categorisation and naming. And probably the most reliable result obtained to date is that if we vary tasks, conditions, instructions, cue availability, and so forth, so also varies the role of language as a determinant of behavior. Thus, the long-assumed effect of language spatial terms, such as "on the left of" or "to the north of,"on space perception proved to be the effect of available spatial cues. Natural peoples, when tested in their natural conditions, use significantly more objective (allocentric) spatial cues than Europeans (Dutch or English) tested in the lab. Also English-speaking people, without changing their mother language, use more allocentric cues when tested outdoors as compared to being in a closed room with blinds pulled down (Li & Gleitman 2002). The availability of potentially useful information appears, therefore, to exert a stronger effect on space perception than the language itself.

Turning back to colours, the data are not very different. For example, most European languages have one basic term for blue, whereas Russian has two; a popular Russian children's song listing "the seven colors of the rainbow" mentions light-blue and dark-blue as two completely different colours, the latter being close, but not identical, to purple. Nevertheless, being presented with a large number of green and blue colour tones, Russian and English subjects did not differ in their classification; particularly, Russians did not tend to group dark and light blue separately (Davies & Corbett 1997). There is no evidence that English speakers are unable to distinguish those hues that Russian speakers do.

Kay and Kempton (1984) developed colour triads, such as one containing two green colours and one blue. One of the green colours (Green 1) was separated from the other green (Green 2) by a larger number of just noticeable differences than from Blue. When asked to choose the stimulus that looked least like the other two, subjects chose Blue. However, when asked to compare stimuli pairwise, they found Green 1 and Green 2 more different than Green 1 and Blue. The issue may be even more complicated because neuropsychological data indicate that a patient who performed like controls in this experiment (and who, therefore, could distinguish between classification and similarity judgment) was nonetheless unable to classify colours according to their names. His sorting was based on superficial perceptual similarity (Robertson et al. 1999). This may indicate that not only the presence of

https://doi.org/10.1017/S0140525X05360080 Published online by Cambridge University Press

verbal cues can substantially affect the result of classification, but also the explicit versus implicit nature of those cues.

To summmarise, the Whorfian question was formulated (Li & Gleitman 2002, p. 267) as follows: "Do the differences in how people talk create the differences in how they think?" The target article, in contrast, answered a quite different question: "Do the differences in how people learn to talk create the differences in how they subsequently talk?" It is not surprising that the answer to the latter question was positive, but this does not permit any conclusion concerning the former one.

Not all categories work the same way

Sidney R. Lehky

Computational Neuroscience Laboratory, The Salk Institute, La Jolla, CA 92037. **sidney@salk.edu**

Abstract: The relative contributions of biological and cultural factors in determining category characteristics almost certainly vary for different categories, so that the results of these simulations on color categories don't necessarily generalize. It is suggested here that categories that pick out structure in the environment of strong behavioral significance to individual agents will be predominantly biologically determined and will converge without interagent communication, whereas those categories that serve primarily to coordinate behavior in a population will require communication to converge.

The computer simulations described in Steels & Belpaeme's (S&B's) article provide an interesting example of a situation in which language communication amongst a population of agents can affect the development of color categories. Although the empirical situation regarding color categories, of course, remains to be determined, these theoretical studies will be valuable in constraining the debate about what is possible.

It seems likely, however, that the potential for cultural shaping of perceptual categories can differ sharply depending on the particular category at hand. Some categories may be more culturally dependent, others may have a stronger learning component, and finally, some may be genetically hard-wired. In other words, there may be different categories of categories, and the results of studying one sort of category won't necessarily generalize to others.

What characteristics might in general distinguish more culturally dependent perceptual categories from the more biologically dependent ones? (Here I lump together genetic evolution and individualistic learning under "biology.") I would suggest that if a category represents structure in the environment that is of strong behavioral significance to each individual agent, then that category will develop in a predominantly biology-dependent manner such that all agents in the population, without communicating, will share the same category. If a category does not directly distinguish any behaviorally critical feature of the environment, but rather serves to coordinate the behavior of agents in the population, then communication between agents will be required to ensure the convergence of category properties.

Let us see how this distinction might operate in the context of color categories. CIE (Commission Internationale de L'Éclairage, or International Commission of Illumination) color space, of whatever variant, is a continuous space. The question arises as to in which situations is it advantageous to discretize this continuum into a small set of fixed categories. Two possibilities will be given here, corresponding to the distinction made earlier.

The first is if there were a small set of special colors that flag aspects of the environment that have overriding significance to the agents behaviorally (perhaps related to mating, food selection, predator evasion, etc.). There could, in that case, be an advantage in creating color categories centered on these special colors in order to highlight them for structures associated with implementing decisions and motor responses. To the extent that the embodiment characteristics of agents within a population are essentially the same (similar sensory apparatus, motor capabilities, etc.) and they have similar behavioral repertoires, it seems a reasonable possibility that all agents will converge to these same color categories independent of interagent communication.

The second situation is if agents needed to communicate information about color to each other. In this case, discretization of the color space reflects the discrete nature of the vocabulary used to describe it. Here the color categories don't correspond to anything that is of behavioral significance to an agent operating in isolation. A different set of categories would neither enhance nor detract from the survival prospects of the isolated agent. Thus, there is little pressure for isolated agents to develop the same categories. It is only within a population that the categories acquire significance, and the categories converge through interagent communication to coordinate behavior within the population.

Without the presence of a set of ecologically "special" colors and without language, the "discrimination game" described in the target article could probably be implemented in a robot by setting receiver-operating characteristics within a signal detection model, without fixed categories. Although statistical clustering of natural inputs can lead to the creation of color categories, it is not clear what benefits arise from building a robot with categories derived in this manner, other than perhaps somewhat more efficient encoding of sensory inputs (in an information-theoretic sense, Simoncelli & Olshausen 2001). It is also possible that characteristics of the sensory apparatus may lead to biases in color category formation in noncommunicating agents, so that there is some degree of correlation in the categories formed by them (as indeed we saw in the simulations in this article). However, if these embodiment-specific effects are confined to the input (sensory) stages of the system and do not translate to something behaviorally meaningful, as one considers the agent in its sensorimotor entirety, they may not provide a sufficient drive to strongly coordinate the color categories of noncommunicating agents (again, as we saw in this article).

Moving away from color categorization, consider more abstract categories, such as animals versus non-animals, or food objects versus non-food objects. Membership in these categories can rapidly be determined visually by both humans and nonhuman primates (Fabre-Thorpe 2003). These are examples of perceptual categories that are of strong behavioral significance to individual organisms, perhaps more so than the color categories formed by humans. The expectation here is that individuals undergoing unsupervised learning in their natural environment will be able to converge to the same visual categorization of food versus non-food items (for example), without any communication amongst themselves, to a greater degree than color categories will converge for non-communicating agents.

On sticking labels

Jan Pieter M. A. Maes

Department of Psychiatry, University Hospital Antwerp, 2650 Edegem, Belgium. **janpietermaes@hotmail.com.**

Abstract: Steels & Belpaeme (S&B) are clearly interested in the possible test their models may provide for human language theories. However, they only superficially address the assumptions underlying their own agent architecture, while these are of crucial relevance to the topic of human language. These assumptions fit an Augustinian picture of language, which Wittgenstein challenges in his *Philosophical Investigations*. It is too early to draw conclusions regarding human language evolution from such models.

> Could a machine think? – Could it be in pain? – Well, is the human body to be called such a machine? It surely comes as close as possible to being such a machine.
> —Wittgenstein (1953)