Cognition 229 (2022) 105223

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Color technology is not necessary for rich and efficient color language

Ewelina Wnuk [a,b], Annemarie Verkerk [c], Stephen C. Levinson [d,e,*], Asifa Majid [f]

[a] Institute of English Studies, University of Warsaw, 00-681 Warsaw, Poland
[b] Department of Anthropology, University College London, London WC1H 0BW, United Kingdom
[c] Department of Language Science and Technology, Saarland University, 66123 Saarbrücken, Germany
[d] Radboud University, 6500 HD Nijmegen, the Netherlands
[e] Language & Cognition Department, Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, the Netherlands
[f] Department of Experimental Psychology, University of Oxford, Oxford OX2 6GG, United Kingdom

ARTICLE INFO

Keywords:
Basic color terms
Secondary color terms
Coding efficiency
Cross-cultural codability
Color technology
Color salience

ABSTRACT

The evolution of basic color terms in language is claimed to be stimulated by technological development, involving technological control of color or exposure to artificially colored objects. Accordingly, technologically “simple” non-industrialized societies are expected to have poor lexicalization of color, i.e., only rudimentary lexica of 2, 3 or 4 basic color terms, with unnamed gaps in the color space. While it may indeed be the case that technology stimulates lexical growth of color terms, it is sometimes considered a sine qua non for color salience and lexicalization. We provide novel evidence that this overlooks the role of the natural environment, and people’s engagement with the environment, in the evolution of color vocabulary. We introduce the Maniq—nomadic hunter-gatherers with no color technology, but who have a basic color lexicon of 6 or 7 terms, thus of the same order as large languages like Vietnamese and Hausa, and who routinely talk about color. We examine color language in Maniq and compare it to available data in other languages to demonstrate it has remarkably high consensual color term usage, on a par with English, and high coding efficiency. This shows colors can matter even for non-industrialized societies, suggesting technology is not necessary for color language. Instead, factors such as perceptual prominence of color in natural environments, its practical usefulness across communicative contexts, and symbolic importance can all stimulate elaboration of color language.

1. Introduction

The idea that technological development matters for color term evolution was first spelled out by Berlin and Kay (1969), and later elaborated by Kay and Maffi (1999). According to the claim, color is highly lexicalized when it is useful for communicative purposes, i.e., to distinguish between otherwise similar objects. The pressure to lexicalize color is less strong, however, when color is not a reliable cue for distinguishing between objects, i.e., when objects differ on many attributes. In line with that, there is said to be little pressure to develop rich color lexica in non-industrialized societies since colors occurring naturally in the environment are rarely the only distinctive object feature: “Except perhaps for a few pairs of closely related species of birds or of fish, it is rare that naturally occurring objects or the artifacts of technologically simple societies are distinguishable only by color” (Kay & Maffi, 1999, p. 746).

Societies with relatively low technological complexity are generally expected to have a lexicon in the general range of 2, 3, or 4 (Casson, 1994), or up to 5 terms (Naroll, 1970). However, these predictions are to be understood in terms of a statistical tendency rather than an absolute rule and many color language researchers have been careful not to suggest that technological development is the sole determinant for acquiring terms beyond the 2, 3, 4, or 5 basic color categories. Yet, the strong focus on the technology-to-lexicalization link in the literature

* Corresponding author at: Radboud University, P.O. Box 9103, 6500 HD Nijmegen, the Netherlands
E-mail addresses: ewelina.wnuk@wn.uw.edu.pl (E. Wnuk), annemarie.verkerk@uni-saarland.de (A. Verkerk), Stephen.Levinson@mpi.nl (S.C. Levinson), asifa.majid@psy.ox.ac.uk (A. Majid).

https://doi.org/10.1016/j.cognition.2022.105223
Received 18 February 2020; Received in revised form 24 June 2022; Accepted 3 July 2022
Available online 13 September 2022

(Kay & Maffi, 1999; Levinson, 2000; Naroll, 1970), and the simultaneous discounting of naturally-colored objects (covering “quite narrow segments of the colour continuum” [Taylor, 2003, p. 13]) as possible stimulants of color term growth (Gibson et al., 2017; Kay & Maffi, 1999; Wierzbicka, 2008), have occasionally resulted in such interpretations. In recent publications and discussions it is stated that in the context of natural objects “an extensive botanical vocabulary (…) might obviate the need for color terms” (Gibson et al., 2017, p. 10789), “Without industrialization, cultures don’t need color terms” (Gibson, 2017), and “discriminating between abstract ‘colours’ (…) [is] a mental habit which makes sense in a world full of manufactured objects” (Wierzbicka, 2008, p. 414), the implication being it does not do so among natural objects. The widely cited case studies such as that of the Bellonese or the Warlpiri, who are “not interested” in color and “don’t talk (much)” about color (Kuschel & Monberg, 1974, p. 213; Wierzbicka, 2008, pp. 411, 421), reinforce this belief, magnifying the expectation that if a society lacks technological control of color or is not exposed to artificially colored objects, it has a simple color lexicon. At the same time, however, a recent cross-linguistic investigation examining 20 diverse languages failed to show a link between color technologies (use of paints and dyes, role of color in ritual, existence of color experts in the culture) and color codability (a measure taking into account number of color terms and their consensual usage) (Majid et al., 2018), casting doubt on the indispensability of technology in the development of color language.

In this paper, we challenge the idea that technological development is necessary for the emergence of color lexica. We examine color terminology among the hunting-gathering Maniq—a nomadic people who inhabit the rainforests of Southern Thailand. This group of about 300 people speaks a distinct language of their own (Maniq, also referred to as Ten’en) of the Austroasiatic language family, Aslian branch. They are culturally and linguistically related to other Orang Asli groups of the Malay Peninsula, many of which are hunter-gatherers of the Semang ethnographic cluster (Benjamin, 1985; Burenhult, Kruspe, & Dunn, 2011; Kruspe, Burenhult, & Wnuk, 2015). Highly mobile, they have simple but highly effective material culture relying on materials such as wood, palm, and bamboo, and no applied color technology such as paints or dyes (Wnuk, 2016) (see §6.1). Using Maniq linguistic and ethnographic data, we show that rich and efficient color language can emerge without technological development. We demonstrate that naturally-colored objects too have the capacity to stimulate lexical growth of color vocabulary and can provide a context in which colors are salient for speakers and useful to talk about.

In order to contextualize our findings within the general picture of color language research and the specific regional and cultural context of Thailand, we compare color naming data from Maniq to other languages using previously published data (Kay, Berlin, Maffi, Merrifield, & Cook, 2009; Majid et al., 2018; Majid & Burenhult, 2014; Majid & Kruspe, 2018) as well as newly collected data from 2 languages in close geographical and cultural context of the Maniq—Mlabri and Thai. While we have detailed ethnographic data on technologies within the Maniq setting, comparative information on color technologies (presence of dyes and paints) for many societies is not readily available. We therefore focus our comparison on hunter-gatherer groups since they are considered to be among the least technologically complex societies (Woodburn, 1982), prototypically associated with mobile settlement pattern and low importance of prestige items (Hayden, 1998; Kelly, 2013). Previous work with hunter-gatherer groups shows that their languages often have few conventionalized color terms with unnamed gaps in the color space (Hill, 2011; Lindsey, Brown, Brainard, & Apicella, 2015) and their color lexica are used with low consensus (Majid & Burenhult, 2014; Majid & Kruspe, 2018). Contrary to expectations, however, our investigation reveals that color in hunter-gatherer languages does not stand out for being especially poorly lexicalized or inefficiently coded. In reality, color language varies considerably among different hunter-gatherer groups and this variation does not seem to be easily explained by differences in color technology.

Our main focus in this paper is the Maniq color lexicon and the prediction that as a technologically “simple” non-industrialized society it should display few terms. We begin by presenting the results of color naming and focal color tasks, standard methods used for systematic elicitation of the extent and best examples of color categories, which puts Maniq in the context of other languages where we have comparable data. In §2, we present results from data collected using a set of 80 color chips and in §3 we replicate and extend this by examining a larger set of 330 color chips. In §4, we provide an interim discussion of coding efficiency for color, and in §5 we present further analysis of the Maniq basic color terms and secondary terms used outside of color naming tasks. In §6, we give further contextualization with a detailed examination of the place of color in everyday life of Maniq people before summing up our findings in the final discussion in §7.

2. Study 1: color naming with 80 color chips

We first explored color vocabulary by carrying out a standard color elicitation task using a stimulus set of 80 color chips and comparing Maniq to other languages, including several spoken by hunter-gatherer groups.

2.1. Methods

2.1.1. Participants

This study presents new data collected from 11 Maniq participants (6 female; age range 25–50) who took part in a color naming task. Focal color data were collected on a separate occasion from 10 participants (4 female; age range 25–50), eight of whom were the same speakers as in the naming task. On both occasions, we tested the maximum number of participants who were able to take part. All data were collected in 2009–2010. Later, naming data were also collected from 14 Mlabri (10 female, age range 16–62) and 10 Thai participants (7 female; age range 23–54) in 2018. The same Mlabri and Thai participants also completed the focal color task. One Mlabri participant was excluded due to failure in completing the color deficiency test. All participants were native speakers of their languages. The Maniq and the Mlabri also had a good command of the regional varieties of Thai, as well as standard Thai in the case of the Mlabri. Most Maniq had only limited contact with the Thai population, while the Mlabri were in regular contact with Thai speakers, as well as Hmong and Mien speakers. Informed consent was obtained in writing or orally as appropriate to each community.

2.1.2. Stimuli

The naming stimuli were presented in the form of a booklet with 80 standardized Munsell color chips of 20 equally spaced hues with 4 degrees of brightness at maximum saturation (Majid & Levinson, 2007). The focal color array was a single sheet with 84 color chips laid out in Munsell color space: 80 colors from the naming task and 4 additional chips with achromatic colors (Majid, 2008). The most common type of color deficiency (i.e., red-green color blindness) was tested with unlettered Ishihara plates (note that color blindness is uncommon in hunter-gatherers; Josserand, Meeussen, Majid, & Dediu, 2021). These were appropriate for testing illiterate participants.

2.1.3. Procedure

Participants were presented with colors in a fixed random order and asked in their native language ‘What color is this?’. Participants were free to use any terms or not give a response. After completing the color naming, they selected focal examples of the basic color terms identified in naming with the instruction ‘Which one is the best/real X?’. All participants also took an Ishihara color deficiency test, which required tracing of winding lines, carried out either before or after the naming task.

2.2. Results

It is predicted that languages of societies with hunter-gatherer mode of subsistence “should be at an early stage of color lexicon evolution” (Lindsey et al., 2015, p. 2441). Previous work with hunter-gatherers in comparable tasks showed relatively few basic color terms (Lindsey et al., 2015) and low codability of color, i.e., low-consensus naming (Majid & Burenhult, 2014; Majid & Kruspe, 2018), with many color chips left unnamed (Hill, 2011). If Maniq were to follow the predicted pattern, we would expect it to have a small color lexicon and display low codability. If, on the other hand, it were distinct, we would expect a relatively large number of color terms and high codability. In the latter case, we would also expect Maniq color system to display high coding efficiency relative to other languages (Lindsey et al., 2015), and be well-formed (i.e., exhibit optimal or near-optimal partitioning of color space) (Regier, Kay, & Khetarpal, 2007, 2009).

2.2.1. Codability of color

To test naming agreement, we compared Maniq to 25 languages spoken in a broad variety of socio-cultural and geographical settings. This includes the new data from Mlabri and Thai described in §2.1.1, as well as previously published color naming data using the same stimulus set and procedure from 23 additional languages (Majid et al., 2018; Majid & Burenhult, 2014; Majid & Kruspe, 2018) (see Fig. 1 for other languages). We first examined codability of color, operationalized as naming agreement measured using Simpson’s diversity index (Majid et al., 2018; Simpson, 1949). The index is calculated for each color stimulus using the formula: $D = \Sigma n(n-1)/N(N-1)$, where n = the total number of responses using that specific name and N = the total number of responses across all names. Simpson’s Diversity Index provides a summary statistic between 1 and 0 where the number of different words produced and the frequency of each word is taken into account. Accordingly, 1 indicates all participants provided the same response for a stimulus and 0 indicates each participant gave a unique response.

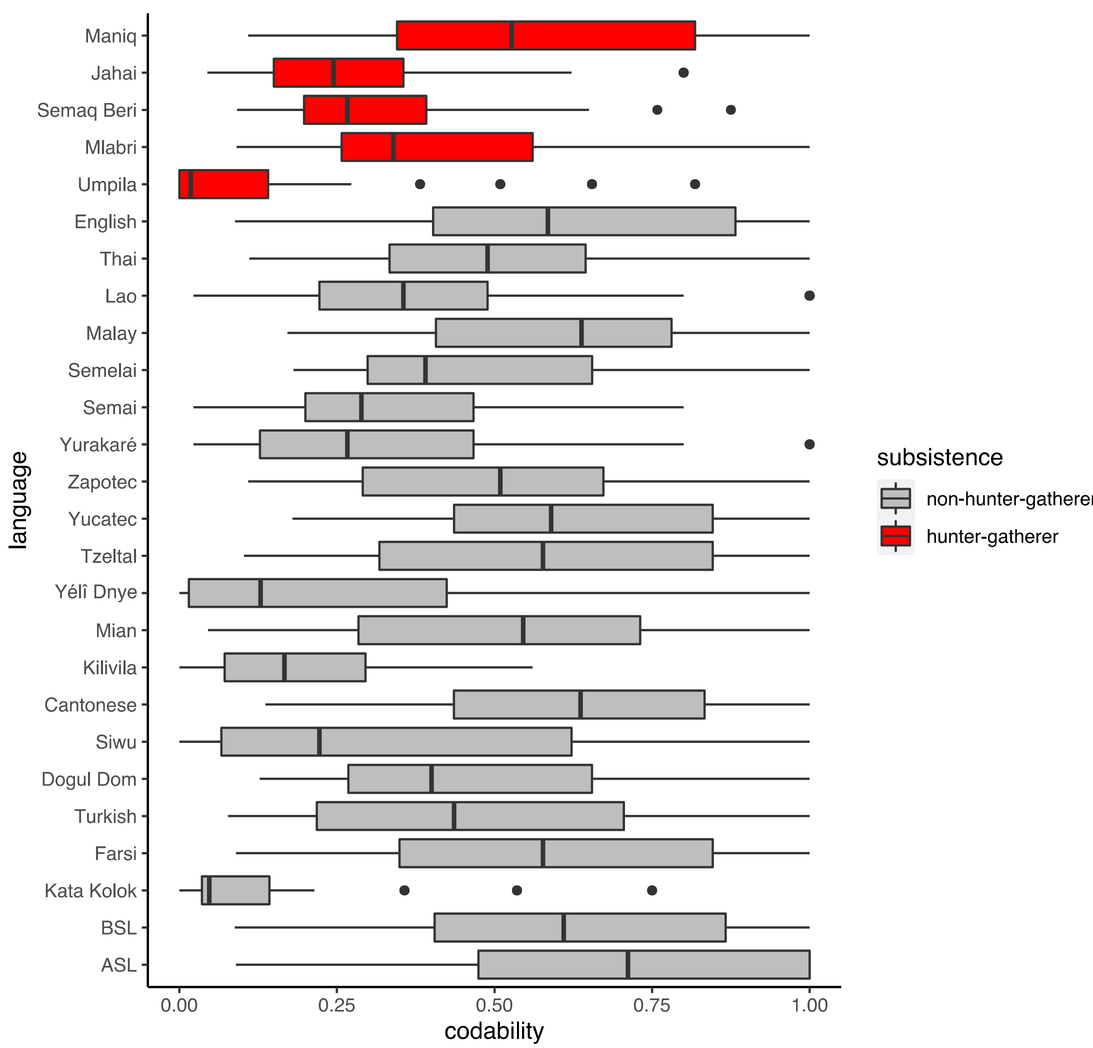

Fig. 1. The boxplot for Maniq, Thai, Mlabri, Jahai (Majid & Burenhult, 2014), Semaq Beri and Semelai (Majid & Kruspe, 2018), and 20 additional languages (Majid et al., 2018). Each language portrays Simpson’s Diversity Index for each color chip in the naming experiment: a value of 1 indicates unanimity in naming responses for a specific color chip, while 0 indicates each participant provided a unique response for a color chip. The boxplots depict the median Simpson’s Diversity scores across color chips for each language, as well as the first and third quartiles, and the minimum and maximum scores. Hunter-gatherers do not differ in codability from non-hunter-gatherers.

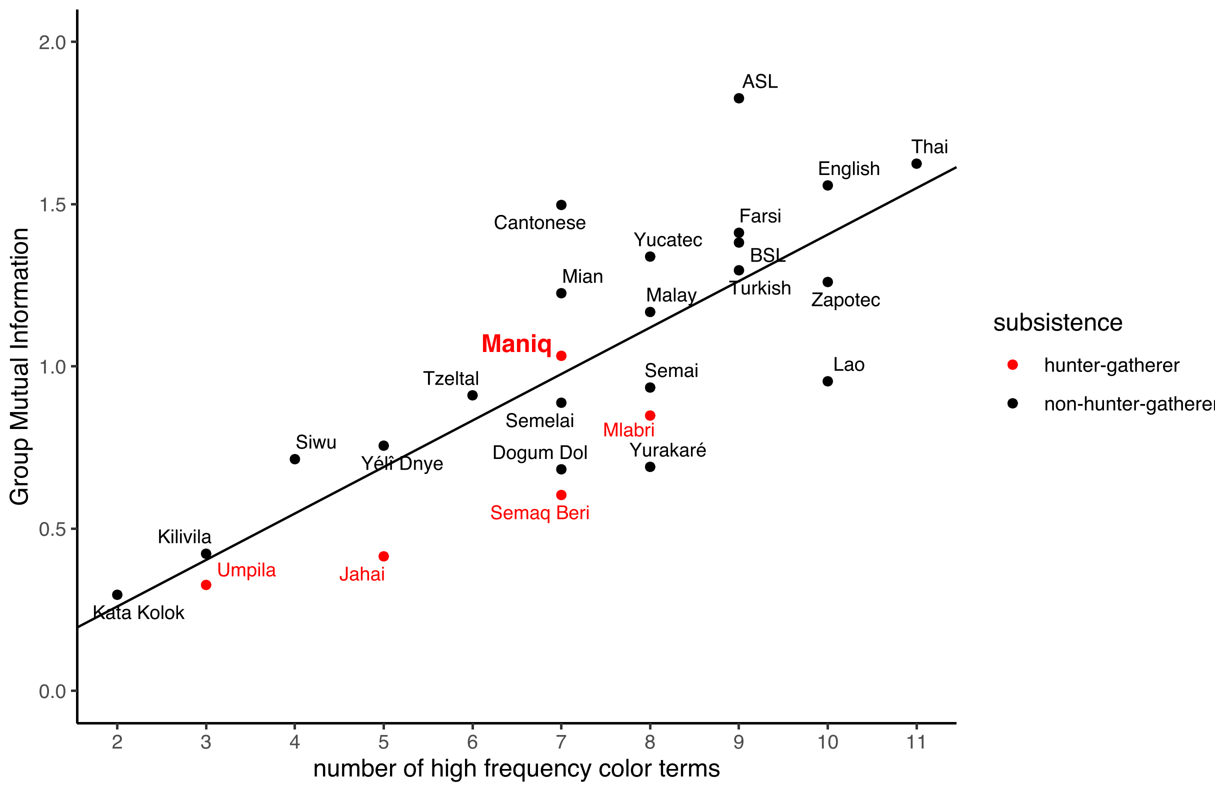

Fig. 2. Group Mutual Information (GMI) for color naming of 80 chips in Maniq in comparison with new Thai and Mlabri data, as well as data from Jahai (Majid & Burenhult, 2014), Semaq Beri and Semelai (Majid & Kruspe, 2018), and 20 additional languages (Majid et al., 2018).

We compared Maniq and the other hunter-gatherers in the dataset, including three other Southeast Asian communities: Jahai, Semaq Beri, and Mlabri, and one Australian community, Umpila, to the non-hunter-gatherer communities. Linear mixed-effects regression analysis conducted in R (R Core Team, 2015) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) with subsistence as a fixed effect (Maniq, other hunter-gatherer, non-hunter-gatherer) and language and stimulus as random effects.[1] We compared models with and without subsistence and language using Maximum Likelihood Estimation with chi-square which showed a significant random effect of language, $\chi^2(1) = 1109.6$, $p = .0001$, indicating considerable variation in color naming agreement between languages, as would be expected. However, there was no significant effect of subsistence on agreement, $\chi^2(2) = 4.15$, $p = .12$. Compared to non-hunter-gatherer groups, other hunter-gatherers were not significantly different in naming agreement ($B = -1.13$, $SE = 0.59$, $t = -1.90$, $p = .07$), and neither were the Maniq ($B = 0.50$, $SE = 1.11$, $t = 0.45$, $p = .65$) (Fig. 1).

2.2.2. Coding efficiency: group mutual information (GMI)

High naming agreement indicates high codability and cognitive salience of color terms among speakers (Brown & Lenneberg, 1954). It has been shown that even inconsistent color naming patterns in a language—while not reaching a theoretical optimum—can be information-theoretically efficient in comparison with other languages (Lindsey et al., 2015). We therefore computed Lindsey et al.'s (2015) formal measure of efficiency—group mutual information (GMI)—to gauge the level of efficiency of the Maniq color system and contextualize it relative to other languages. GMI is an information-theoretic measure that takes into account the full distribution of naming responses given the set of color chips; on the basis of these responses, it quantifies the effort needed to maintain that color system. Lindsey et al. (2015: 2442-2443) explain the calculation of GMI as a way of conceiving of individual data as if it originated from a director-matcher task, i.e., a "color communication game", in which a Sender names color chips according to her color lexicon, and a Receiver tries to identify which color chip the sent term could refer to based on the Receiver's color lexicon. Mutual information (MI) indicates the extent to which a Sender-Receiver pair can be successful and is calculated on the basis of probability the Receiver will identify color sample $X_{1,2,3,\ldots n}$ given the term provided by the Sender. Mutual information thus provides an estimate of how informative the Sender's term is in light of the Receiver's color lexicon. GMI, then, is an aggregate of the color communication games played by all pairs of participants in an experiment: the higher the value, the more efficient the system. Formally:

$$GMI = I_N(C_R; C_S) = \Sigma_{s,r}\, p_N(s,r)\log 2\left(\frac{p_N(s,r)}{p_N(s)p_N(r)}\right)$$

where $p_N(s,r)$ is the average joint probability distribution across all N pairwise color games played within the group of participants. The marginal probabilities $p(r)$ and $p(s)$ are sums of columns and rows of $p(s, r)$, which is a matrix of the joint probability distribution of the variables C_S and C_R, where C_S are the sender's samples and C_R are the receiver's test samples (see Lindsey et al., 2015; SI p. 24, for a full description).

We use GMI as a measure of efficiency in order to compare Maniq to other languages. Note, there is no expectation that optimum GMI is reached by any language (Lindsey et al., 2015). In fact, no language reaches optimum efficiency, since this requires all participants give identical responses, i.e., no variation in responses. Instead, we use this measure to explore how Maniq compares to other hunter-gatherer and non-hunter-gatherer languages in color naming.

GMI is dependent on the number of color terms frequently used by participants, with languages with larger color term systems having higher possible optimum GMI values (Lindsey et al., 2015: 2445). Hence, GMI is

[1] Participants cannot be included as random effects since an aggregate score across participants is calculated for codability.

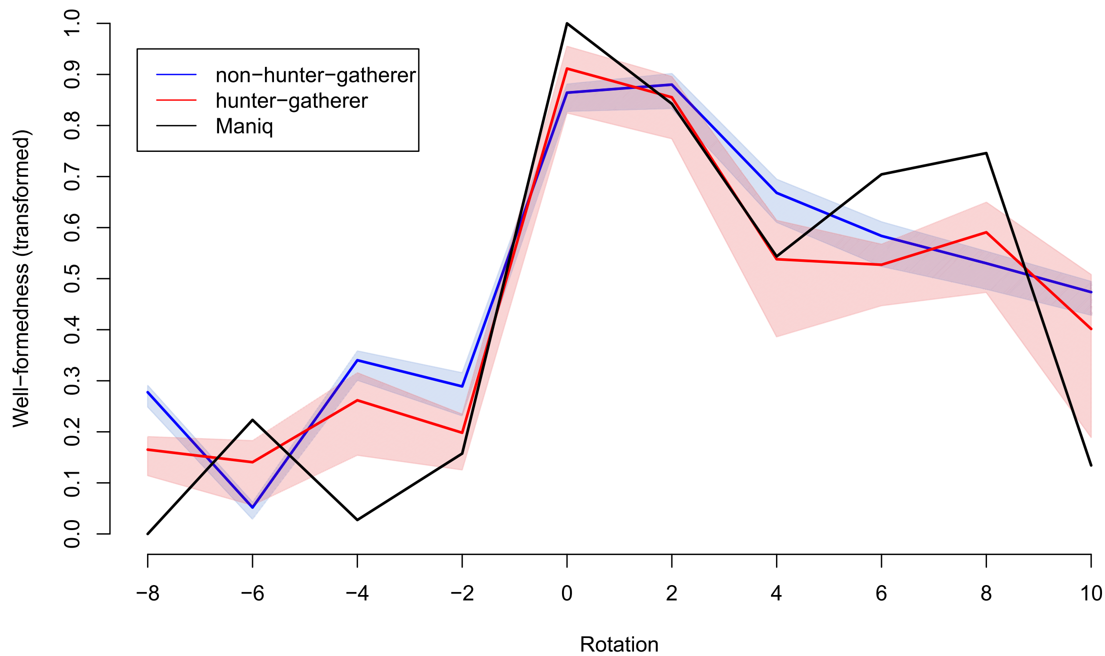

Fig. 3. Well-formedness for Maniq, Thai, Mlabri, Jahai (Majid & Burenhult, 2014), Semaq Beri and Semelai (Majid & Kruspe, 2018), and 20 additional languages (Majid et al., 2018) when rotated 0, 2, 4, 6, etc. hue columns. The most well-formed color-naming scheme across Maniq, non-hunter-gatherer, and hunter-gatherer languages is the unrotated one or the one rotated +2 hue chips.

assessed and plotted in relation to the number of high frequency color terms. After visual examination of a scree from the histograms of term usage in each language, we defined high frequency terms as those which were used by at least 70% of speakers.[2] Fig. 2 depicts GMI across all 25 languages in relation to the number of frequent color terms, which ranged from 2 to 11. Three hunter-gatherer languages appear to have a sizeable number of high frequency color terms, Maniq ($n = 7$), Semaq Beri ($n = 7$), and Mlabri ($n = 8$), but only Maniq has a GMI value above the regression line, indicating it is also relatively efficient.

2.2.3. *Well-formedness: rotation analysis*

Finally, we assessed whether the Maniq color system is well-formed separately from other hunter-gatherer and non-hunter-gatherer groups. To do this, we followed Regier et al. (2007) and Regier et al. (2009) to determine whether Maniq color categories maximize similarity within category while minimizing similarity across categories. The well-formedness of an attested color system is compared to that of hypothetical color-naming schemes by rotating the data by 2, 4, 6, etc. hue columns in the stimulus array. Well-formedness itself is based on the distance between a pair of colors in CIELAB space (Regier et al., 2007: 1437–1438), with dist(x,y) being the CIELAB distance between colors x and y:

$$\mathrm{sim}(x, y) = exp(-0.001 \times [\mathrm{dist}(x, y)]^2)$$

The well-formedness function W takes into account both chips that are labelled identically and chips that are labelled differently:

$$W = S_w + D_a$$

With Sw being a measure of similarity within category:

$$S_w = \Sigma\, \mathrm{sim}(x, y)$$

And D_a being a measure of dissimilarity across categories:

$$D_a = \Sigma\, (1 - \mathrm{sim}(x, y))$$

Regier et al. (2007) showed the color-naming schemes used by most languages are positioned optimally such that—if rotated along the hue dimension—well-formedness decreases. Critically, well-formedness is not foreordained since some attested color systems (e.g., Waorani) are less well-formed than their rotated hypothetical variants (Regier et al., 2009).

Fig. 3 shows the attested color system of Maniq, other hunter-gatherer languages, and non-hunter-gatherer languages. For Maniq and the other hunter-gatherer languages, the most well-formed color system is the one that is attested; for non-hunter-gatherer languages well-formedness is also comparably high with a rotation of +2 hue chips. We return to this in §3.3.3. This shows the Maniq color naming system fits the structure of perceptual color space in much the same way that color systems in other languages do.

3. Study 2: color naming with extended stimulus set

Study 1 demonstrated the Maniq color language is richly lexicalized, codable, and efficient, with 7 high frequency terms and a GMI score exceeding the average value for 7-term languages. It is also well-formed, reflecting optimal partitioning of color space. We next undertook a replication examining color language in Maniq against a larger language sample, with more hunter-gatherer languages included, and an extended stimulus set of 330 color chips. This allowed us to obtain data using the methods of the World Color Survey (WCS) and situate Maniq within the context of a larger set of languages that have been studied extensively previously (e.g., Kay et al., 2009).

3.1. *Methods*

3.1.1. *Participants*

Seven Maniq participants (4 female, age range 7–55) took part in the color naming and focal color task. Five were adults and 2 were minors (of about 7 and 16 years). Four participants were entirely new to the task, but three had taken part in Study 1 in 2009–2010. The data for this

[2] Lindsey et al. (2015) used an 80% threshold, but in this task there are only 80 stimuli and no achromatic chips and the scree better supported a 70% threshold.

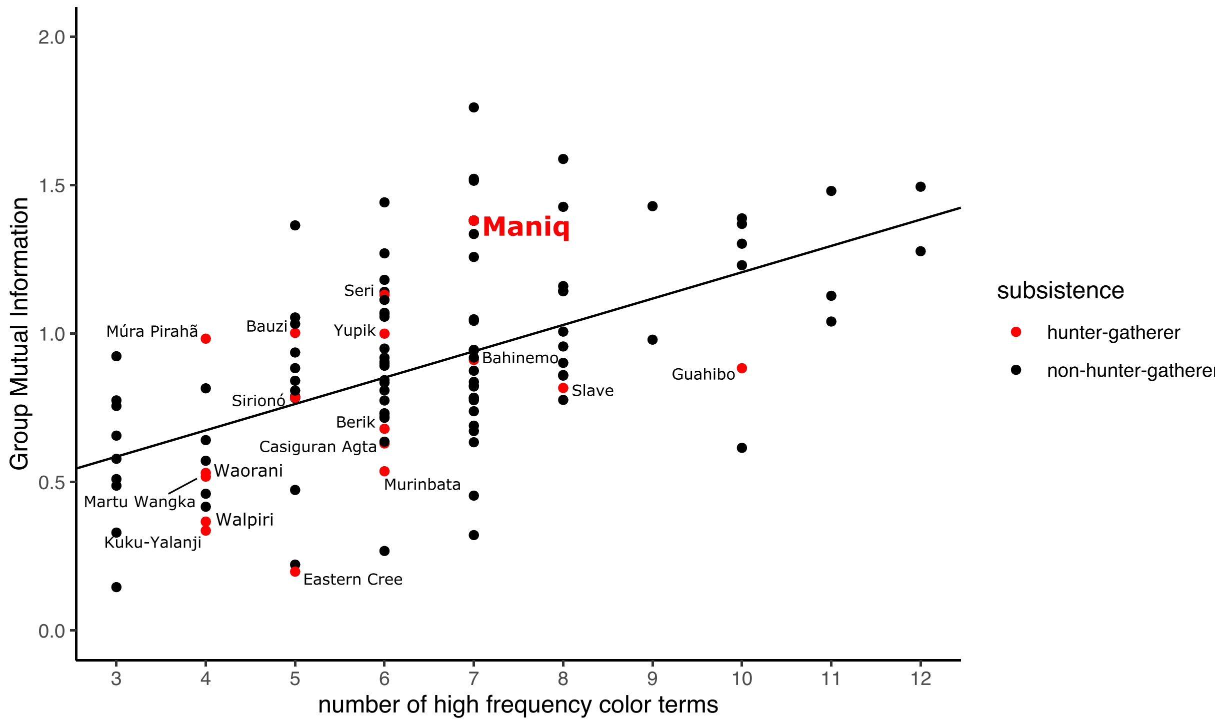

Fig. 4. Group Mutual Information (GMI) for color naming of 330 chips in Maniq in comparison with the World Color Survey languages (110 languages) with red dots indicating hunter-gatherer societies and black dots non-hunter-gatherer societies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

study were collected in 2012. All participants were native speakers of Maniq and had a good command of Southern Thai. None had formal schooling. Informed consent was obtained orally. Though the absolute number of participants (15 speakers in two tasks) is relatively small, the sample can be considered representative. It constitutes about 5% of the total population (estimated at 300 people; Wnuk, 2016) and is reasonably varied since only about half the participating speakers were connected to the place of data collection (i.e., they normally resided in this area) while the other half came from other locations in neighboring districts/provinces.

3.1.2. Stimuli

The stimulus was the World Color Survey (WCS) set of 330 Munsell color chips mounted on cards (Kay et al., 2009).[3] It included 40 equally spaced hues with 8 degrees of brightness at maximum saturation, and 10 achromatic chips. The WCS focal color array was a sheet with 410 Munsell color chips: 330 patches from the WCS naming task, 40 copies of the white chip (N9.5), and 40 copies of the black chip (N0.5). Color blindness was tested with Ishihara plates, similar to Study 1 (§2.1.2).

3.1.3. Procedure

The task followed the same procedure as in §2.1.3. Colors were presented in a fixed random order and participants were asked 'What color is this?'. Participants were free to use any terms or not give a response. After completing the color naming, they selected best examples of the basic color terms identified in naming task with the instruction 'Which one is the best/real X?'. All participants also completed an Ishihara color deficiency test.

3.2. Results

We compared the newly collected Maniq color naming data for 330 color chips to the World Color Survey data which include previously collected data from 110 unwritten languages. We identified those languages which were spoken by traditionally hunter-gatherer communities, based on primary mode of subsistence, to provide a point of comparison to the Maniq. They included a broad range of societies classified as hunter-gatherers (Epps, Bowern, Hansen, Hill, & Zentz, 2012; Kelly, 2013; Lee & Daly, 1999; Roscoe, 2002), including groups with traditionally hunting and gathering subsistence who in relatively recent history (mostly the last few decades or the last century) were forced to adapt their lifeways in some way by colonial or other external forces, e.g., by taking up agriculture or waged labor. We followed the same analysis procedure as in Study 1.

3.2.1. Codability of color

We first examined codability as reflected in Simpson's diversity index across all subsistence groups (Maniq, other hunter-gatherer, non-hunter-gatherer). Using the same approach as described previously in §2.2.1, we found a significant random effect of language, $\chi^2(1) = 11{,}091$, $p = .0001$, but no significant effect of subsistence on color naming agreement, $\chi^2(2) = 3.63$, $p = .16$. Compared to non-hunter-gatherer groups, other hunter-gatherers were not significantly different in naming agreement ($B = -0.14$, $SE = 0.08$, $t = -1.85$, $p = .07$), nor were the Maniq ($B = 0.10$, $SE = 0.29$, $t = 0.35$, $p = .72$).

3.2.2. Coding efficiency: Group mutual information (GMI)

Next, we examined GMI across all 111 languages in relation to the number of high frequency color terms (see Fig. 4). We find languages with 3–12 high frequency color terms. Across the board, hunter-gatherer languages range from 4 to 10 terms. In comparison with other languages, Maniq can be considered highly efficient (a higher GMI than average against the regression line). Moreoever, Maniq appears not to be unique among hunter-gatherer languages for having highly efficient color coding. Several other languages score well above (Múra Pirahã, Bauzi, Seri, and Yupik) or close to the regression line (Sirionó, Bahinemo), with the latter result also reported previously for Hadza hunter-

[3] The stimuli were kindly provided by the WCS team.

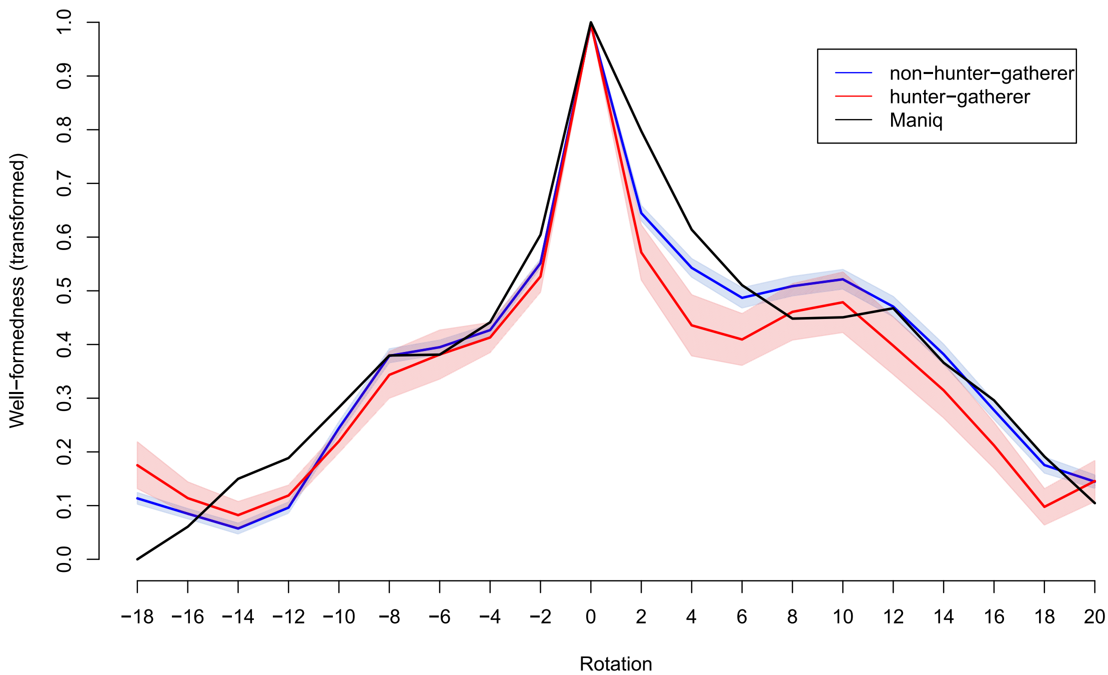

Fig. 5. Well-formedness for Maniq and World Color Survey languages when rotated 0, 2, 4, 6, etc. hue columns. The most well-formed color-naming scheme for Maniq, other hunter-gatherer, and non-hunter-gatherer languages is the one that is actually attested.

gatherers (Lindsey et al., 2015). This echoes earlier work suggesting color naming systems in languages of non-industrialized groups can be highly efficient (Zaslavsky, Kemp, Regier, & Tishby, 2018).

3.2.3. *Well-formedness: rotation analysis*

Maniq, other hunter-gatherer, and non-hunter-gatherer languages all display greatest well-formedness for their actual color naming systems rather than some hypothetical rotated variant (see Fig. 5).

4. The Maniq color naming system in comparative perspective

Both Study 1 and Study 2 demonstrate that in comparison to other languages, the Maniq color system is highly codable, relatively efficient, and well-formed (i.e., optimally partitions color space).

4.1. *Maniq color coding efficiency*

What accounts for the comparatively higher Maniq color coding efficiency as reflected in its GMI score? We consider three possibilities: (1) superior color technology; (2) linguistic affiliation (i.e., there is exceptional prominence of color in the Aslian branch of the Austroasiatic family); (3) recent contact with Thai (i.e., the Maniq color lexicon reflects Thai categories). The technological account would predict that high GMI could be due to more advanced color technology in Maniq compared to hunter-gatherers with lower GMI scores. Though hunter-gatherers generally tend to possess relatively simple technologies, use of technology across hunter-gatherer groups varies depending on factors such as ecology, residential mobility, and cultural importance of ornaments and prestige items (Hayden, 1998; Kelly, 2013). The Maniq, however, are predominantly mobile, do not use prestige items, and have among the least complex technologies in our dataset. They therefore have less, not more, technological control of color and typically also less exposure to artificially colored objects than some traditionally hunter-gatherer groups. For example, other groups with relatively high GMI scores, Seri and Yupik, have more complex color technologies (see e.g., Crowell, 1992 on Yupik painted face masks), linked to their settled residence. At the other end of the GMI distribution, too, there are groups with relatively complex color technologies such as Eastern Cree (who make sophisticated painted crafts, Graburn, 2006), Australian Aboriginal communities – Martu Wangka, Kuku Yalanji, Walpiri, Murinbata (who traditionally use ochres to obtain red, yellow, and white paint and more recently also use acrylic paints) (see e.g., Higgs, 2016), and groups with less technological complexity (with no mention of paints or dyes) like the Casiguran Agta (Headland, 1993). While we do not have directly comparable information on technologies in the communities interviewed in the World Color Survey and we cannot confirm what the situation was at the time of testing, it is clear the technological variation documented in the general ethnographic record does not map onto variation in color language. The technological account, therefore, cannot explain the high GMI score in Maniq.

Crucial for the second hypothesis, we were able to compare Maniq to four of its close Aslian relatives: Jahai, Semaq Beri, Semai, and Semelai. If the high GMI score of Maniq was due to the exceptional linguistic prominence of color within the Aslian language group, we would expect all of these languages to exhibit Maniq-like GMI. This, however, is not the case since other Aslian languages including those spoken by the hunting-gathering Jahai and Semaq Beri had lower GMI scores falling below the regression line (Fig. 2). This difference can be largely attributed to differences in codability. Maniq scored significantly higher on codability than Jahai, $t(79) = 9.56$, $p < .0001$, Semaq Beri, $t(79) = 10.52$, $p < .0001$, Semai, $t(79) = 8.02$, $p < .0001$, and Semelai, $t(79) = 2.68$, $p < .01$, as revealed by pairwise t-tests with Bonferroni correction. Since groups like the Jahai live in a similar rainforest environment to the Maniq, this makes an explanation based primarily on environmental factors less plausible.

Finally, we examined the role of contact with Thai as a possible explanation of the Maniq color system. Maniq is subject to a unique contact situation in the Aslian context, being the only member of the branch spoken entirely in Thailand, so its high GMI score could be an outcome of recent Thai contact. If Thai contact alone was linked to efficient color coding, we would expect to observe the same effect in other languages exposed to Thai such as Mlabri—the language spoken by the only other traditionally forest-dwelling hunter-gatherer group of Thailand. However, Mlabri has a low GMI score, so the mere fact of being in contact with Thai was not associated with efficient color coding. As in the case of the Aslian languages examined above, this can be explained by low codability in Mlabri compared to Maniq, $t(79) = 4.27$, $p < .0005$. The specifics of the contact situation are different in the two communities, i.e., living in permanent settlements (Nimonjiya, 2013), Mlabri people are in more frequent contact with Thai than Maniq people, most of whom live in temporary camps inside the forest. In addition, Mlabri children receive formal schooling in Thai, while most Maniq children do not go to school (Bishop & Peterson, 2003; Wnuk, 2016). Not surprisingly there are multiple Thai color term borrowings in Mlabri, while there are no borrowed Thai color terms in Maniq: all Maniq color terms are part of the indigenous lexicon with typical Maniq phonological segments, syllable, and word structures (Wnuk & Burenhult, 2014) and without any formal resemblance to Thai terms.

While there is no evidence for borrowing of the terms themselves, it is possible the Thai color categories affected the shape of the Maniq color system. Influence of this kind is difficult to rule out entirely, but—even if it exists—it is subtle compared to Mlabri, which has multiple Thai color borrowings and a number of color categories similar to Thai, some of which closely resemble in shape Thai color categories (Fig. 6). Critically, however, Mlabri speakers display limited color naming agreement, resulting in a low GMI score, so ultimately Thai influence on Mlabri does not give rise to high coding efficiency.

4.2. Summary

In sum, Maniq displays high color coding efficiency compared to many other hunter-gatherer groups. This is in spite of the fact that Maniq society does not stand out in having an advanced color technology, arguing against technology being a necessary condition for the development of large and efficient color lexicons. While it could be expected that hunter-gatherer groups related to Maniq (including the closest linguistic relatives from the Aslian branch—Jahai and Semaq Beri—and the more distant relative from the Khmuic branch—Mlabri) would be similar in this respect, we find even though they have between 5 and 8 high frequency terms, their coding efficiency is below what is typically found in languages with the same number of color terms. This, in our view, has to do with the fact that the Maniq color system is relatively stable and does not seem to be undergoing any major reorganization. This is contrary to Jahai, Semaq Beri, and Mlabri which in the recent past have expanded their color lexicons with borrowings from their respective contact languages (Malay and Thai). This has caused a reorganization of their color systems and a decrease in stability of color categories, as reflected in lower agreement (see Grimm, 2014, for similar observations among Gyeli Pygmy hunter-gatherers who borrowed color terms from neighboring Bantu languages). For example, Mlabri was reported to have just 4 color terms in 1985 (Luangthongkum, 1985), whereas now it has 8 high frequency terms and several marginal terms, many of which are recent Thai loans (e.g., *sɔm* 'orange', *chompu* 'pink', *muaŋ* 'purple', *namŋɯn* 'dark blue', *thaw* 'gray'). These have been acquired due to intensified contact with Thai or increased participation in schooling after Mlabri people settled in permanent villages in the late 1990s (for a description of similar developments among the Jahai, see Burenhult, 2005). Maniq, on the other hand, did not borrow color terms from Thai (or Malay, with which it was in contact earlier in history), as it experienced less intense contact, i.e., no similar large-scale regroupment program—the majority of Maniq are still nomadic.

5. Maniq color vocabulary

The previous sections show color in Maniq is efficiently coded and richly lexicalized in a set of 7 high frequency color terms. We now turn to examining Maniq color terms in greater detail, including establishing which terms qualify as basic and discussing secondary color terms used primarily outside of color naming tasks.

5.1. Basic color lexicon

First we consider the basic color lexicon. According to the results of color naming tasks in Study 1 and Study 2, Maniq has 7 high frequency terms: *bəlɛɲ, blaʔɛm, hayet, haŋɔt, bagı̃ɛ̃c, paliek,* and *panuk*. Two terms—*paliek* and *panuk*—are synonymous terms for 'white' and represent a single color category. Based on the full set of criteria for basicness (Berlin & Kay, 1969) Maniq thus has a 6-term system. Fig. 7 shows the partitioning of the color space in Maniq based on the color naming tasks with 80 and 330 chips, as reported in Studies 1 and 2.

These are in every sense fully-fledged abstract color terms, satisfying all main and subsidiary criteria for basicness (Berlin & Kay, 1969). Formally, they are monolexemic. They are semantically general and their meanings are not included under the meaning of other color terms. In addition, they have broad application—they are not restricted to a narrow class of objects, but are applicable to different kinds of things. All speakers employed all 6 terms in the task showing considerable agreement. Their reference therefore appears stable across participants, suggesting they are psychologically salient. All of the items also satisfy the subsidiary criteria for the basic status. Namely, they have the same distributional potential, i.e., they belong to the same word class of stative verbs, occur with the same set of modifiers, themselves act as modifiers in the noun phrase, and take verbal morphology (Wnuk, 2016). None of the terms are loanwords. They also do not have a transparent connection to specific objects. Some terms have cognates in other Aslian languages (e.g. the Maniq 'grue' and 'red' terms *bəlɛɲ* and *bagı̃ɛ̃c* are cognate with the *bəlʔəɲ* and *blakãc* in the related language Batek spoken in Malaysia (Burenhult, 2009)), suggesting that they are of Aslian origin and were present in the ancestral stages of the Aslian language group, i.e., precursors of present-day Maniq. These terms therefore evolved over a long period, and most likely lost their connection to concrete entities becoming fully abstract in the course of language evolution. One of the terms—*bagı̃ɛ̃c* 'to be red'—has a secondary sense of a collective noun meaning 'blood and raw meat', but the color term is not derived from the noun. On the contrary, comparative evidence from Batek suggests the object sense is a metonymic extension of the color sense, specific to Maniq. In Batek, the cognate of *bagı̃ɛ̃c* refers only to the red color and does not have a sense of raw meat and blood (T.P. Lye, personal communication). In addition to being monolexemic, the terms are synchronically monomorphemic, i.e., they are not divisible into smaller meaningful units. Taken together, all criteria suggest that indeed these terms are basic color terms.

An inventory of 6, while not close to the original maximum 11 proposed by Berlin and Kay (1969), is considered large in a broader context (Casson, 1994; Naroll, 1970), and is attested in languages of societies with advanced color technologies such as Vietnamese and Hausa (Naroll, 1970). The exact response times were not measured, but overall speakers were reasonably quick and showed little hesitation. The responses were concise, further indicating high codability of color (cf. Majid & Burenhult, 2014). Most frequently, they consisted of a single word—the color term, e.g., *hayet* 'yellow', or a simple phrase, e.g., *si hayet* 'yellow color' (color be.yellow), *ʔuʔ ʔɛn pãʔ hayet* 'this one is also yellow' (3S DEM also be.yellow). Some participants employed simple modifiers, e.g., *naki* 'to be real' as in *hayet naki* 'real, genuine yellow', but their use was limited – only about 1% of responses in Study 1 and 5% of responses in Study 2 were modified (for a full summary of color term modification in both studies, see Supplementary Material).

Contrary to Kay and Maffi (1999), we found terms employed in the task captured color distinctions of objects from the local environment, including culturally salient plants and animals, e.g., wild yams and gibbons (Fig. 8). These entities come in a variety of colors and constitute implicit contrast sets (starchy tuberous foods and gibbons). While members of these sets are not distinguishable only by color, color is a

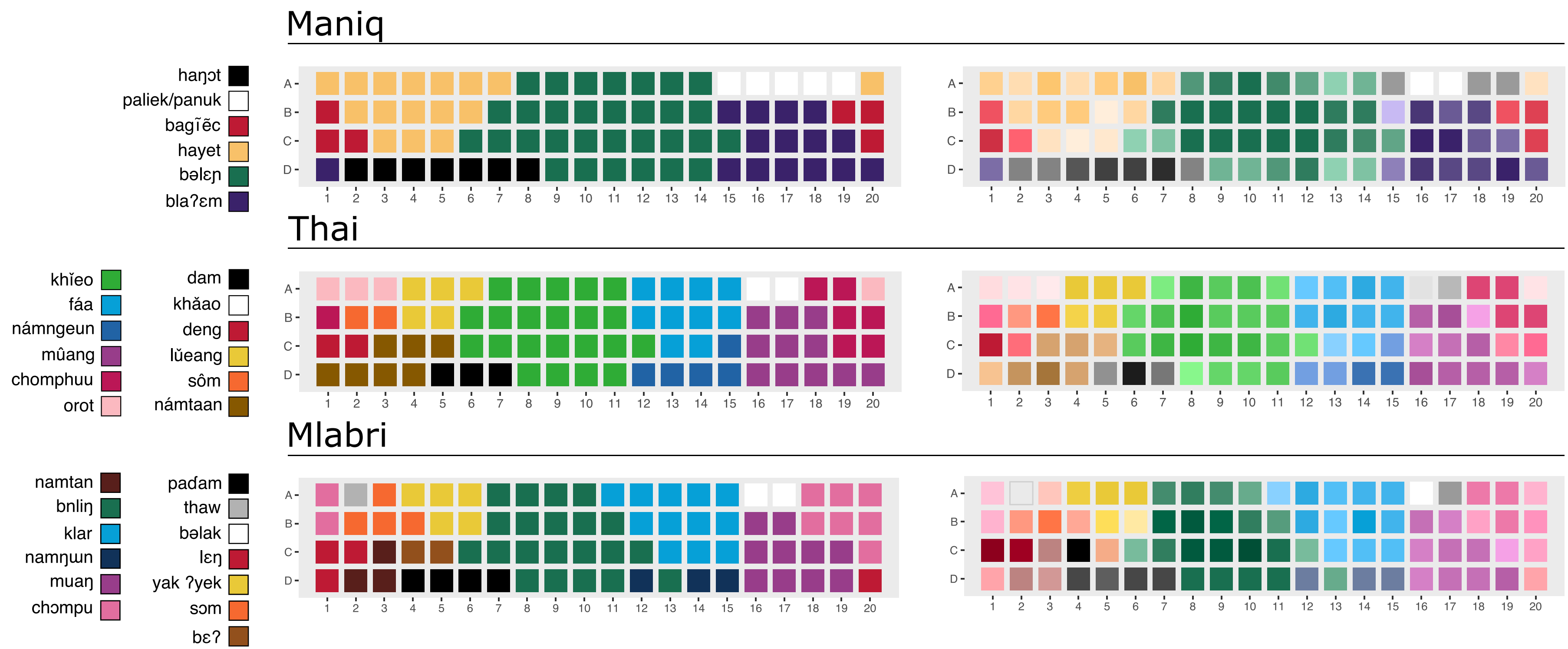

Fig. 6. Maniq, Mlabri, and Thai color systems. Left: Color arrays showing focal colors. Right: Consensus maps with relative consensus coded as brightness. Darker shades indicate higher consensus. The Mlabri color terms *namtan, thaw, sɔm, chompu, namŋɯn* and the Thai term *orot* were modal responses (used by the majority of participants) for one or more chips, but do not qualify as high frequency terms (used by at least 70% of speakers).

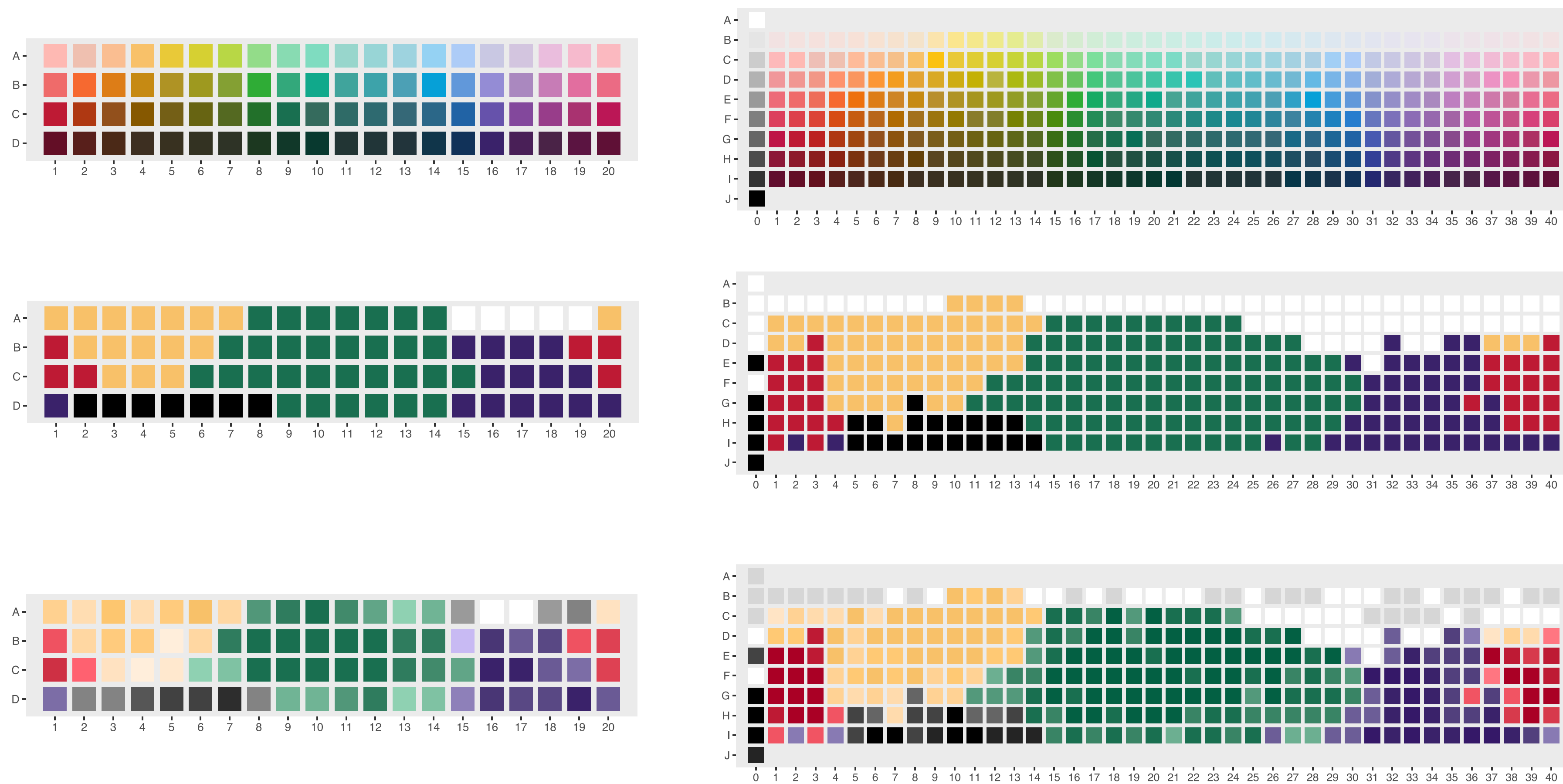

Fig. 7. Maniq color naming results for 80 (left) and 330 chips (right). Top: Colors approximating test stimuli. Middle: Color categories (determined by modal terms for each chip) and focal hues (indicating span of category) for each color term. Bottom: Consensus maps with relative consensus coded by brightness (darker hues indicate higher consensus).

visually prominent dimension of difference between them which is referred to in everyday conversations (see further §6.2). So while these objects are not part of the industrialized world, their cultural salience shapes patterns of communicative need for color within the Maniq community (Zaslavsky, Kemp, Tishby, & Regier, 2019) affecting the shape of the color lexicon. Even purple—considered a late-stage term in the color evolutionary sequence (Kay et al., 2009) and a term whose usefulness is posited to be particularly closely linked to technological innovation (Conway, Ratnasingam, Jara-Ettinger, Futrell, & Gibson, 2020)—is commonly employed. This runs contrary to predictions based on perceptual structure on color lexica (Kay & McDaniel, 1978), but is not entirely uncommon since presence of purple in the absence of the blue-green distinction has previously been reported for a number of other languages, e.g., earlier forms of Japanese and related Japonic languages (Huisman, van Hout, & Majid, 2022), Setswana (Davies et al., 1992), and some languages in the World Color Survey sample (e.g. Cofán; Kay et al., 2009).

5.2. Secondary color lexicon

The color language literature is dominated with discussions of basic color terms, but linguistic elaboration of color is further indicated by secondary color terms. Though secondary terms are sometimes viewed as indicators of lesser focus on color as an abstract property (Kuschel & Monberg, 1974), in reality such terms are often present in a language along with basic color terms (as in English, for instance) and can also be abstract (e.g., the English *beige*). Their presence is an indication of general interest in color and a need to make fine-grained distinctions beyond basic color categories. Thus, if technology is a sine qua non for lexicalization of color, we would expect to find few secondary color terms in Maniq. However, we find Maniq has an elaborate secondary color lexicon, with over two dozen terms. These capture fine color nuances of cultural relevance (Table 1) and reveal a close link with the natural world.

5.2.1. Study 3: exemplar listing

To probe secondary color terms, we explored the type and range of exemplars most commonly associated with each term by Maniq participants. There were eight secondary terms that could not be included in exemplar listing as they emerged after the task. They included: three "red" (*laŋɔɲ, yəhəy, katəɲ*), two "black" (*haʔɔ̃ŋ, haŋɨt*), one "green" (*bayel*), one "yellow" (*bayɛϕ*), and one "white" term (*blahut*) (for further details on the meaning of all secondary terms, see Supplementary Material).

5.2.1.1. Methods

5.2.1.1.1. Participants. A total of 8 participants (4 female; approximate age 20–55) took part in exemplar listing. All were native speakers of Maniq.

5.2.1.1.2. Stimuli. The stimuli were the 21 terms in Table 1. Being fairly semantically specific and context-bound, these terms did not feature prominently in the naming tasks (only 7 secondary terms were used by fewer than 50% of speakers, of which 5 were used by only one speaker), but emerged in spontaneous interactions and semi-structured interviews over the course of long-term fieldwork.

5.2.1.1.3. Procedure. The instruction used to elicit verbal responses was 'What is x like?', where x was a target term. Speakers listed a variety of items: objects, body parts, animals, and plants or their elements. All provided responses were taken into account. Participants were encouraged to list multiple exemplars by additional prompts, e.g., 'What else?', or repeating the initial instruction. Since elicitation was carried out over the course of two field trips and newly discovered verbs were added to the list in later stages, some speakers provided exemplars for more terms than others.

5.2.1.2. Results and discussion. The results are presented in Table 1 below. Numbers in brackets indicate the number of participants who provided the response. The list includes a variety of animals and plants from the local environment. Animal identifications were mostly based on earlier elicitations with zoological field guides (e.g. Francis, 2001) or Thai translations. Most plants, notably the various species of wild yams (*Dioscorea* spp.), were identified with the vernacular labels in Maneenoon (2001, 2008). In cases where identification was impossible, the Maniq forms are given in square brackets, along with an approximate gloss in English.

The relationship between terms used for paraphrases and secondary terms is similar to hyponymy (i.e., inclusion) where the basic color term has a function similar to a superordinate term and the secondary term to a hyponym. However, since in standard hyponymy the meaning of the hyponym should be fully included under the meaning of the superordinate term and this not the case here (e.g., some secondary terms carry non-color information; see Supplementary Material), the secondary terms cannot be considered true hyponyms (cf. Levinson, 2000). We therefore refrain from using the terms "superordinate" and "hyponym".

The terms express a variety of distinctions with color information which are highly relevant in the everyday life of the Maniq, relating to culturally significant aspects of the environment and material culture. Similar to secondary terms for color reported in other languages (Kuschel & Monberg, 1974; Wierzbicka, 2008), they are typically associated with specific objects (similar to *blond* being associated with hair) and sometimes imply additional non-color information such as reflectance, spatial arrangement, age, and plant growth stage. For example, red objects are typically described with the basic red term *bagĩɛ̃c*, but the specific type of red characteristic of impregnated bamboo containers is referred to with the secondary term *lɲŋaɲ*. All verbs, with the exception of *taʔum* and *batgɨt*, were linked to several exemplars. Some of them had strong prototypes listed by the majority of participants, e.g., *bayul* (ashes), *bayiϕ* (smoke). Others did not have a single prototypical exemplar, but were associated with a range of objects, e.g., *batɛ̃ŋ* (trousers, child's hair, flying squirrel, civet sp., sky), *haɲɛp* (bearcat, wild pig's fur, trousers, candy/biscuit, '*lamŋɨm*' plant). Such variation in response patterns suggests the terms differ in the level of context-restrictedness. In addition, even in the case of terms with a narrow meaning, e.g., *bayɛϕ*, associated with the bright yellow color of a baby stump-tailed macaque, the meaning is not set in stone and speakers may apply such terms spontaneously in novel contexts (i.e., with objects that they're not normally exposed to), e.g., *bayɛϕ* was used to describe blond hair of a child viewed in a photograph.

On a continuum of abstractness, these terms are less abstract than basic terms because of a narrower range of reference, but they also have some typical features of abstract vocabulary. First, their forms are not linked to specific concrete entities (except for *haʔɯŋ*, derived from a name of a civet), so it's not transparent for speakers what entity they derive from (as in English terms *ruby, scarlet*, etc.). Second, even though their range of applicable contexts is typically limited, in most cases they are not tied exclusively to a single exemplar (possible exceptions here are *taʔum* and *batgɨt*, which seem to be associated with only one thing). And third, they are formally similar to basic color terms—they belong to the class of stative verbs, they can be used in similar syntactic frames, and take a similar set of derivational morphemes.

5.2.1.3. Summary. The examination of secondary color terms reveals color is even more expressible than the initial exploration based on the naming tasks would suggest. These terms express fine nuances in color and aspects of visual surfaces. Altogether, evidence from the lexicon suggests color is salient in Maniq, but leaves open the question of the place of color in Maniq people's everyday life. The following section explores color in relation to Maniq culture and everyday discourse in order to provide a more comprehensive characterization of the color domain among Maniq speakers.

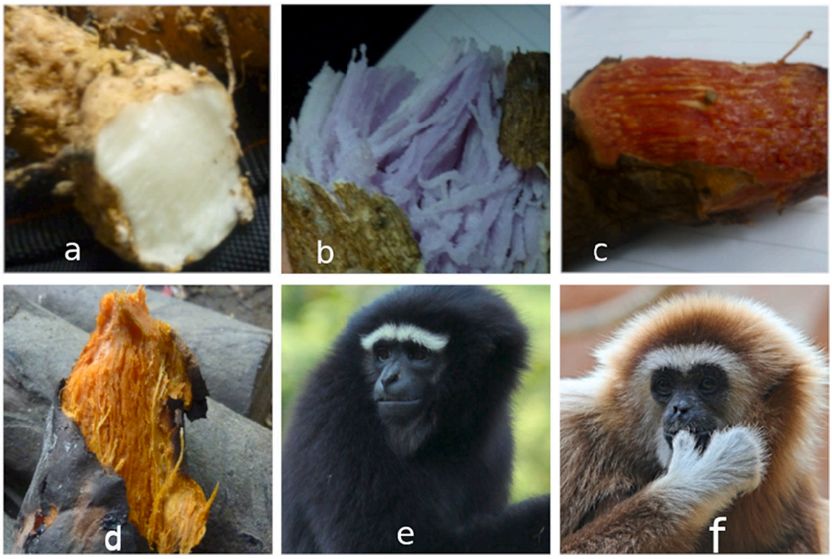

Fig. 8. Natural color objects from the Maniq setting routinely described by basic color terms (provided in brackets) **a** *Dioscorea orbiculata* (*paliek/panuk* 'to be white') **b** *Dioscorea glabra* (*blaʔɛm* 'to be purple') **c** *Dioscorea* cf. *piscatorum* (*bagĩẽc* 'to be red') **d** Sweet potato (*hayet* 'to be yellow') **e** Agile gibbon (*Hylobates agilis*) (*haŋɔt* 'to be black') **f** Lar gibbon (*Hylobates lar*) (*blaʔɛm* 'to be purple'/*bagĩẽc* 'to be red'). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6. Color in everyday life of the Maniq

The place of color in the community was explored through participant observation and semi-structured interviews relating to various aspects of the natural world carried out during long-term fieldwork. The ethnographic method employed here follows a long tradition in cultural and social anthropology. It has been widely applied in color language research and resulted in influential contributions, e.g., Conklin (1955), Kuschel and Monberg (1974), Levinson (2000). As a participant in everyday activities, the first author was able to observe the use of color terms in ordinary situations. While these observations are not quantifiable as they were not systematically recorded, they provide an interpretive background to the experimental tasks. Their role here is to showcase the kinds of contexts in which color is spontaneously mentioned by speakers of Maniq and show that color terms are routinely used not just for referential discrimination, but also to describe objects, structure knowledge, and symbolically represent complex cultural notions. Several examples from audio-recorded sessions (verbatim quotes of statements by Maniq speakers) are provided as illustrations.

6.1. Material culture and color

As a mobile hunter-gatherer group, The Maniq have a relatively simple material culture and eschew the accumulation of objects (Wnuk, 2016), as is typical for highly mobile hunter-gatherer groups with egalitarian social organization (Woodburn, 1982). There is no use of elaborate color technology in the community. Common artifacts like baskets, quivers, blowpipes, etc. are not painted or decorated with color. Colors of the raw materials become transformed in the production process, e.g., bamboo impregnated with beeswax becomes *lɲŋaɲ* 'to be orange/red/brown', or they change naturally as artifacts age, but there is no deliberate color application. Similarly, there is little traditional bodily adornment involving color. Traditional personal ornaments are made of simple objects such as vines, seeds, bones, turtle shells, etc. with no applied coloring. This may have been different historically since early ethnographic reports on various Semang groups (e.g., in the Malaysian states of Perak and Kedah) describe face and body painting practices involving at least black, red, yellow, and white (Evans, 1937; Skeat & Blagden, 1906). Less is known about the Maniq specifically, but the scarce reports mention face painting with charcoal for ritual dance (Bernatzik, 1938), wearing of flowers as ornaments and dyeing of bamboo combs, however no details are provided about possible colors of dyes (King Chulalongkorn, 1907). Applying charcoal and ashes on the forehead is still practiced, but its main purpose seems to be medicinal rather than decorative. The other practices do not seem to be commonly followed anymore. For instance, no use of combs has been observed among the Maniq, and this seems to have been the case at least since the 1960s (Brandt, 1961).

6.2. Color term use in daily contexts

Although use of color technology is limited, the Maniq are attentive to color and routinely include references to color in descriptions of the surrounding world. These might relate to, for instance, atmospheric phenomena, geological features, as well as various plants and animals, most notably some culturally important taxa such as wild yams or gibbons, as referred to in Fig. 8. Tubers of wild yams, for instance, vary in color from white, yellow, purple to red (Maneenoon, 2008). Two main species of gibbons found in the area—the agile gibbon (*Hylobates agilis*) and the lar gibbon (*Hylobates lar*) (also known as the white-handed gibbon)—can have black, dark brown, reddish, buff, and blonde fur coloring (Francis, 2001). These differences are noted by the Maniq (see (1) below) and further connected with other specific properties, e.g., habitat, as expressed in example (2).[4] Similarly, atmospheric phenomena can be defined with reference to color, as in (3) in which the speaker uses color to talk about the difference between clouds (*ʔac*) and mist (*ʔal*).

(1)

panuk	*biʔ=cas (…),*	*panuk*	*mɛt,*
be.white	PL=hand	be.white	eyes
kaʔɔʔ	*cɨŋ*	*blaʔɛm*	
back	belly	be.purple	

'Its hands are white, its eyes are white, its back and belly dark brown/purple.'

(2)

hay	*tawɔh*	*kayɔm*	*haɲɛp,*	*tawɔh*	*kahɨp*	*ʔaʔ*	*panuk*
like	gibbon	down	be.dark	gibbon	forest	CONTR	be.white

'Like the gibbons living down are dark, and the gibbons in the forest are white.'

[4] Glossing abbreviations used for grammatical elements: 1 'first person', 3 'third person', S 'singular', D 'dual', PL 'plural', DEM 'demonstrative', CONTR 'contrastive', CAUS 'causative', FOC 'focus', MULT 'multiplicity', NEG 'negative', RECOG 'recognitional demonstrative', PROG 'progressive', IMFV 'imperfective'.

Table 1
Secondary terms, general terms used to paraphrase them, and example objects described with the terms. Numbers in the third column indicate the total number of different objects elicited. Numbers in brackets following exemplars indicate the number of speakers who produced that exemplar. The basic Maniq terms used in paraphrases were: *paliek/panuk* for ‘white’, *haɲɔt* for ‘black’, *bag̃ĩẽc* for ‘red’, and *hayet* for ‘yellow’.

Term	More general term used in paraphrase	Number of objects	Exemplars
palak	white	8	muntjac (*Muntiacus muntjak*) or its body parts (3), hair of animal (1), black giant squirrel (*Ratufa bicolor*) (1), cheeks of pig-tailed macaque (*Macaca nemestrina*) (1), dirty white socks (1), gray shirt (1), paper (1), tree (1)
haliek	white	5	basket (4), cucumber skin (1), bamboo tube for water (1), liver of Sunda flying lemur (*Galeopterus variegatus*) (1), liver of Asian leaf turtle (*Cyclemys dentata*) (1)
halãk	white	7	basket (3), cucumber skin (1), skin of [*kunu* - gourd type] (1), head of a Maniq person (1), paper (1), [*taduk* – plant sp.] (1), Prevost’s squirrel (*Callosciurus prevostii*) (1)
hlawãk	white	9	shirt (2), wings of great hornbill (*Buceros bicornis*) (1), plumage of [*cakɛp* - bird sp.] (1), muntjac (*Muntiacus muntjak*) (1), lesser mouse deer (*Tragulus kanchil*) (1), butterfly/moth (1), civet’s hair (1), Prevost’s squirrel (*Callosciurus prevostii*) (1), camouflage trousers (1)
bayɔ̃ɸ	white	10	mist (3), smoke (2), cloud (2), gray/white hair (2), paper (2), rotten wood (2), ashes (2), flower (1), nose of sun bear (*Helarctos malayanus*) (1), head of dusky leaf monkey (*Trachypithecus obscurus*) (1)
bayĩɸ	white	3	smoke (5), ashes (1), mist (1)
bayul	white	5	ashes (5), soil (2), mist (1), white-crowned hornbill (*Berenicornis comatus*) (1), feces of Sunda flying lemur (*Galeopterus variegatus*) (1)
bayek	white	10	rice (4), wild yam (*Dioscorea orbiculata*) (3), soil (2), other wild yams: *Dioscorea filiformis* (1), *Dioscorea daunea* (1), *Dioscorea stemonoides* (1), *Dioscorea pentaphylla* (1), *Dioscorea pyrifolia* (1), [*sac* – wild yam sp.] (1), inside of tree (1)
sakwɨk	white	9	jeans (3), mist (2), stalactite/stalagmite (1), tip of *Dioscorea orbiculata* shoot (1), *Dioscorea daunea* (1), sky (1), mist (1), clear water (1), deep water (1)
pataw	white	4	soil (4), unripe petai (*Parkia speciosa*) (1), [semi-ripe *hak* – type of fruit] (1), [*tayuʔ* – type of fruit] (1)
lahiy	white	2	rock (2), sky (2)
lŋgieŋ	white	4	rock (2), white bowl (1), knife (1), mushroom (1)
lalɛ̃ŋ	white	2	eyes (3), water (2)
bakay	white	4	sun (1), water (1), road (1), soil (1)
batɛ̃ŋ	black	5	trousers (1), hair of a Maniq child (1), flying squirrel (1), [*kadieʔ nɨŋsuŋ* - civet sp.] (1), sky (1)
haɲɛp	black	5	bearcat (*Arctictis binturong*) (1), fur of wild pig (*Sus scrofa*) (1), trousers (1), candy/biscuit (1), [*lamŋɨm* – plant sp.] (1)
taʔum	black	1	flying squirrel (4)
batgɨt	black	1	sky (2)
lɲŋaɲ	red	5	bamboo quiver (2), muntjac (2), soil (2), outer shaft of a blowpipe (1), tree (1)
talaɲ	red	2	fruit (4), leaves (1)
haʔuŋ	yellow	2	mud (1), yellow package (1)

(3)

ʔɔʔ	*ʔac*	*ʔɛʔ*	*haɲɔt*	*pãʔ,*	*ʔɛʔ*	*hayet*	*pãʔ,*
RECOG	cloud	3	be.black	also	3	be.yellow	also
ʔɛʔ	*blaʔɛm*	*pãʔ,*	*bɛm*	*ʔɛʔ*	*ʔɔʔ*	*ʔac,*	
3	be.purple	also	many	3	RECOG	cloud	
ʔaʔ	*ʔɔʔ*	*ʔal*	*ʔɛʔ*	*nay*	*ʔɛʔ*	*baʔ*	
CONTR	RECOG	mist	3	one	3	itself	
naʔ	*lɔʔ*	*naʔ*	*bayul*				
FOC	float	FOC	be.white				

'There are clouds which are black, also yellow ones, and purple ones, there are many (kinds of) clouds. Mist/fog, on the other hand, is of one kind, it floats and it's white.'

Although previous literature has stressed the object-distinguishing function of color terms, i.e., when two otherwise similar objects can be told apart by referring to color (cf. Wierzbicka, 2008), in everyday life of the Maniq color terms are used for a variety of purposes, e.g., to talk about objects for which color is not a stable feature or to metonymically represent objects or actions for which there are no dedicated labels. For example, in (4) below, the speaker describes the edible and non-edible parts of a tuber as the 'white' and 'black'.

(4)

ʔiɲ	*kɔs*	*pɔk*	*haɲɔt,*	*ʔiɲ*	*hãw*	*paliek*
1S	cut	remove	be.black	1S	eat	be.white

'I cut off the black (part). I eat the white (part).'

Some spontaneous uses of color terms reveal a conceptualization of more complex actions or entities via color. For instance, when talking about writing (something the Maniq themselves do not do, and do not have a word for), a Maniq speaker used the term *h<l>aɲɔt* 'to be black (here and there)' (be.black<MULT>) to describe handwritten text. Similarly, color terms are employed to attract someone's attention to a specific object, as in (5), in which a Maniq speaker referred to a green snake under her shelter with the general label *yəkɔp* 'snake' and specified its color as *bəlɛɲ* 'grue' when alerting her male relative about it. In another situation, a speaker participating in a director-matcher task used a color term *lalɛ̃ŋ* 'to be white (of eyes)' to direct the matcher to the right stimulus, cf. (6). The stimuli were images of simple scenes with the same actor. They differed on parameters such as body posture and gaze direction, and consequently, the appearance of the actors' eyes.

(5)

yəkɔp (…),	*ʔɛʔ*	*b<m><i>lɛɲ*
snake	3	be.grue<PROG><CAUS>

'A snake (…), it's making itself grue!'

(6)

ʔɔʔ	*c<m>i<k>yɛ̃k*	*mɛt*	*lalɛ̃ŋ*	*ʔahaw*
RECOG	look.sideways<PROG><IMFV>	eyes	be.white	be.big

'The one that is looking sideways with big white eyes.'

Examples like this suggest color is a salient element of objects and scenes, and spontaneous use of color terms is a strategy for coordinating reference.

Colors are also significant for the Maniq because of their powerful symbolism. Red is particularly meaningful since it evokes the idea of blood, a potent symbol of life and death. The special significance of red in the context of blood is widespread among Aslian-speaking groups (Dentan, 1979; Endicott, 1979). For instance, blood let from lower legs (calves) is subject to offerings in blood-throwing ceremonies that are meant to appease the thunder being *Kaʔey* (Kricheff, 2019; Needham, 1967). There are also connections to other culturally elaborated domains—for example, color plays a role in religious and ethnomedical beliefs. It forms part of culturally salient sensorially-defined notions of well-being and safety, e.g., the orange/yellow color of the sun during hot weather is explicitly associated with disease transported down with heat and a dangerous smell. The sunlight is also believed to turn Maniq people's skin yellow, which, as explained by the speaker in (7), would pose a risk of death, while the usual black color of skin is a sign of safety.

(7)

ʔɔʔ	*katɨʔ*	*biʔ=hayet*	*ʔɛʔ*	*sɔʔ, (…)*	*ʔaʔ*	*katɨʔ*	
RECOG	skin	PL=be.yellow	3	die	CONTR	skin	
miʔ	*batɛ̃ŋ,*	*hay*	*hiy*	*m-ɲɔk,*	*ʔɛn*	*bah*	*sɔʔ*
Maniq	be.black	like	1D	PROG-sit	DEM	NEG	die

'If our skin became yellow, we'd die, (…) but with black skin, like this, we don't die.'

The sun-related threat is mitigated by a ritual involving burning hair of forest animals and release of a good smell *caɲɛs* with power to dispel the danger (Wnuk & Majid, 2014). Color is thus interwoven with other percepts and constitutes an important component of notions central to Maniq belief system.

7. General discussion

Color in Maniq is well lexicalized, codable, and expressed efficiently. These findings largely go against the expectation that non-industrialized societies with limited color technology should have small color lexica and limited ability to communicate color efficiently. As our analysis shows, however, Maniq is not alone among groups with a similar profile in having a large and efficient color lexicon. There are other hunter-gatherer languages which have six, seven, or more high frequency color terms and, similarly, other hunter-gatherer languages which have relatively high coding efficiency given their number of color terms. In addition, hunter-gatherers seem to constitute an internally varied grouping in terms of color language and this variation does not seem to map onto variation in technology. This does not deny the possibility that technology can stimulate lexical growth of color, but together with earlier work (Majid et al., 2018), our data suggest the exact relationship between technology and color language needs to be examined more closely.

Maniq constitutes an especially clear example challenging the tight link between technology and color language—despite limited color technology, it codes color efficiently and has rich lexical color resources. An opposite case has also been reported. Kata Kolok—a rural sign language of North Bali is used by a community with complex culturally embedded color technology, yet it has a rudimentary color lexicon with low codability and low efficiency (Fig. 1 and Fig. 2) (de Vos, 2011). Together, these extreme cases show technology by itself is not a decisive force in shaping color language. They also point to the complex nature of color term evolution and the fact that the same factors can be associated with different outcomes depending on the broader context. For instance, Kata Kolok has been argued to lack the pressure to develop a more complex color lexicon due to the small size and tightly knit nature of the community, but this is not true of many spoken language communities of similar size and structure which have evolved larger color systems (de Vos, 2011). Similarly, language contact is not a significant evolutionary pressure for the Kata Kolok color lexicon, but it has been a crucial change-driving force for numerous other languages, including the previously reported cases of Kilivila (Senft, 2011), Gyeli (Grimm, 2014), and cases of Mlabri, Jahai, and Semaq Beri discussed here.

Although language contact does not always impact color lexicons, it is an important force to reckon with in the evolution of color naming systems. Lexical data in the World Loanword Database (WOLD) (Haspelmath & Tadmor, 2009) suggest borrowing of color terms, in particular late-stage terms could be relatively common cross-linguistically. The term for "blue", for instance, has been borrowed in about a quarter of the 41 surveyed languages, including English (for details; see Biggam, 2006). The examination of Maniq vis-à-vis its relatives shows that borrowing impacts not just the size of the lexicon, but also the general efficiency of the system. While Mlabri, Jahai, and Semaq Beri have all expanded their lexicons with borrowed color terms, many of these terms are not high-frequency and are used with low agreement, suggesting their meaning has yet to stabilize. The general efficiency with which colors can be communicated in these languages is therefore relatively low. This means that even though language contact can stimulate lexical growth of color terms, it does not automatically lead to more efficient communication.

The Maniq data further contradict the idea that societies in non-industrialized settings do not need color terms because other vocabulary, e.g., ethnobotanical terminology, obviate the need for such terms (Gibson et al., 2017). While intuitively appealing, this view is problematic because—contrary to stereotypical portrayals, ethnobiological nomenclatures of hunter-gatherers are not as rich as typically assumed. In fact, ethnobiological research has shown the inventories of labelled plants and animals of hunter-gatherers are on average three times smaller than those of small-scale cultivators (Brown, 1985), so ethnobiological vocabulary alone does not suffice to distinguish between some plants and animals. For instance, there are several hundred butterfly species in the Khao Banthad area where the Maniq live (Basset et al., 2013), but in Maniq most butterflies are referred to with the same general label *ŋwãk* 'butterfly, moth'. Even in cases where specific names for different species exist, e.g., for snakes (*yəkɔp*) and for birds (*kawaw*), people still use generic names in many everyday contexts, unless talking about particularly salient species. It is thus part of routine language use to refer to different animals with generic labels combined with color terms (e.g., *yəkɔp hayet* 'yellow snake', *kawaw bagɛ̃ɛ̃c* 'red bird'). In addition, while certain plants and animals are associated with a single predictable color, it is quite common for plants and animals to vary in color with different stages of maturation and as a result of genetic variation within species/population (color polymorphism) (White & Kemp, 2016), e.g., floral colors, diverse coloring in birds, butterflies, frogs, and toads. This type of variation entails lack of color predictability, which is meaningful because, as Kay and Maffi (1999) point out, when color is not fully predictable, it is more informative and more relevant in communication.

Beyond the sheer existence of such variation, however, it is also relevant that these plants and animals are actually talked about. Based on raw color statistics of the environment, especially taking into account the exuberant variation in coloring of butterflies, birds and frogs, one might predict an even richer color system in Maniq. Critically, it matters how people *engage* with the environment. In the Maniq context, variation in color concerns some of the most culturally salient biological taxa, including the traditional staple food—wild yams (described with basic terms for white, yellow, red, and purple)—and commonly hunted animals such as gibbons and macaques (described with basic terms for white, black, red and purple) and the Maniq system is well-suited to talking about these objects. Our data thus align with the suggestion that the community's communicative needs might be shaped by the environment (Josserand et al., 2021; Twomey, Roberts, Brainard, & Plotkin, 2021) in combination with local factors such as culture-specific preoccupations, yet again underscoring the importance of ethnographic work in color term research (Conklin, 1955; Levinson, 2000).

In addition, in real life many animals are often spotted momentarily or from a large distance and therefore may not be immediately identified, so color information becomes highly relevant. Accordingly, color is commonly mentioned to aid visual search, e.g., when talking about animals in trees or undergrowth. Thus, our Maniq data also put into question the proposition that discriminating between abstract colors is irrelevant in the context of naturally-colored objects (Wierzbicka, 2008) suggesting that while technology can stimulate color term growth, it is not a necessary condition for developing a large color lexicon.

Our long-term observational data suggest that unlike the Bellonese or Warlpiri (Kuschel & Monberg, 1974; Wierzbicka, 2008), the Maniq talk about color on a regular basis. Color terms (both basic and secondary) are widely used to describe aspects of the natural world, including when color is not the only variable aspect of the scene. These observations align with experimental psycholinguistic research in English and other languages showing color is a salient object property that is frequently mentioned in referential expressions, even when redundant (Pechmann, 1989; Tarenskeen, Broersma, & Geurts, 2015). Given that color is accessible in early stages of visual processing (Livingstone & Hubel, 1988) and speakers might be under time pressure when formulating utterances, specifying color can help initially confine the space of possibilities when zeroing in on a referent. This seems especially relevant for real-world environments, such as rainforest, where perceptual ambiguity abounds. However, even in low-noise lab settings, color words are still used more often than required for unambiguous reference. In fact, participants in unscripted director-matcher tasks often start producing color modifiers for critical objects *before* visually fixating on distractor objects, so color terms (unlike, for instance, size terms) are often applied promptly without evaluating their informativeness in the specific context (Brown-Schmidt & Konopka, 2011). This suggests color terms may be applied in a wide range of communicative contexts and our impressionistic observations from the field align with such observations.

Much of the color language literature has focused on the referential function of color and has tried to explain color term growth primarily in terms of its distinguishing function. Yet, as exemplified by the Maniq data, color is routinely mentioned not just to distinguish between objects, but also to characterize objects, structure knowledge, and symbolically represent cultural notions connected to color through a rich web of associations. Color language thus has wider applicability and is necessarily associated with a wider range of communicative pressures.

As such, it is similar to other lexical domains, notably ethnobiology, whereby a purely utilitarian perspective on vocabulary growth (i.e., one which views lexicalization as being driven by "practical consequences" of knowing or not knowing a referent; Hunn, 1982) does not fully account for extant lexicalized distinctions. As argued by Berlin (1992), lexical distinctions can be the result of inherent human interest and a drive to acknowledge perceptually prominent categories in language. A similar argument can be made for color. Lexicalization of color can be the result of the perceptual prominence of color, as suggested by Berlin, combined with cultural preoccupation with visual appearance of objects. While such preoccupation can take different forms (e.g., relate to chromatic or non-chromatic surface features, cf. Wierzbicka, 2008) and be expressed in diverse ways, it is not determined by practical usefulness alone.

Take for instance the following practice reported for the Aṉangu people of Australia. When driving through country, the Aṉangu "slow the car to examine groups of wild donkeys more closely and invariably remark on the different colours among the group of animals" (Young, 2011, p. 359). The interest in donkeys is linked to them being considered "fortuitous presence" due to being "marked with 'the cross of Jesus' in the fur pattern on their backs" (Young, 2011, p. 372). The color naming practice, however, does not have any immediate practical outcome, but serves to express shared interests and coordinate beliefs (cf. Enfield, 2022). To evoke a first-hand example, the Maniq often commented on the visual appearance of audio-recording equipment, jokingly likening a microphone to a dusky leaf monkey (*basiŋ*) and describing it as *kuy bayɔ̃ɔ̃ɸ* 'white/mottled gray head'. This spontaneous analogy, again, is not intended for practical purposes, but serves as an expression of knowledge the Maniq share with the researcher.

Though such examples of color term use have not received much attention in the color language literature, they are likely ubiquitous across cultural contexts. The focus on referential discrimination has overshadowed other uses of color terms. It has also biased discussions of communicative pressures towards objects "distinguishable only by color" (Kay & Maffi, 1999, p. 746), resulting in narrow interpretations of the role of technological complexity as sine qua non for lexical growth. Looking at color more broadly, however, we see that human preoccupation with color is not limited to its potential for distinguishing between similar objects, but often relates to the opposite, i.e., creating analogies between disparate objects (Young, 2013) (as in the leaf monkey example above). While it is ordinary practice for groups like the Maniq, or Aṉangu ("a culture where searching for and creating mimesis is paramount"; Young, 2011, p. 360), to remark on such analogies in everyday discourse, their broader linguistic expression remains understudied.

All in all, rich and efficient color language can emerge without industrialization and technological development. Color salience, while being promoted by color technology, does not rely on it, but can arise in the context of naturally-colored objects. Perceptual prominence of color, its practical usefulness across communicative contexts, and symbolic importance can all be factors that promote elaboration of color language.

More generally, this paper calls attention to the fact a language's vocabulary is an important cognitive tool. It is hard to formulate simple drivers of vocabulary elaboration, given the diversity of cultural interests and the many ways in which languages are used.

CRediT authorship contribution statement

Ewelina Wnuk: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing. **Annemarie Verkerk:** Formal analysis, Visualization, Writing – review & editing. **Stephen C. Levinson:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing. **Asifa Majid:** Conceptualization, Methodology, Resources, Formal analysis, Visualization, Supervision, Writing – review & editing.

Data Availability

Data and code are available at: https://zenodo.org/record/7085267#.YyRJgy8RqL8.

Acknowledgments

Thanks to Yuma Ito, Shu Nimonjiya, Jaroon Thotsagool, and Prajoub Kongkaew for help with logistics and recruitment of participants, and Delwin Lindsey for helpful discussions on GMI and providing us with a sample of Hadzane responses to test our GMI implementation. This work was funded by the Max Planck Gesellschaft and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 800004.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2022.105223.

References

Basset, Y., Eastwood, R., Sam, L., Lohman, D. J., Novotny, V., Treuer, T., … Osorio-Arenas, M. A. (2013). Cross-continental comparisons of butterfly assemblages in tropical rainforests: Implications for biological monitoring. *Insect Conservation and Diversity, 6*(3), 223–233. https://doi.org/10.1111/j.1752-4598.2012.00205.x

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. In *R package version 1.1-9*. https://CRAN.R-project.org/package=lme4.

Benjamin, G. (1985). In the long term: Three themes in Malayan cultural ecology. In K. L. Hutterer, A. T. Rambo, & G. W. Lovelace (Eds.), *Cultural values and human ecology in Southeast Asia* (pp. 219–278). Center for South and Southeast Asian Studies.

Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton University Press.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. University of California Press.

Bernatzik, H. A. (1938). *Die Geister der gelben Blätter. Forschungsreisen in Hinterindien (Auflage: 1)*. F. Bruckmann.

Biggam, C. P. (2006). Political upheaval and a disturbance in the colour vocabulary of Early English. In C. P. Biggam, C. Kay, & N. Pitchford (Eds.), *Progress in colour studies* (pp. 159–179). Benjamins.

Bishop, N., & Peterson, M. (2003). *Northern Aslian language survey: Trang, Satul and Phatthalung provinces*. Thailand: TU-SIL-LRDP Thammasat University.

Brandt, J. H. (1961). The negrito of Peninsular Thailand. *Journal of the Siam Society, 49* (2), 123–158.

Brown, C. H. (1985). Mode of subsistence and folk biological taxonomy [and comments and reply]. *Current Anthropology, 26*(1), 43–64.

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49*(3), 454–462. https://doi.org/10.1037/h0057814

Brown-Schmidt, S., & Konopka, A. E. (2011). Experimental approaches to referential domains and the on-line processing of referring expressions in unscripted conversation. *Information, 2*(2), 302–326. https://doi.org/10.3390/info2020302

Burenhult, N. (2005). *A grammar of Jahai*. Pacific Linguistics.

Burenhult, N. (2009). Semang archive: Survey, Malaysia. http://hdl.handle.net/1839/00-0000-0000-0009-79A4-7@view.

Burenhult, N., Kruspe, N., & Dunn, M. (2011). Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula. In N. J. Enfield (Ed.), *Dynamics of human diversity: The case of mainland Southeast Asia* (pp. 257–277). Pacific Linguistics.

Casson, R. W. (1994). Russett, rose, and raspberry: The development of English secondary color terms. *Journal of Linguistic Anthropology, 4*(1), 5–22.

Chulalongkorn. (1907). *Ngo Pa: Romance of the Sakai (Malithat Pramathatavedi, Trans.)*. Department of Curriculum and Instruction Development.

Conklin, H. C. (1955). Hanunóo color categories. *Southwestern Journal of Anthropology, 11*(4), 339–344.

Conway, B. R., Ratnasingam, S., Jara-Ettinger, J., Futrell, R., & Gibson, E. (2020). Communication efficiency of color naming across languages provides a new framework for the evolution of color terms. *Cognition, 195*, Article 104086. https://doi.org/10.1016/j.cognition.2019.104086

Crowell, A. (1992). Postcontact Koniag ceremonialism on Kodiak Island and the Alaska Peninsula: Evidence from the fisher collection. *Arctic Anthropology*, 18–37.

Davies, I. R. L., Macdermid, C., Corbett, G. G., Mcgurk, H., Jerrett, D., Jerrett, T., & Sowden, P. (1992). Color terms in Setswana: A linguistic and perceptual approach. *Linguistics, 30*(6), 1065–1104.

Dentan, R. K. (1979). *The Semai: A nonviolent people of Malaya*. Rinehart and Winston: Holt.

Endicott, K. M. (1979). *Batek negrito religion: The world-view and rituals of a hunting and gathering people of Peninsular Malaysia*. Clarendon Press.

Enfield, N. J. (2022). *Language vs. reality: Why language is good for lawyers and bad for scientists*. The MIT Press.

Epps, P., Bowern, C., Hansen, C. A., Hill, J. H., & Zentz, J. (2012). On numeral complexity in hunter-gatherer languages. *Linguistic Typology, 16*(1), 41–109. https://doi.org/10.1515/lity-2012-0002

Evans, I. H. N. (1937). *The negritos of Malaya*. The University Press.

Francis, C. M. (2001). *A photographic guide to mammals of South-East Asia*. New Holland Publishers.

Gibson, E. (2017). *Language lab MIT Disqus profile*. Disqus. https://disqus.com/by/langua gelabmit/.

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., … Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences, 114*(40), 10785–10790.

Graburn, N. H. H. (2006). Arts of the fourth world. In M. Perkins, & H. Morphy (Eds.), *The anthropology of art: A reader* (pp. 412–430). Blackwell.

Grimm, N. (2014). Color categories in language contact: 'Pygmy' hunter-gatherers and Bantu farmers. In K. Carpenter, et al. (Eds.), *Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society* (pp. 31–46). Berkeley Linguistics Society.

Haspelmath, M., & Tadmor, U. (Eds.).. (2009). *WOLD*. Max Planck Institute for Evolutionary Anthropology. http://wold.clld.org/.

Hayden, B. (1998). Practical and prestige technologies: The evolution of material systems. *Journal of Archaeological Method and Theory, 5*(1), 1–55.

Headland, T. (1993). Agta. In P. Hockings (Ed.)*, Vol. 5. Encyclopedia of world cultures* (pp. 4–6). G. K. Hall.

Higgs, S. (2016). Martu paint country. The archaeology of colour and aesthetics in Western Desert rock art and contemporary acrylic art [PhD dissertation]. Australian National University.

Hill, C. (2011). Named and unnamed spaces: Color, kin, and the environment in Umpila. *The Senses and Society, 6*(1), 57–67. https://doi.org/10.2752/174589311X12893982233759

Huisman, J. L., van Hout, R., & Majid, A. (2022). Stability and change in the colour lexicon of the Japonic languages. *Studies in Language, 46*(2), 323–351. https://doi.org/10.1075/sl.20032.hui

Hunn, E. (1982). The utilitarian factor in folk biological classification. *American Anthropologist, 84*(4), 830–847.

Josserand, M., Meeussen, E., Majid, A., & Dediu, D. (2021). Environment and culture shape both the colour lexicon and the genetics of colour perception. *Scientific Reports, 11*(1), 19095. https://doi.org/10.1038/s41598-021-98550-3

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The world color survey*. Center for the Study of Language and Information Publications.
Kay, P., & Maffi, L. (1999). Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist, 101*(4), 743–760.
Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language, 54*(3), 610–646. https://doi.org/10.2307/412789
Kelly, R. L. (2013). *The lifeways of hunter-gatherers: The foraging spectrum (2nd ed)*. Cambridge University Press.
Kricheff, D. A. (2019). *Action and identity among Maniq people in Southern Thailand* [PhD dissertation]. University College London.
Kruspe, N., Burenhult, N., & Wnuk, E. (2015). Northern Aslian. In P. Sidwell, & M. Jenny (Eds.), *Handbook of the Austroasiatic languages* (pp. 419–474). Brill.
Kuschel, R., & Monberg, T. (1974). "We don't talk much about colour here": A study of colour semantics on Bellona Island. *Man, 9*(2), 213–242.
Lee, R. B., & Daly, R. H. (1999). *The Cambridge encyclopedia of hunters and gatherers*. Cambridge University Press.
Levinson, S. C. (2000). Yélî Dnye and the theory of basic color terms. *Journal of Linguistic Anthropology, 10*(1), 3–55.
Lindsey, D. T., Brown, A. M., Brainard, D. H., & Apicella, C. L. (2015). Hunter-gatherer color naming provides new insight into the evolution of color terms. *Current Biology, 25*(18), 2441–2446. https://doi.org/10.1016/j.cub.2015.08.006
Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science, 240*(4853), 740–749. https://doi.org/10.1126/science.3283936
Luangthongkum, T. (1985). Minority languages of Thailand. *Science of language papers 5: Languages and dialects* (pp. 29–74). Chulalongkorn University.
Majid, A. (2008). Focal colours. In A. Majid (Ed.)*, volume 11. Field manual* (pp. 8–10). Max Planck Institute for Psycholinguistics.
Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition, 130*(2), 266–270. https://doi.org/10.1016/j.cognition.2013.11.004
Majid, A., & Kruspe, N. (2018). Hunter-gatherer olfaction is special. *Current Biology, 28* (3), 409–413.e2. https://doi.org/10.1016/j.cub.2017.12.014
Majid, A., & Levinson, S. C. (2007). The language of vision I: Color. In *, volume 10. Field manual*. Max Planck Institute for Psycholinguistics.
Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., O'Grady, L., … Levinson, S. C. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences, 115*(45), 11369–11376. https://doi.org/10.1073/pnas.1720419115
Maneenoon, K. (2001). *Ethnobotany of Sakai tribe in Trang, Phatthalung and Yala provinces [Master's thesis]*. Prince of Songkla University.
Maneenoon, K. (2008). *A study on the genus Dioscorea L. (Dioscoreaceae), a major food plant of Sakai tribe at Banthad Range, Peninsular Thailand [PhD dissertation]*. In Prince of Songkla University.
Naroll, R. (1970). What have we learned from cross-cultural surveys? *American Anthropologist, 72*(6), 1227–1288. https://doi.org/10.1525/aa.1970.72.6.02a00030
Needham, R. (1967). Blood, thunder, and mockery of animals. In J. Middleton (Ed.), *Myth and cosmos: Readings in mythology and symbolism*. Natural History Press.
Nimonjiya, S. (2013). From 'ghosts' to 'hill tribe' to Thai citizens: Towards a new history of the Mlabri in Northern Thailand. *Aséanie, 32*, 155–176.
Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics, 27*(1), 89–110. https://doi.org/10.1515/ling.1989.27.1.89
R Core Team. (2015). R: A language and environment for statistical computing. www.R-project.org.
Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences, 104*(4), 1436–1441. https://doi.org/10.1073/pnas.0610341104
Regier, T., Kay, P., & Khetarpal, N. (2009). Color naming and the shape of color space. *Language, 85*(4), 884–892. https://doi.org/10.1353/lan.0.0177
Roscoe, P. (2002). The hunters and gatherers of New Guinea. *Current Anthropology, 43* (1), 153–162.
Senft, G. (2011). Talking about color and taste on the Trobriand Islands: A diachronic study. *The Senses and Society, 6*(1), 48–56. https://doi.org/10.2752/174589311X12893982233713
Simpson, E. H. (1949). Measurement of diversity. *Nature, 163*, 688. https://doi.org/10.1038/163688a0
Skeat, W. W., & Blagden, C. O. (1906). *Pagan races of the Malay Peninsula*. Macmillan.
Tarenskeen, S., Broersma, M., & Geurts, B. (2015). Overspecification of color, pattern, and size: Salience, absoluteness, and consistency. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01703
Taylor, J. R. (2003). *Linguistic categorization*. Oxford University Press.
Twomey, C. R., Roberts, G., Brainard, D. H., & Plotkin, J. B. (2021). What we talk about when we talk about colors. *Proceedings of the National Academy of Sciences, 118*(39). https://doi.org/10.1073/pnas.2109237118
de Vos, C. (2011). Kata Kolok color terms and the emergence of lexical signs in rural signing communities. *The Senses and Society, 6*(1), 68–76. https://doi.org/10.2752/174589311X12893982233795
White, T. E., & Kemp, D. J. (2016). Colour polymorphism. *Current Biology, 26*(13), R517–R518.
Wierzbicka, A. (2008). Why there are no 'colour universals' in language and thought. *The Journal of the Royal Anthropological Institute, 14*(2), 407–425. https://doi.org/10.1111/j.1467-9655.2008.00509.x
Wnuk, E. (2016). *Semantic specificity of perception verbs in Maniq* [PhD dissertation]. Radboud University.
Wnuk, E., & Burenhult, N. (2014). Contact and isolation in hunter-gatherer language dynamics: Evidence from Maniq phonology (Aslian, Malay Peninsula). *Studies in Language, 38*(4), 956–981. https://doi.org/10.1075/sl.38.4.06wnu
Wnuk, E., & Majid, A. (2014). Revisiting the limits of language: The odor lexicon of Maniq. *Cognition, 131*(1), 125–138. https://doi.org/10.1016/j.cognition.2013.12.008
Woodburn, J. (1982). Egalitarian societies. *Man, 17*(3), 431–451.
Young, D. (2011). Mutable things: Colours as material practice in the northwest of South Australia. *The Journal of the Royal Anthropological Institute, 17*(2), 356–376. https://doi.org/10.1111/j.1467-9655.2011.01684.x
Young, D. (2013). The colours of things. In C. Y. Tilley, W. Keane, S. Küchler, M. Rowlands, & P. Spyer (Eds.), *Handbook of material culture* (Paperback ed., pp. 173–185). SAGE.
Zaslavsky, N., Kemp, C., Regier, T., & Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences, 115*(31), 7937–7942.
Zaslavsky, N., Kemp, C., Tishby, N., & Regier, T. (2019). Color naming reflects both perceptual structure and communicative need. *Topics in Cognitive Science, 11*(1), 207–219. https://doi.org/10.1111/tops.12395