Abstracts for Scheduled Presentations

(1) From Basic Color Terms to the World Color Survey: A self-indulgent chronical

Brent Berlin
(obberlin@gmail.com)
Anthropology, University of Georgia

&

Paul Kay
(pauldeykay@gmail.com)
Linguistics and International Computer Science Institute, University of California, Berkeley

In 1964-65 two recent Anthropology PhDs discovered in a casual conversation that in their dissertation field investigations in widely separated places with unrelated languages each had made a surprising observation regarding the local color terms. Moreover, the two observations were strikingly similar and they jointly challenged the hitherto unchallenged doctrine that, “Each culture has taken the spectral continuum and has divided it upon a basis which is quite arbitrary (Ray 1952).” That conversation led to a research seminar in Berkeley two years later and eventually to the World Color Survey. This presentation chronicles the events that led from the original conversation up to the undertaking of the WCS.

(2) The World Color Survey (WCS) Online Digital Archive

Richard Cook

(rscook@rscook.org)

International Computer Science Institute, University of California, Berkeley

The World Color Survey (WCS) set out in the mid-to-late 1970's to test theories advanced by Berlin and Kay (1969) regarding the existence of: (1) universal constraints on cross-language color naming; (2) an evolutionary progression by which languages gain color terms over time.

The WCS Digital Archive which went online for the first time in 2002 at the *International Computer Science Institute* (ICSI) included digital transcriptions and documentation of the elicitation materials (coding sheets) for all 110 WCS languages.

<http://www1.icsi.berkeley.edu/wcs/> <http://www1.icsi.berkeley.edu/wcs/data.html>

The original paper documents from which those digital transcriptions derived are old, fragile, nuanced and priceless. Imaging these documents would make them available to future researchers, to help experts resolve questions in the digital transcription data. But the paper archive is large and complex, with nearly 6,000 pages of elicitation coding sheets and associated fieldworker notes, in a variety of paper sizes and configurations (stapled and loose-leaf booklets). Only recently have advancements in digital scanning technology made it possible to produce high-resolution color scans of the complete WCS paper archive, in a quick, error-free and cost-effective way. In this presentation we will discuss some of the background of the decades-long WCS data digitization project, outlining stages of development, methods and technologies used, and plans to publish a new augmented WCS archive online. Despite decades of work, a number of issues remain to be addressed before the archive can be made generally available.

Sample images of the original WCS coding sheets:
<http://www1.icsi.berkeley.edu/wcs/images/samp/Chiquitano7samp.pdf>
<http://www1.icsi.berkeley.edu/wcs/images/samp/Chiquitano7samp_85x11.pdf>

The WCS book:
<https://web.stanford.edu/group/cslipublications/cslipublications/site/9781575864150.shtml>

The online version of the WCS book:
<http://www1.icsi.berkeley.edu/wcs/analyses/PDF/index.html>

The Index of WCS languages (and ISO 639-3 Codes):
<http://www1.icsi.berkeley.edu/wcs/WCS_SIL_codes.html>

Instructions to WCS fieldworkers:
<http://www1.icsi.berkeley.edu/wcs/images/WCS_instructions-20041018/jpg/border/index.html>

(3) Evolutionary models of color categorization

Natalia Komarova
(komarova@uci.edu)
Mathematics Department, University of California, Irvine

Methods of evolutionary biology and population dynamics are used to study the problem of color categorization. What is an optimal partitioning of a color space into color categories? How do individuals manage to arrive at a shared color categorization system that allows successful communication? What are symmetry properties of this mathematical problem, and what is the role of population heterogeneity? Elements of game theory are used to address some of these questions.

(4) A mathematical approach to color categorization studies

Nicole Alexis Abello Fider
(nfider@uci.edu)
Mathematics Department, University of California, Irvine

Color categorization in humans is a topic in psychology and linguistics which can shed light on human thought and perception in general. Although individuals can divide the color space in different ways, it is accepted that in a linguistically unified society there exists a specific set of basic color categories which speakers use when categorizing the color space. These categories give members of the population the ability to communicate color information with each other, and can evolve over time as the culture and language evolve. We believe that dynamic changes are less likely to occur within categories and more likely to occur on or around category boundaries. We present a mathematical method of identifying a language's set of color categories and boundaries based on color-naming data provided by the World Color Survey Data Archives, and we present interesting cases that appear when we study color categorization trends on gender-separated data. We discuss the possible dynamics of category evolution and how they can be related to the numerical data.

(5) Further evolution of natural categorization systems: A new approach to evolving color concepts

Maryam Gooyabadi
(mgooyaba@uci.edu)
Institute for Mathematical Behavioral Sciences
University of California, Irvine

To assess the effectiveness of a simulation-based approach to linguistic category evolution, a language evolution model is applied to the natural color categorizations of 108 linguistic communities from the World Color Survey (2009). This dynamic model is derived from Komarova et al. [Journal of Mathematical Psychology 51, 359–382 (2007)], and evolves color-naming systems to a stable equilibrium through agent interactions. This simulation-based approach remedies the sparseness of empirical, diachronic data by broadly approximating evolutionary trends in an idealized form. Additionally, we present novel explorations of evaluating equilibrium stability and determining category boundaries. Results show that all 108 systems evolve to a stable equilibrium while maintaining key features of its original categorization, suggesting that our simulations are a suitable representation of natural evolutionary processes. This approach can have valuable insights and implications for research where diachronic data is sparse. For example, weighing in on the linguistic debate between the Emergence and Partition Hypotheses of category evolution, our analyses finds evidence in favor of the Emergence Hypothesis.

(6) Two Italian basic blue categories: Visualizing diatopic variation of referential meanings

Galina V. Paramei
(parameg@hope.ac.uk)

Department of Psychology, Liverpool Hope University, UK

Standard Italian has more than one "blue" basic color term (BCT): *blu* is unanimously glossed as "dark blue", while *azzurro* is referred to as "light blue" and/or "medium blue". We explored diatopic variation in denotata of *blu*, *azzurro* and *celeste* "sky blue" in a psycholinguistic experiment, with unconstrained color naming, conducted in Alghero (Sardinia) and Verona (Veneto region). Participants named Munsell chips (N=237), embracing the BLUE area (8 charts: 7.5BG—5PB), under fixed luminance conditions. For each Italian modal blue term (*blu-, azzurr-, celest-*), a referential volume of colors with naming consensus was fitted by a convex hull, visualized in CIELAB space. We found that, while the referential extent and foci of *blu* are similar, the referential extents of *azzurro* and *celeste* differ markedly between the two regions. Verona speakers denote by *azzurro* "light-and-medium blue"; in contrast, Alghero speakers use mainly *celeste* for a similar color space extent, with *azzurro* being constrained to darker "medium blue". We argue that the prominence of *celeste* in the Algherese Catalan dialect is contact-induced (cf. Catalan *celeste* "light blue"). *Celeste* (from Latin *caeruleus*) also appears to be reinforced by the insularity and conservatism of Sardinian dialects, to the detriment of mainland BCT *azzurro*.

Acknowledgments: The author is grateful to Cristina Stara and Mauro D'Orsi for collecting data in Alghero and Verona, respectively, and to Gloria Menegaz, University of Verona, for invaluable guidance in applying the convex hull algorithm.

(7) The sparse, distributed, diverse mapping of color terms onto color space, Part I: Motifs and Hering theory

Angela Brown
(brown.112@osu.edu)
College of Optometry, The Ohio State University
&
Delwin Lindsey
(lindsey.43@osu.edu)
Psychology, The Ohio State University

The study of color naming is implicitly based on a shared mental representation of color that prevails throughout a language community. In our work, we found that the naming of colors is distributed across a community of speakers, with each individual adopting the names of a subset of the colors in that shared representation. Thus, individual color naming is diverse even within a language community. However, the striking similarity of individual color naming patterns around the world suggests that some aspects of the shared representation are common to all people.

The diversity among named color categories observed within a language is mainly due to variation in how the color terms are mapped onto color space. These mappings can be grouped into about four—six "motifs," which are structurally similar to some of the "stages" of color term evolution, classically described 50 years ago by Berlin&Kay. These motifs re-occur worldwide, in completely unrelated languages. Here, we report recent data on Somali, English, and Japanese, which were collected prospectively within this framework.

One influential approach to the universality of color naming is to link the named color categories to the six elemental color sensations of Hering's theory of color appearance. However, terms for Hering's six color sensations are neither all necessary (some people use fewer than six terms) nor jointly sufficient (others use many more than six color terms). Furthermore, many Somali observers, as well as many informants in the World Color Survey, use a single color term for both yellowish and bluish colors. This violates the mutual exclusivity of blue and yellow required by Hering's theory. Thus, it is not obvious to us how the Hering theory of color appearance can be identified with the common sensory experience that has guided the naming of colors over the millennia since language began.

(8) The sparse, distributed, diverse mapping of color terms onto color space, Part II: Hadzane, the color communication game, and color sorting

Delwin Lindsey
(lindsey.43@osu.edu)
Psychology, The Ohio State University

&

Angela Brown
(brown.112@osu.edu)
College of Optometry, The Ohio State University

In the quest to understand the processes governing the evolution of color lexicons, it is instructive to consider languages at the earliest stage of color term evolution. Hadzane, spoken in Tanzania, is a language isolate, and is among the oldest continually spoken languages on earth. Just as throughout the World Color Survey, color naming in Hadzane is highly diverse across informants. Strikingly, the Hadza use only three high-consensus color terms; the commonest response in our data set was “don’t know.” However, the Hadzane-speaking community uses a pooled lexicon containing words for most of the color categories found in English and other world languages. We discuss this result in the context of two methodologies that we have developed in our lab over the last several years. First, we performed an information-theoretic analysis of color communication, and followed up with experiments where Somali- and English-speaking players actually played the color communication game. And second, we had subjects perform a non-lexical color sorting protocol, which allowed us to examine the color categories defined with, and without, the color terms in an informant’s language.

Our work suggests that the naming of colors has always been linked to a pan-human sensory experience of color. This link between color terms and color experience, however, is only weakly related to the structure of Hering’s color space. At first, colors were named using idiosyncratic, low-consensus labels, and the representation of color in language was sparse and distributed across the language community. Modern color lexicons evolved as individuals communicated, teaching each other common terms for the colors that were important to the society, until everyone in the community used a common color lexicon that was capable of mediating better color communication.

(9) When does ‘Bright’ mean ‘Prototypal’? Color-term modifiers in five European languages, examined with color-survey data

David Bimler

(D.Bimler@massey.ac.nz)

Massey University, Australia

&

Mari Uusküla

(muuskyla@tlu.ee)

Linguistics and Translation Studies
Tallinn University, Estonia

The Berlin-Kay program belongs within a color-research tradition of determining what members of a language community collectively mean when they use a given color term (its denotata) by eliciting unconstrained descriptions of standardized color stimuli, e.g. Munsell tiles. Here we extend that approach to verbal modifiers and qualifiers. In a multi-language corpus of color-naming data, elicited with 65 ColorAid-Corporation tiles, a significant minority of Estonian respondents described one or more tiles with qualified terms of the form ‘ere-X’ or ‘erk-X’. Finnish, Hungarian, Latvian and Lithuanian respondents used comparable forms ‘kirkkaan-X’ or ‘kirkas-X’; ‘világos-X’; ‘spilgti-X’; and ‘ryški-X’ or ‘ryškiai-X’ respectively. These have been glossed into English as ‘bright (or vivid) X’, but the translation of these particles into the precisely defined terms of colorimetry is not straightforward. We addressed the question empirically by examining the denotata of each modified term, treating its uses as a distribution when the tiles are located as points in a metric color space, and determining the distribution’s centroid. These were compared with the distributions of each ‘X’ per se. We argue that the unmodified term-centroids can be identified with the focus of X: the within-language consensus about the ideal, most prototypal examplar of that term. In some cases ‘bright-X’ shares its centroid with ‘X’ although the distribution is more focused, allowing the modifier to be interpreted as ‘bright’ in the sense of ‘intense’ or ‘very’; in other cases the centroids differ significantly, implying that ‘bright’ here has the sense of ‘saturated but lighter / darker’.

(10) Color naming and cognition in computational perspective

Terry Regier

(terry.regier@berkeley.edu)

Linguistics
University of California, Berkeley

Why do languages have the color categories they do, and what do those categories reveal about cognition and communication? Color categories vary widely across languages, but this variation is constrained. I will argue that this pattern reflects a range of language-specific solutions to a universal functional challenge: that of communicating precisely while using minimal cognitive resources. I will present a general computational framework that instantiates this idea, and will show how that framework accounts for cross-language variation in color naming. I will then address the Sapir-Whorf hypothesis - the claim that such language-specific categories in turn shape cognition. I will argue that viewing this hypothesis through the lens of probabilistic inference has the potential to resolve two sources of controversy: the challenge this hypothesis apparently poses to the widespread assumption of a universal groundwork for cognition, and the fact that some findings supporting the hypothesis do not always replicate reliably.

(11) Quantifying diachronic change in category meaning using evolutionary models of learning

Kirbi Joe
(joek@uci.edu)
Institute for Mathematical Behavioral Sciences
University of California, Irvine

Recent work in color categorization has used agent-based modeling to develop an language-learning dynamic to further evolve existing color naming categorizations to a stable equilibrium. This simulation approach has been shown to be a suitable representation of the natural evolutionary processes of color categories. A natural question to ask, then, is: how does the color naming system of an agent population change as a result of this simulated evolution? In an effort to measure this diachronic shifting of color categories, we develop two measures called *category strength* and *category agreement*. Together, these two measures indicate, for a single category, the direction of change (i.e. whether a category is becoming more contiguous or more disperse) and whether the observed change is salient and meaningful to the population as a whole. We simulated the evolution of 108 linguistic communities from the World Color Survey (2009) to see if these two measures would yield any cross-cultural variation. Preliminary results show that most languages exhibit an increasing trend in both category strength and agreement over time, indicating that the learning dynamic causes agents to form cohesive categories and populations to converge on shared meaning for all its categories (i.e. agents assign the same color terms to very similar subsets of stimuli). Further analysis comparing the magnitudes of change in category strength and analysis over time hopes to provide insight into the evolutionary path of these linguistic societies' naming conventions and where they fit relative to the evolutionary pattern detailed in Berlin and Kay's *Basic Color Terms: Their Universality and Evolution* (1969).

(12) Mechanisms of color, uncovered by neuroscience and language

Bevil Conway
(bevil@nih.gov)
National Institutes of Health, Bethesda, MD

What is color for and how are color operations implemented in the brain? In my talk I will draw upon experiments using neurophysiological recording, imaging, color naming, and object color statistics to develop the idea that color comes about through the activity of a richly interconnected network spanning all levels of visual processing from retina to frontal cortex. I will argue that the network serves as a trainable system that enables the rapid detection of objects of likely behavioral relevance. I will present data suggesting that the neural implementation depends on a multi-stage network that encodes color initially through an opponent process and ultimately as a uniform representation of color space at a mid-level in visual processing. This network provides input to a series of stages associated with object vision and decision making that are responsible for decoding color signals and guiding action. The provocative theme of my talk, which will hopefully spark some productive discussion, is the idea that color categories reflect an emergent property of the nervous system arising from the needs we place on it, rather than a constraint on how color is encoded.

(13) Color categories and the perceptual representation of color

Michael A. Webster

(mwebster@unr.edu)

Cognitive Sciences

University of Nevada, Reno

It is common to represent color in terms of a two- or three-dimensional space in which the stimuli are defined by their coordinates along two or three privileged axes. But to what extent is the perceptual representation of color consistent with a "space"? We have explored this representation by examining the relationships between different color categories. Untrained individuals have little intuition about the metrical relationships between different colors, and for example are poor at judging the complement of a hue or the average of a set of colors. Between observers, differences in color categories are large and reliable. These differences show little dependence on the spectral sensitivity of the observer, and thus may be unrelated to physiological variations between or within observers. Moreover, the inter-observer variations in color appearance are uncorrelated across categories, suggesting that these differences do not reflect variations in a set of underlying global dimensions. Our results instead suggest that color categories may reflect qualitative rather than quantitative differences – that different hues are treated more like different objects than different vectors – and that these "objects" are defined or learned independently. This perceptual representation is fundamentally different from other low-dimensional representations such as the direction of motion.