Report

Current Biology

Hunter-Gatherer Color Naming Provides New Insight into the Evolution of Color Terms

Highlights

- Hadza hunter-gatherers use high-consensus color terms for black, white, and red
- Other Hadza color terms are low consensus, and "don't know" is commonly used
- Each Hadza names his/her own subset of the color categories of world languages
- A complete color lexicon is distributed across the Hadzane-speaking community

Authors

Delwin T. Lindsey, Angela M. Brown, David H. Brainard, Coren L. Apicella

Correspondence

lindsey.43@osu.edu (D.T.L.),
capicella@psych.upenn.edu (C.L.A.)

In Brief

Lindsey et al. show that color naming by Hadza hunter-gatherers, while individually idiosyncratic, is remarkably structured across the language community: Hadza share few color terms and often respond "don't know" when naming colors. Yet collectively, Hadzane color terms represent most color categories found in English and other world languages.

Lindsey et al., 2015, Current Biology *25*, 2441–2446
September 21, 2015
http://dx.doi.org/10.1016/j.cub.2015.08.006

Current Biology
Report

Hunter-Gatherer Color Naming Provides New Insight into the Evolution of Color Terms

Delwin T. Lindsey,[1,2,*] Angela M. Brown,[2] David H. Brainard,[3] and Coren L. Apicella[3,*]
[1]Department of Psychology, Ohio State University, Mansfield, OH 44906, USA
[2]College of Optometry, Ohio State University, Columbus, OH 43210, USA
[3]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA
*Correspondence: lindsey.43@osu.edu (D.T.L.), capicella@psych.upenn.edu (C.L.A.)
http://dx.doi.org/10.1016/j.cub.2015.08.006

SUMMARY

Most people name the myriad colors in the environment using between two and about a dozen color terms [1], with great variation within and between languages [2]. Investigators generally agree that color lexicons evolve from fewer terms to more terms, as technology advances and color communication becomes increasingly important [3]. However, little is understood about the color naming systems at the least technologically advanced end of the continuum. The Hadza people of Tanzania are nomadic hunter-gatherers who live a subsistence lifestyle that was common before the advent of agriculture (see Supplemental Experimental Procedures, section I; [4]), suggesting that the Hadzane language should be at an early stage of color lexicon evolution. When Hadza, Somali, and US informants named 23 color samples, Hadza informants named only the black, white, and red samples with perfect consensus. Otherwise, they used low-consensus terms or responded "don't know." However, even low-consensus color terms grouped test colors into lexical categories that aligned with those found in other world languages [5]. Furthermore, information-theoretic analysis showed that color communication efficiency within the Hadza, Somali, and US language communities falls on the same continuum as other world languages. Thus, the structure of color categories is in place in Hadzane, even though words for many of the categories are not in general use. These results suggest that even very simple color lexicons include precursors of many color categories but that these categories are initially represented in a diverse and distributed fashion.

RESULTS

Responses of 55 Hadza informants to each of 23 Munsell color samples are shown in Figure 1. To put the Hadza data in context, we also show results from 48 monolingual Somali immigrants living in Columbus, Ohio, and 43 university undergraduates who spoke US English as their first language. All informants named each color sample, one at a time, with a single color term, or else responded "don't know" (DK). Explicitly allowing DK was a crucial methodological difference between this study and much [6], but not all [7], of the previous work that involved naming individual colors.

The Distribution of Color Terms across the Stimulus Set

The first central result in the Hadzane dataset was that only black, white, and red (BWR) samples were named with perfect consensus (Figures 1B and 1E). Non-BWR color terms were never used by more than 60% of informants, and many informants used idiosyncratic color terms (column "O" in Figure 1B). The sparse distribution of non-BWR color terms across the stimulus set was different for different Hadza informants (Figure 2A), with each informant naming his or her own subset of the colors. DK was the most common response overall (56.6% of non-BWR trials; "DK" in Figures 1B and 1E; Supplemental Experimental Procedures, section III, part D). Based on taxonomies proposed by others [1, 2, 5, 7], these results suggest that Hadzane is at a very early stage of color term evolution.

In contrast to the Hadza, many Somali informants named the colors of all 23 samples (DK = 12.8% of non-BWR trials), as did almost all US informants (DK = 0.56% of non-BWR trials) (Figures 1C, 1D, 2B, and 2C, respectively). Pairs of Somali informants used the same color term to name a given sample more often than Hadza informants did, but the Somalis reached 100% consensus only for the white sample (Figure 1F). In contrast, US informants showed 100% consensus for all 11 of the classic basic color terms of Berlin and Kay [1](Figures 1D and 1G), and even for lower-consensus terms, agreement was high in English as compared to Hadzane and Somali. The inventory of terms used by 80% or more informants was 5 for Somali and 11 for English, compared to 3 for Hadzane.

Hadza Color Terms and the Color Terms of the WCS

The second central finding was that even the low-consensus Hadzane color terms generally grouped together colors falling within the universal color categories: black, white, red, yellow-or-orange, green, blue, green-or-blue, brown, pink, purple, and gray [5]. Informally, many Hadzane color terms are easily translated into Somali or English. For example, in Figures 1B–1D, color terms "d" in Hadzane, "e" in Somali, and "d" in English clearly mean "red," and Hadzane term "e" is easily translated to "yellow." More formally, we define a "lexical color group" (LCG) to be a set of colors receiving the same term by an informant, and we define 11 partially overlapping universal color

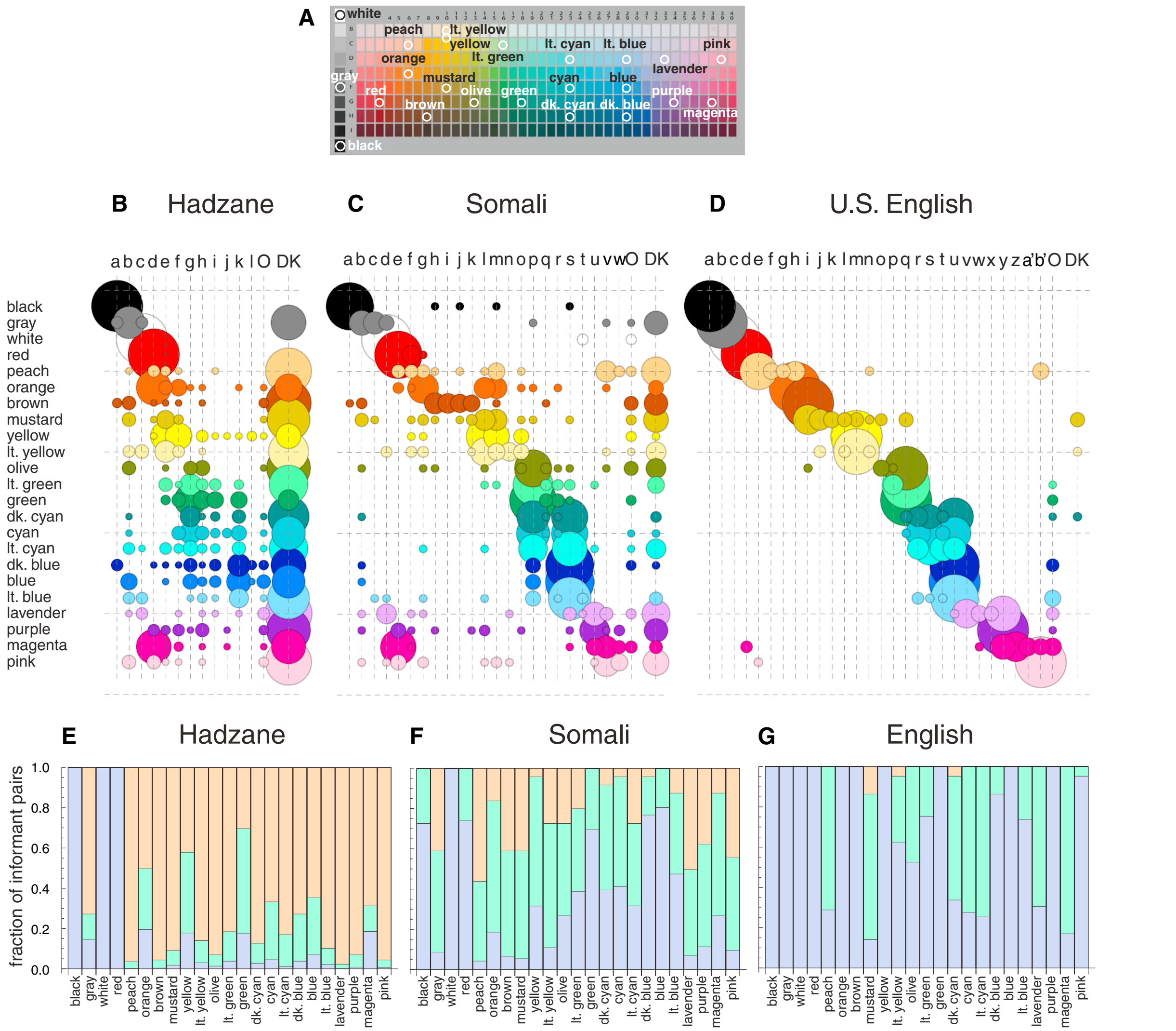

Figure 1. Stimuli and Group Data

(A) 23 color samples, shown within the 330-sample World Color Survey (WCS; [5]) palette (see also Figure S1).

(B–D) Group color naming data. Disk sizes show the fraction of informants who provided each color term (letters at top; key to the color terms in Table S1) for each sample (disk colors; descriptively named on the left). "O" indicates "other" terms used by ≤2 informants; DK, don't know.

(E–G) Fraction of pairs of informants where both members of the pair provided the same non-DK term (blue), different non-DK terms (green), or at least one DK response (orange), for the samples listed on the x axis.

categories from our analysis [5] of the World Color Survey (WCS), a database of color terms for 330 Munsell colors provided by 2,616 informants of 110 unwritten world languages [6]. 74% of the Hadzane LCGs were wholly contained within the boundaries of single universal color categories (colored polygons in Figures 2A and 2D), whereas 26% spanned more than one such category (gray polygons in Figures 2A and 2D; Supplemental Experimental Procedures, section III, parts A–C). This result was unlikely to have occurred by chance ($p < 10^{-5}$). Thus, although Hadzane color naming is sparse, it is also highly organized. Results for Somali (Figures 2B and 2E) and English (Figures 2C and 2F) show similar patterns. Moreover, the hue alignment of all three languages' LCGs with the universal WCS categories is optimal: cyclic shifting of the WCS categories along the hue dimension [8] always reduced the proportion of LCGs contained in the shifted categories (see Supplemental Experimental Procedures, section III, part B.1.H and Figures 2G–2I).

Color Term Diversity and Color Communication

A third feature of the Hadzane dataset was the diversity of responses to a given color and the diversity of colors named by a given color term. Yellow and blue-to-green regions of the color chart (rows of bubbles in Figure 1B) were named with multiple essentially synonymous terms, and single terms were often used to name many colors (columns of bubbles in Figure 1B). Does this diversity, along with the high frequency of DK responses, imply that color communication among speakers of Hadzane is poor compared to color communication in other languages?

To evaluate this issue, we performed an information-theoretic analysis by simulating a color communication game in which a

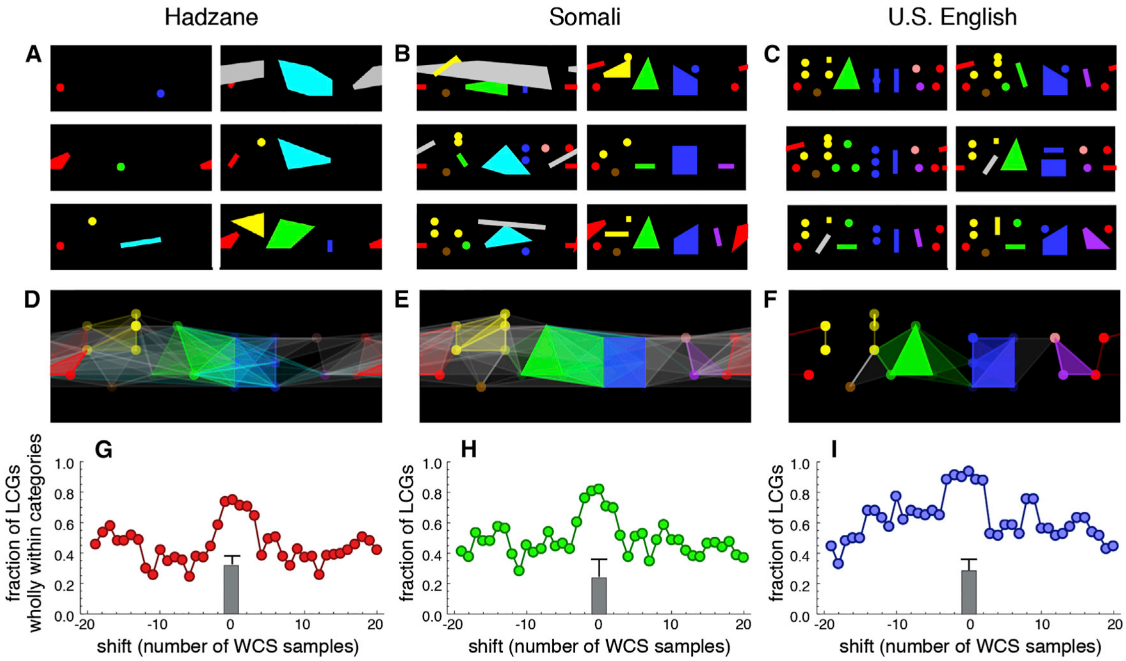

Figure 2. Color Naming Data

(A–C) Representative individual datasets for Hadzane, Somali, and US English, including individuals who used many terms, few terms, and terms that straddled the color category boundaries. Color terms are shown within the palette from Figure 1A. Single named colors are shown as colored dots. Multiple colors receiving the same color term (lexical color groups [LCGs]) are shown as lines or polygons (colors: LCGs falling wholly within the universal color categories of the WCS, color key in Figures 1B–1D; gray: LCGs straddling multiple color categories). The proportion of non-BWR responses that were DK differed across all three languages (all three pairwise comparisons: $p < 0.0005$ on a one-way ANOVA, after Tukey HSD correction for multiple comparisons).

(D–F) Data collated within languages; opacity shows the level of consensus.

(G–I) Fraction of chromatic LCGs contained within universal color categories. Data points show LCGs within cyclically shifted chromatic color categories. At zero shift, the Hadza data are below the Somali data ($p = 0.017$) and the US data ($p < 0.0005$), but the Somali and US data did not differ significantly ($p = 0.228$) on a one-way ANOVA, after Tukey HSD correction. Bars show the prediction of our permutation test at zero shift. Error bars are +95% confidence intervals (see Figure S2 for complete analysis of chromatic and achromatic data and simulations).

"sender"(S) views a set of color samples and names randomly chosen samples in turn, based on his or her color idiolect. A simulated "receiver" (R) attempts to identify each named sample from his or her duplicate set of color samples, based on his or her own color idiolect. Mutual Information ([9]; Supplemental Experimental Procedures, section IV) is a measure of the amount by which S's utterance can improve R's chances of identifying S's color sample selection. We determined the empirical group mutual information (GMI) by aggregating the results of games played by all pairwise permutations of a language's informants. See also [10], and see Supplemental Experimental Procedures, section IV, part A for the details of our analysis and related information-theoretic approaches to understanding color naming and categorization.

For a given set of colors, a language's GMI is limited by the number of color terms used by its informants and also by the level of consensus among informants in the deployment of these terms across colors. To distinguish these two effects, we also calculated optimum GMI, which assumes perfect consensus. Optimum GMI was calculated for representative languages from the WCS with 2–8 high-frequency color terms ([8]; Supplemental Experimental Procedures, section IV), plus English. Perfect consensus was assured by assigning, for each language, the same prototypic color lexicon to both S and R. Optimum GMI (purple dots in Figure 3A) always fell well above empirical GMI (black dots in Figure 3A), indicating that level of consensus is an important limitation on the GMI of color lexicons of all sizes.

The results of the color communication game analysis unite the present datasets with those of the WCS, even though the data collection methods differed. The empirical GMI of our Hadza informants fell close to the regression line fitted to the WCS data, so the Hadzane GMI is typical for a language with three frequently used color terms. A similar result was obtained for Somali (with five color terms), while English (with 11 color terms) was somewhat above the extrapolated regression line. In the Hadzane dataset, the high frequency of DK responses, together with the low consensus for the color samples that did receive names, accounted for the low Hadza GMI, compared to the optimum value. In contrast, the low GMI in the WCS languages occurred both because of the low overall consensus in most languages and because some WCS informants responded seemingly randomly to some of the color samples. It seems likely that WCS informants would also have used DK more often, had they not been discouraged from doing so. In any case, the low consensus shown in all the languages in Figure 3A agrees with the striking diversity we have previously reported in the overall patterns of color naming, within WCS languages [2], which we

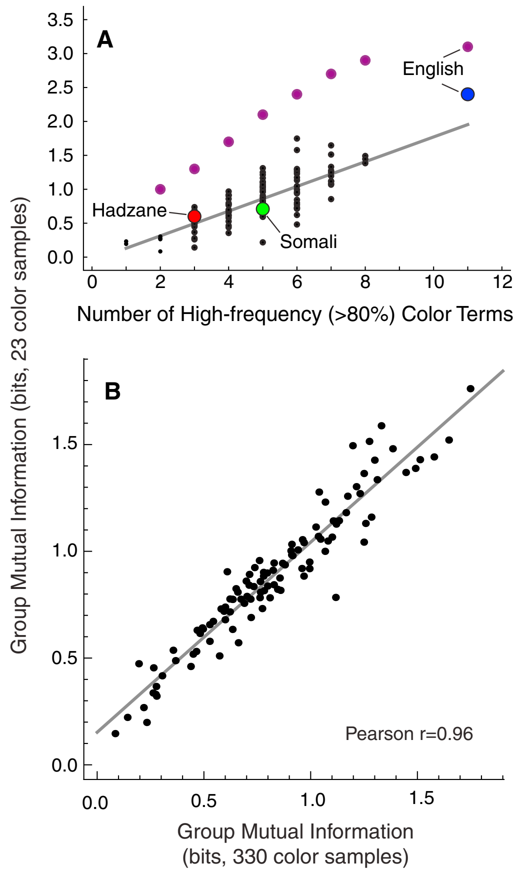

Figure 3. Group Mutual Information

(A) Group mutual information (GMI) in 110 WCS languages (black dots), Hadzane (red disk), Somali (green disk), and English (blue disk). Purple dots indicate optimum GMI for representative languages (see also Supplemental Experimental Procedures, section IV, Figure S3, and Table S2 for a worked example of mutual information).

(B) GMI for the 23 samples used here, compared to GMI for the full 330-sample set in the WCS.

refer to as "motifs." The WCS GMI results shown in Figure 3A were based on our 23-sample stimulus subset, but GMI analysis of the whole 330-sample WCS dataset yielded highly correlated results (Figure 3B; Supplemental Experimental Procedures, section IV).

Summary and Relation to Controversies in the Field of Color Naming

Our results show that Hadza color naming is sparse, distributed, and diverse. It is sparse in that non-BWR terms were not produced by many informants and were used with low consensus, while DK was used frequently. It is distributed in that a systematic lexical representation of non-BWR colors was distributed across the idiolects of the Hadza-speaking language community. It is diverse in that the dataset included many terms for each named color. Each informant used his or her own color terms for his or her own subset of samples but tended to unite these colors into LCGs according to universally observed color categories. As a group, the Hadzane lexicon carries information about color about as effectively as that of any other group of people who use only three high-frequency color terms. For individual informants, the representation of color was incomplete, but collectively, Hadza usage of color terms showed the beginnings of a more complete color naming system.

These three results hold even as we consider the possible origins of the non-BWR terms used in modern Hadzane. Hadzane is a language isolate [11], but the Hadza people have interacted with their Cushitic, Nilotic, and Bantu language speaking neighbors for hundreds of years and probably much longer [11], with increasing contact with Swahili speakers since the 1950s [12]. Such interactions are, in general, important for language change [13] and have almost surely introduced loanwords into the Hadzane color lexicon [14] (Supplemental Experimental Procedures, section III, part D). For example, in the present dataset, the term *buluwa* (term "k" in Figure 1B) is undoubtedly a loanword, while *manjanowa* (term "e" in Figure 1B) originated in neighboring Bantu languages (Swahili or Nikamba [1, 15]). Regardless of their origins, these loanwords are now used alongside other Hadzane color terms that name similar ranges of color samples. All non-BWR color terms were similarly sparse, diverse, and distributed in the idiolects of the Hadza informants who used loanwords and those who did not. Importantly, the distribution of LCGs within the WCS universal categories was also similar for Hadza informants who did and did not use loanwords. Thus, the essential features of our results are not driven by the data from individuals who use loanwords. Moreover, demographic analyses show that the use of loanwords was not concentrated with respect to informant age, education level, or exposure to outside cultures through tourism and village life, as might be expected of incompletely adopted loanwords of recent origin (see Supplemental Experimental Procedures, section III, part D for further discussion of loanwords).

Color naming and color communication among Hadza informants provide new insight into the evolution of lexical representations of color. Although there are major controversies in this field, investigators generally agree that color naming systems tend to evolve over time from simpler to more complex. According to this view, early color lexicons spoken in preindustrial societies included few terms, and more terms were added over time, up to a maximum of about 11 terms for basic color categories. Consistent with this view, Hadza society is technologically very simple, and the Hadzane dataset shows only three high-frequency color terms. However, not necessarily predicted by that view, the Hadzane dataset also contains a larger set of less common color terms that collectively name many of the color categories found in the WCS and in English [16]. These non-BWR terms are distributed across the idiolects of the members of the language community rather than being fully represented within a single, unified lexicon. Furthermore, both the Hadza and the WCS GMI data fall far short of the optimal GMI that could be achieved for a given number of high-consensus color terms. These two results suggest the insight that the lexical representation of color within preindustrial societies such as the Hadza does not evolve directly from simple, high-consensus representations to more complex, high-consensus representations. Instead, consensus for new terms is initially low, and color term evolution probably involves not only the introduction of new terms but also a gradual process through which consensus develops about which terms are preferred for which colors. This consensus builds through the interaction of multiple lexical representations that coexist across individuals within a culture [2]. Even in our English dataset, the variety of terms deployed for small numbers of samples (e.g., mustard, magenta) suggests a distributed and perhaps evolving color lexicon for samples at the boundaries of well-established categories [16]. Interestingly, there is some evidence in our data that close familial contact influences Hadza color idiolect: the frequency with which an individual used DK was related to his or her spouse's frequency of DK, but not to the frequencies of DK of other camp members. The latter may be because of the fluidity of camp membership (see Supplemental Experimental Procedures, section III, part E).

The data presented here challenge all three major accounts about the early stages of color term evolution. The universalist hypothesis [1, 6] holds that the earliest color naming systems contained only two color terms, which were sufficient to name all visible colors. That is, even the early color naming systems were complete. As the color systems evolved, color space was divided into smaller and smaller parcels, eventually reaching about 11 named color categories, which were based on an innate, universal representation of color [3]. Our results are consistent with the view that certain aspects of the representation of color are innate and universal, and the arrival of loanwords from other languages suggests that the Hadzane color lexicon is evolving. However, our results are probably also consistent with other, non-innate explanations of the universality of color categories as our data do not speak directly to the origin of the structure we observe. More importantly, the sparseness of individual Hadzas' idiolects is at variance with the idea that named color categories partition color space exhaustively. Moreover, the universalist view posits a color lexicon that is shared across informants at each stage of color term evolution, a view that is inconsistent with the diversity of color naming we observe in Hadzane and, more generally, the WCS [2].

The linguistic relativity hypothesis is that "the language we speak affects the way we think" [17]. As applied to color, this is generally taken to mean that there is a color lexicon that is shared, with high consensus, across the speakers of a language, and that there is a direct, causal link between that color lexicon and the organization of color in the minds of the speakers of that language. Under this view, color lexicons and their corresponding named color categories can vary freely across cultures, and it is the language's color lexicon that determines the locations of the specific boundaries between the colors that partition color space [18]. Perhaps a relativist explanation for our results could be found. However, linguistic relativity does not easily explain the universality of the color categories within the WCS and within the distributed structure of the Hadza color naming system or the fact that the non-BWR categories occur without the corresponding presence of high-consensus color terms that name them. In addition, like the partitioning aspect of the universalist view, the boundary partition hypothesis of the relativist view is not easily squared with the high frequency of DK in the Hadza dataset.

The emergence hypothesis [7], which is also a relativist view, holds that the lexical representations of color are sparse in the simplest color lexicons because they are metaphorically linked to items in the culture or the natural environment whose colors span limited regions of color space. Consistent with this view, Hadza informants often responded DK to colors in non-BWR regions of color space. However, in spite of sharing the same environment and culture, the Hadza community has not arrived at a shared lexical representation of the sparsely distributed non-BWR color terms. Instead, color terms are distributed with low consensus across the Hadzane-speaking language community. Moreover, we find excellent alignment of Hadza LCGs and the universal WCS color categories, suggesting important influences on color naming that transcend those supplied by local culture.

DISCUSSION

Some aspects of the present datasets are also apparent in previous work: the existence of universal color categories is central to Berlin and Kay's seminal work [1, 6]; other investigators have reported that color naming is sparse [7, 19], and the diversity of color term usage across individuals [3, 20] and the distributed representation of color categories across the lexicons of many informants are implicit in the existence of multiple motifs occurring within most languages of the WCS [2]. However, sparseness, diversity, and the distributed representation of color across individuals, embraced in a unified way, not only provide new insight into how color terms might evolve but also suggest a simple framework for understanding how adjectives related to other object properties come into use. In this view, there were commonalities in human sensory experience that initially lacked common lexical representation. Standard labels for the properties of this common sensory experience were initially unknown to many people (sparseness), and many non-standard labels, whether invented locally or acquired as loanwords, were used idiosyncratically or on the spur of the moment (diversity). However, general knowledge of the properties of objects was in place and was distributed over the lexicon of the language community (distribution). Modern vocabularies came into being as individuals communicated, over a period of time, teaching to others the terms and associated categories that they knew and learning from others the terms and categories they did not know, until every person's vocabulary came increasingly to contain distinct high-consensus words for all the properties of objects known to the culture. This increasing vocabulary improved the ability of people to communicate about the properties of objects in their world. The Hadza provide a striking example of an early phase of this process.

EXPERIMENTAL PROCEDURES

Hadza informants were tested *en scene* in Tanzania by co-author C.L.A. Co-authors D.T.L. and A.M.B. tested comparison groups of monolingual Somali immigrants living in Columbus, Ohio, and Ohio State University undergraduates who spoke US English as their first language. All test procedures followed a protocol approved by University of Pennsylvania and Ohio State University Institutional Review Boards. Participants were determined to be color normal by the Richmond HRR (Hardy, Rand, Rittler) pseudo-isochromatic plate test [21]. Each informant provided a single color term in his or her native language—or DK—for each of 23 Munsell samples presented one at a time, in fixed order. See Supplemental Experimental Procedures, section I for further information on the participants and the testing procedure. The Munsell colors were chosen to provide a representative sample of colors spanning color space. See Figure S1, Table S1, and Supplemental Experimental Procedures, section II for further details on the stimuli.

Statistical analysis of each experimental group's chromatic and achromatic LCGs was based on permutation tests ([22]; Supplemental Experimental Procedures, section III). We estimated the likelihood that the numbers of LCGs falling wholly within universal categories were greater than those expected by chance by repeated simulations, in which the colors associated with each informant's LCGs were randomly reassigned, without replacement, while preserving the overall numbers of the observed LCGs and the number of color samples they contained. Bars shown in Figures 2G–2I show the mean and 95% confidence intervals for 100,000 simulations for each experimental group. See Figure S2 and Supplemental Experimental Procedures, section III, parts A–C for further details.

Calculations of GMI were based on the classic equation for mutual information [9]:

$$GMI(C_S; C_R) = \sum_{s,r} p_N(s,r)\log_2\left(\frac{p_N(s,r)}{p_N(s)p_N(r)}\right).$$

GMI is the reduction in uncertainty in the identification of the test samples C_R by the receiver in our communication game, given the utterances by the sender associated with the samples C_S, aggregated across a language group's color naming responses. Here, $p_N(s,r)$ is a 23 × 23 (or 330 × 330) matrix of the joint probability distribution on the random variables C_S and C_R, averaged across the results of games played by all N pairwise permutations of a language's informants. GMI assumes that sender S samples the 23 colors (or 330 colors) in the test set randomly, with replacement, and with equal frequency. The entries in the matrix giving $p_N(s,r)$ are the probabilities associated with the samples that S names and the samples selected by R in response to S's names. The probabilities $p_N(r)$ and $p_N(s)$ are the marginal distributions on C_R and C_S, respectively. See Supplemental Experimental Procedures, section IV for further details.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2015.08.006.

ACKNOWLEDGMENTS

This project was supported by grants NSF BCS-1152841 to D.T.L. and NIH RO1-EY10016 to D.H.B. and by the University of Pennsylvania. We thank Audax Mabulla and Ibrahim Mabulla for logistical and field assistance in Tanzania; Charles Endeko for interpreter services (Hadzane, Swahili); Abdirizak Isse for interpreter services (Somali, English); and G.L. Mitchell for statistical advice. We thank Kirk Miller for linguistic advice about the Hadzane language.

Received: April 4, 2015
Revised: June 25, 2015
Accepted: August 6, 2015
Published: September 10, 2015

REFERENCES

1. Berlin, B., and Kay, P. (1969). Basic Color Terms: Their Universality and Evolution (Los Angeles: University of California Press).
2. Lindsey, D.T., and Brown, A.M. (2009). World Color Survey color naming reveals universal motifs and their within-language diversity. Proc. Natl. Acad. Sci. USA *106*, 19785–19790.
3. Kay, P., and Maffi, L. (1999). Color appearance and the emergence and evolution of basic color lexicons. Am. Anthropol. *101*, 743–760.
4. Marlowe, F. (2010). The Hadza: Hunter-gatherers of Tanzania (University of California Press).
5. Lindsey, D.T., and Brown, A.M. (2006). Universality of color names. Proc. Natl. Acad. Sci. USA *103*, 16608–16613.
6. Kay, P., Berlin, B., Maffi, L., Merrifield, W.R., and Cook, R.S. (2009). The World Color Survey (Stanford: CSLI).
7. Levinson, S.C. (2000). Yélî Dnye and the theory of basic color terms. J. Linguist. Anthropol. *10*, 3–55.
8. Regier, T., Kay, P., and Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. Proc. Natl. Acad. Sci. USA *104*, 1436–1441.
9. Cover, T.M., and Thomas, J.A. (1991). Elements of Information Theory (Hoboken: Wiley).
10. Baddeley, R., and Attewell, D. (2009). The relationship between language and the environment: information theory shows why we have only three lightness terms. Psychol. Sci. *20*, 1100–1107.
11. Marlowe, F. (2002). Why the Hadza are still hunter-gatherers. In Ethnicity, Hunter-gatherers, and the Other, S. Kent, ed. (Washington: Smithsonian Institution Press), pp. 247–281.
12. Woodburn, J. (1979). Minimal politics: the politcal organization of the Hadza of North Tanzania. In Politics in Leadership: A Comparitive Perspective, W.A. Shack, and P.S. Cohen, eds. (Oxford: Clarendon Press), pp. 244–267.
13. McMahon, A.M.S. (1994). Understanding Language Change (Cambridge: Cambridge University Press).
14. Lusekelo, A. (2015). The consequences of the contacts between Bantu and non-Bantu languages around Lake Eyasi in Northern Tanzania. International Journal of Society, Culture & Language *3*, 62–75.
15. Yukawa, Y. (1989). A Classified Vocabulary of the Nilamba Language (Tokyo: Bikosha Printing Co., Ltd.). Cited in The Tower of Babel: The Global Lexostatistical Database, G. Starostin (2006–2013), http://starling.rinet.ru/cgi-bin/response.cgi?root=new100&morpho=0&basename=new100/bco/bnf&limit=-1.
16. Lindsey, D.T., and Brown, A.M. (2014). The color lexicon of American English. J. Vis. *14*, 17.
17. Lucy, J.A. (1997). Linguistic relativity. Annu. Rev. Anthropol. *26*, 291–312.
18. Roberson, D., Davies, I., and Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. J. Exp. Psychol. Gen. *129*, 369–398.
19. Hill, C. (2011). Named and unnamed spaces: color, kin, and the environment in Umpila. Senses & Society *6*, 57–67.
20. Webster, M.A., and Kay, P. (2005). Variations in color naming within and across populations. Behav. Brain Sci. *28*, 512–513.
21. Bailey, J.E., Neitz, M., Tait, D.M., and Neitz, J. (2004). Evaluation of an updated HRR color vision test. Vis. Neurosci. *21*, 431–436.
22. Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. *15*, 1–25.

