A Colorful Formalization of the Typological Prevalence Hypothesis

Jules Torgue[†] **(jtorgue@nolots.net)**
Matthew Spike[‡] **(matspike@gmail.com)**
Francis Mollica[†] **(mollicaf@gmail.com)**
[†]School of Informatics, University of Edinburgh, UK
[‡]Centre for Language Evolution, University of Edinburgh, UK

Abstract

Languages vary in the way they categorize semantic domains. Incidentally certain semantic systems appear more often than others across the world. Recent research has shown that the attested variability can be explained as the result of languages being a plurality of optimal solutions to efficiency constraints. However, the question of the prevalence remains open. Assuming that languages are a form of culturally transmitted cognitive technology, the Typological Prevalence Hypothesis proposes that the prevalence of a linguistic system is explained by how cognitively natural it is to learn and use. We aim to formalize and test this hypothesis by proposing an information-theoretic measure of communicative and developmental naturalness applied to color typology. While controlling for phylogenetic relatedness, we find that both communicative and developmental naturalness are important predictors of typological prevalence.

Keywords: semantic typology; efficient communication; language acquisition; color perception

Introduction

Characterizing the diversity of the world's languages has been a foundational puzzle in linguistics. Despite diversity in environments, goals and cultural history, why do some features of language appear universal? At the same time, if we enumerate the logically possible linguistic systems, very few of them are attested in the world's languages. Why? Recent work has suggested that the languages of the world reflect diverse, yet optimal, solutions to a common set of communication problems (e.g., Kemp, Xu, & Regier, 2018; Mollica et al., 2021).

When we make the assumption that languages efficiently solve these communication problems, efficiency considerations define a *plurality* of optimal communication systems, trading off the complexity and accuracy of the communication system (e.g., Kemp et al., 2018). Using this framework, linguists have begun to explain some universals and, to a large extent, characterize the diversity of attested communication systems. Specifically, we know which systems are likely to exist and, to a first approximation, appear to explain the typology of those systems in terms of communicative need and conceptual structures. However, we don't know: why are some solutions more prevalent than others?

While there are many approaches to language evolution, we approach language evolution using the same lens applied to other cognitive technologies (Heyes, 2018). A cognitive technology is socially learned and, thus, culturally evolves, shaped by acquisition pressures, social transmission and the goals they help us achieve. The importance of acquisition on typological prevalence has not been lost on language acquisition researchers (Gentner & Bowerman, 2009). Typological prevalence is thought to reflect psychological naturalness, which predicts acquisition. Specifically, the *Typological Prevalence Hypothesis* states: "All else being equal, within a given domain, the more frequently a given way of categorizing is found in the languages of the world, the more natural it is for human cognizers, hence the easier it will be for children to learn" (Gentner & Bowerman, 2009). This has been evidenced empirically in the domains of space and evidentiality (e.g., Gentner & Bowerman, 2009; Saratsli, Bartell, & Papafragou, 2020).

At first glance, the TPH may appear circular; however, the framing is likely to due to the developmental hypotheses that were under investigation. The underlying logic is that linguistic systems that are difficult to learn should die out and, thus, the prevalence of a linguistic system should be proportional to its ease of acquisition. That said, recent modelling work on color and quantifiers has demonstrated that for a large range of the communicatively optimal frontier of possible languages, there is no difference in the ease of acquisition between languages for models that reflect human learning patterns (Steinert-Threlkeld & Szymanik, 2019, 2020; Steinert-Threlkeld, 2021; Gyevnar, Dagan, Haley, Guo, & Mollica, 2022).

Leaping off the definition of a cognitive technology, we know that (1) languages have to be learnable and (2) languages have to be socially transmitted. We can view each of these components as representational constraints on a cognitive technology. We want a representation that is efficient both to learn and to communicate. This is a Rate-Distortion (RDT) problem (Berger, 2003) with each constraint and the relative importance of constraints implying a plurality of optimal solutions. We plan on using the equivalences between acquisition and communication under a RDT formalization (e.g., Gyevnar et al., 2022) to motivate measures of naturalness that will predict how often a particular system is used and provide a new explanatory framework for language evolution, and the cultural evolution of cognitive technologies more generally.

In this paper, we illustrate our proposal using color. First, we describe typological trends in color and how we calculate the prevalence of different color systems. Then, we introduce

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

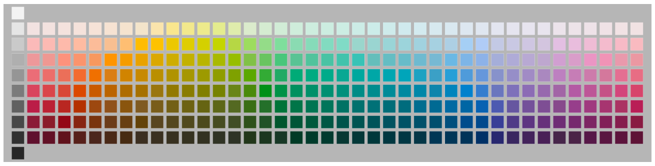

Figure 1: The Munsell stimulus palette used by the WCS.

our proposal for naturalness. Using phylogenetic regression, we compare several hypotheses about the influence of developmental and communicative naturalness on prevalence, and find evidence for an interaction such that communicatively and developmentally natural systems are more prevalent.. We finish by discussing the limitations of our modelling assumptions and outlining future directions.

Color Typology

In the late 1900s, linguistic anthropologists began a massive effort to collect color-naming data in non-industrialized societies. In each society, approximately twenty-five informants were asked to provide a name for all 330 paint chips in the Munsell color system (see Figure 1). The compiled dataset across 110 languages was released as the Word Color Survey (WCS; Cook, Kay, & Regier, 2005). Based on an examination of an earlier subset of the data, Berlin and Kay (1969) proposed several patterns for the evolution of color systems. For example, if a language has only two basic color terms, it should separate "warm" colors from "cool" colors.

In Zaslavsky, Kemp, Regier, and Tishby (2018, henceforth ZKRT), the authors explore the hypothesis that attested color systems in the WCS answer to a general principle of efficient communication. By formalizing this hypothesis as a RDT problem, the authors show that attested color systems achieve near optimal communicative efficiency trade-offs, the variation in the trade-off characterizes the diversity in attested color systems and as we move along the pareto-frontier of efficient systems, the optimal color systems recapitulate many of the patterns proposed by Berlin and Kay (1969).

In order to use the WCS to test the typological prevalence hypothesis, we need to devise a method for grouping similar languages. The real challenge is that the dimensionality of the perceptual space, expressed by 330 color chips, makes it effectively impossible to find two languages that have the same mappings chip for chip, even though they may operate similarly in practice. Thus, we develop a framework to compare and group languages so that we may compute the prevalence of different color systems.

Calculating Prevalence

Following ZKRT, we represent a language as a conditional probability distribution $P(w|u)$ mapping color chips u and words w (see next section for more details). As a result, we can measure the *informational similarity* between two languages using the generalized Normalized Information Distance (gNID), which captures the extent to which words can

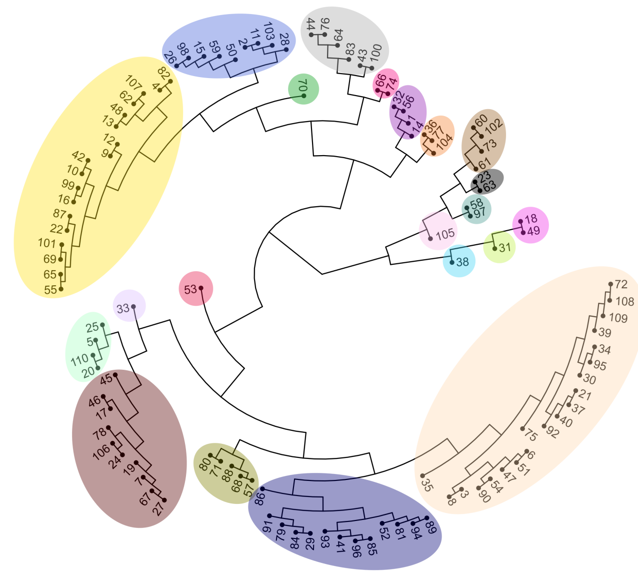

Figure 2: The agglomerative clustering of color systems. The length of branches is representative of the gNID. The shading denotes the 21 clusters validated by GED. The numbers correspond to the WCS identifiers of the 110 languages, see Figure 6 or `osf.io/wextd/` for the corresponding names.

be mapped across languages[1] (Zaslavsky et al., 2018). While we would like to cluster words according to consistent cross-language mappings, we also would like to ensure the mappings respect the topological structure of a color system's perceptual space, which is not required by gNID. For example, three languages might have sufficiently aligned words to be clustered together under gNID; however language $\mathcal{L}_1$ and language $\mathcal{L}_2$ align words in the pink-orange part of the space and language $\mathcal{L}_1$ and language $\mathcal{L}_3$ align words to the same extent in the green-blue part of the space. To better account for color topology, we also calculate a coarse measure of *topological similarity*. First, we created graph structures for each language by mapping edges between each chip to its modal term according to the WCS and between each word to adjacent words in chip space. We then compute the Graph Edit Distance (GED) (Abu-Aisheh, Raveaux, Ramel, & Martineau, 2015) between each language graph. Languages that share topological structure will have smaller GEDs.

To cluster languages, we first compute an agglomerative clustering tree based on informational similarity using the gNID (Figure 2). We then incorporate topological similarity into how we determine the cut-off value for clusters. Following the intuitions of Prototype Theory (Rosch, 1973, 1978), we assume that an optimal partitioning scheme would maximize the in-cluster similarity while minimizing the between-clusters similarity. We use the GED to compute this similarity trade-off at different cut-off values $[\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7},$

[1]To calculate gNID, we use the least informative prior following (Zaslavsky et al., 2018) over color chips.

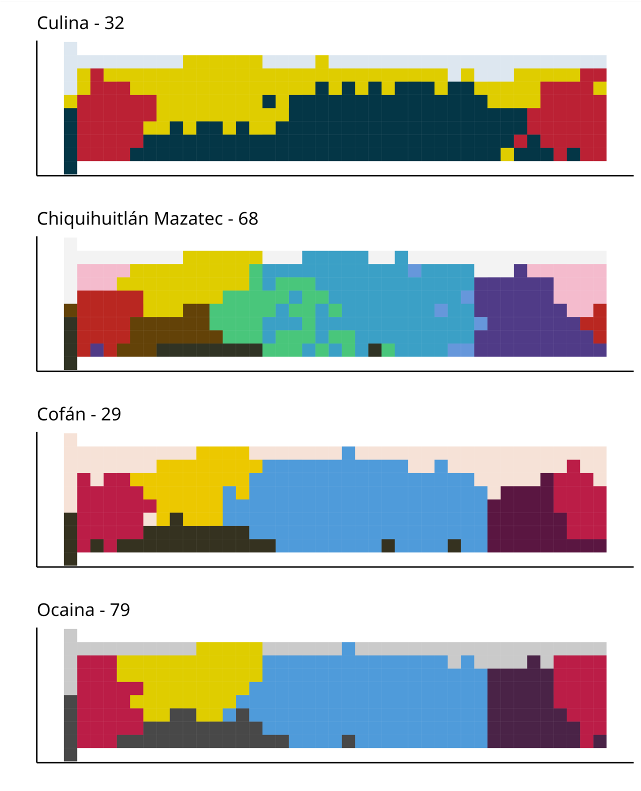

Figure 3: The mode maps show how a language structures the perceptual space. Chips are colored based on their mode answer. Culina and Mazatec are each grouped independently. Cofán and Ocaina are grouped together in the same cluster.

$\frac{1}{8}$, $\frac{1}{9}$]. The best trade-off was achieved at a cut-off value of $\frac{1}{7}$ which produces 21 clusters whose sizes range from 1 to 20 languages, with an average of 5.24 languages per cluster. The prevalence of a color system is given by the size of its cluster.

To briefly illustrate our clustering, Figure 3 reports the color maps of four languages. Culina and Mazatec were assigned to two different clusters, while Cofán and Ocaina were grouped together in another cluster. More details about our clusters are provided here: `osf.io/wextd/`.

Formalizing Naturalness

Our general approach is that cognitive representations $\mathcal{X}$ are abstractions over the states of the world $\mathcal{U}$. The optimal way to represent the world is dependent on the task that you hope to achieve with your representation (Markman, 1998). For example, conceptual development can be equated with learning a representation that allows you to re-construct the world subject to a constraint on representational complexity (Rosch, 1978; Ullman & Tenenbaum, 2020). Social transmission can be equated with learning a representation that allows a speaker to reconstruct a speaker's intended meaning subject to a constraint on the complexity of the language (Kemp & Regier, 2012; Zaslavsky et al., 2018). We can think of a given cognitive representation x as a probability distribution over the states of the world $P(\mathcal{U}|x)$. Similarly, we can think of a word's meaning as distribution over the states of the world[2] $Q(\mathcal{U}|w)$. We then define the unnaturalness of a language as the total amount of information lost by using a word w instead of another cognitive representation x to encode the world, measured as the Kullback-Leibler divergence, assuming the best possible mapping (minimal loss of information) between words and representations. Formally,

$$\mathcal{N}(\mathcal{W},\mathcal{X}) = \sum^{\mathcal{X}} \min_{w} \mathrm{KL}[P(u|x)||Q(u|w)]. \quad (1)$$

Following our proposed hypotheses, we will motivate cognitive representations—i.e., some distribution $P(\mathcal{U}|x)$, from communicative and developmental goals, where the representation corresponds to speakers' intentions and a learners' hypotheses respectively.

Communicative (Perceptual) Naturalness

In the context of a communication game, we argue that the relevant representations to be encoded into words reflect perceptual precision, following ZKRT. Therefore, we adopt the distributions involved in communicative naturalness following the same assumptions about communication and the perceptual space as ZKRT's framework. For full details of the communication model see Zaslavsky et al. (2018).

In communication, speakers have intentions s that they are trying to convey. These intentions are the cognitive representations $P(u|s)$ that we will use for determining naturalness. Formalizing communication as a Rate Distortion Problem, ZKRT demonstrated that the diversity of the world's color systems can be explained as efficient compression of these speaker intentions onto words. We follow their model in defining $P(u|s)$ as a Gaussian distribution in the color space centered on a given color chip, which captures that human color perception is continuous rather than *a priori* categorical (Regier, Kay, & Khetarpal, 2007; Regier, Kemp, & Kay, 2015; Zaslavsky et al., 2018). In the WCS, the set of possible speaker intentions corresponds to the 330 color chips.

Figure 4 shows the communicative naturalness calculation for Culina. The left column reflects the speaker intentions for each chip $P(u|s)$. The rightmost column is the word meanings $P(u|w)$ for each word in Culina. For each intention, we calculate the KL divergence between that intention and every word, which allows us to determine which word best maps to each intention by taking the minimum. We then sum these minimum values across intentions as our measure of naturalness. Lower values reflect greater naturalness.

[2]We define $Q(\mathcal{U}|w)$ by marginalizing over a speaker's intentions s when uttering a word, where intentions are themselves a distribution over states of the world $P(u|s)$. Formally, $P(u|w) = \sum_s P(u|s)P(s|w)$ where $P(s|w)$ is an application of Bayes Rule using the WCS languages and least informative need distribution (see Zaslavsky et al., 2018).

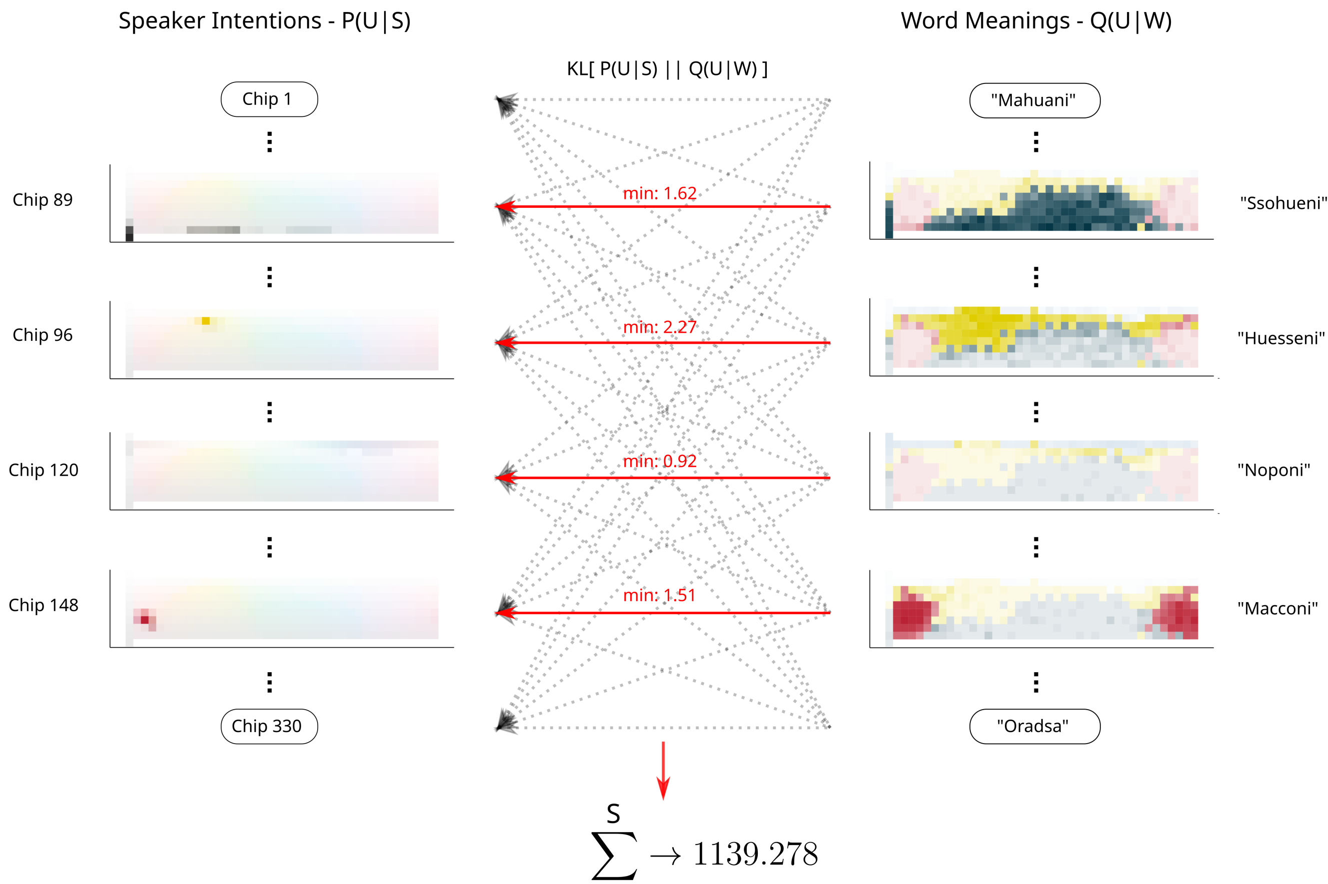

Figure 4: Calculation of communicative naturalness for Culina. The colors chips are displayed with varying opacity to visualise their probability mass in their respective distributions. For each chip, we find the mapping to the word that minimizes the KL divergence between the speaker intention $P(u|s)$ and the word meaning $P(u|w)$. We then sum these divergences to obtain a measure of the unnaturalness of the system. The lower the score the more natural the system.

Developmental Naturalness

Across development, learners have hypotheses h about the semantic categories in their culture. These hypotheses $P(u|h)$ are the cognitive representations we will use for developmental naturalness. As noted in Gyevnar et al. (2022), development can be viewed as a RDT problem where learners compress their experiences onto hypotheses about semantic categories, which allow them to reconstruct the relevant structures in the world. The RDT trade-off is between the complexity of the hypothesis and the fidelity with which a hypothesis reconstructs the environment. In an ideal learning model, this trade-off determines a developmental trajectory where a learner starts out with a simple hypothesis that over-extends a semantic category–i.e., has poor reconstruction; however, as more data is observed learners will converge on more complex hypotheses that better reconstruct the environment (e.g., Mollica & Piantadosi, 2021). As Beekhuizen and Stevenson (2018) demonstrated that a Self-Organizing Map (SOM; Kohonen, 1990) model captures the patterns of color term learning seen in humans, we will also use SOMs to model color development.

In brief, a SOM initializes a square of cells, with each cell containing a vector reflecting input to the model–i.e., color-word pairs. As a word is observed, the most similar cell is identified and updated based on the data point. Additionally, adjacent cells specified by a free parameter also update according to the data point. Over time, the free parameter decreases in magnitude resulting in less adjacent cells updating and more precise updates. In the limit of data, an expressive SOM will eventually converge on the input distribution. For illustration, see left panel of Figure 5. Recently, Gyevnar et al. (2022) implemented SOMs for all languages in the WCS. We use this re-implementation for the current work so see Gyevnar et al. (2022) for full details of the developmental model.

To calculate naturalness, we would like a set of hypotheses that capture the developmental stages exhibited by children. As it is difficult to tell when two distributions over color chips are the same, we decided to sample hypotheses at 124 different data amounts between 1 and 51000, with greater resolution at fewer data amounts (e.g., 1, 2, 3 . . . 50, 60, 70,

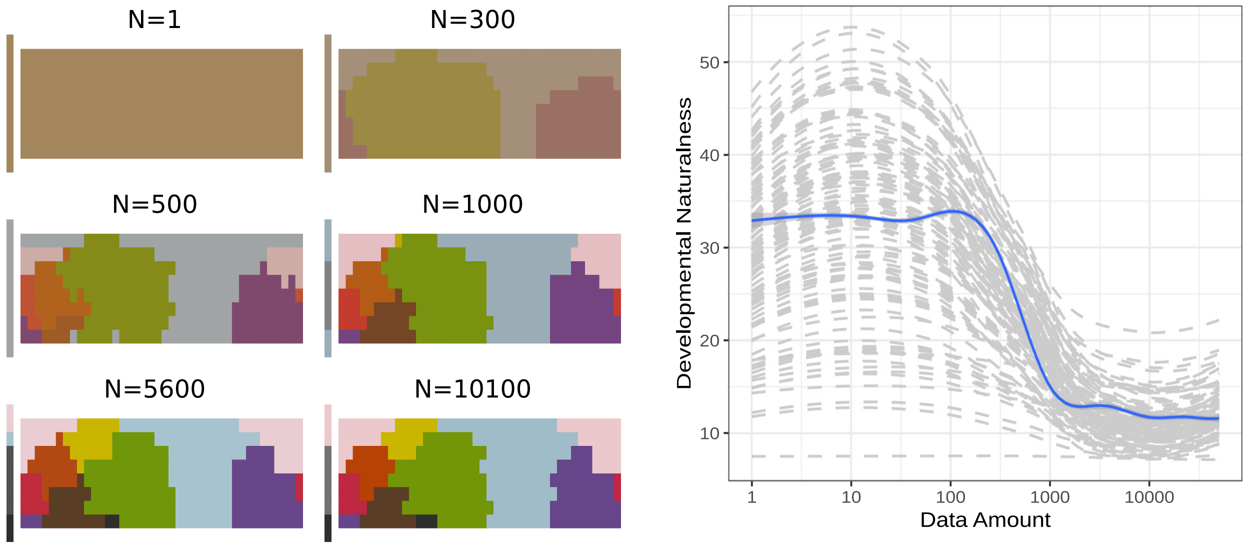

Figure 5: (left) Mode maps reflecting developmental hypotheses at different data amounts N for a model learning English. Each cell corresponds to the Munsell color chips as in Figure 1 with the fill color reflecting the centroid color of the modal word for the cell. (right) Developmental naturalness decreases as more data is observed; although, there is considerable variability in the naturalness of the trajectories. The dashed grey lines reflect individual languages. The solid line and ribbon reflect mean and standard error of all languages.

...1000, 1100, 1200...). As a result, we would have 124 hypotheses for each language $|\mathcal{H}| = 124$. To account for variations in data sampling when training the SOM, we simulate five different learners and average them for each language in the WCS. Naturalness is then calculated following Equation 1 and Figure 4 with the left-most column reflecting learners' hypotheses at different data amounts. As a check on our naturalness measure, we can see that naturalness of hypotheses increases over time (right panel of Figure 5).

Results

To evaluate whether our measures of naturalness are good predictors of typological prevalence, we conduct several phylogenetic mixed-effect regression models (de Villemereuil & Nakagawa, 2014) and compare them using k-fold (k=10) cross-validation to estimate predictive performance. As prevalence is count data, we assume a Poisson link function (Lawless, 1987).

Our baseline model embodies the hypothesis that the prevalence of a color system is guided by inheritance and random historical accidents, such that when languages diverge they retain the color system of their common ancestor despite changes in environments and goals. To take into account this ancestry, we compute the phylogenetic tree corresponding to the 110 languages of the WCS (Figure 6) using phylogenetic data provided by Glottolog 4.6 (Hammarström et al., 2022). Based on the tree, we compute a phylogenetic variance-covariance matrix (Garland & Ives, 2000) that will serve as a baseline random effect structure in all of our models. We also add a random intercept for each language to control for the possibility that languages themselves may introduce unexpected information. One could also look at these random intercepts as an estimate of over-dispersion.

To examine the influence of our measures of naturalness, we conduct four additional nested regression models: (1) fixed effect for communicative naturalness, (2) fixed effect for developmental naturalness, (3) fixed effects for both communicative and developmental naturalness and (4) fixed effects and interaction for both naturalness measures. In all analyses, we centered and scaled both naturalness measures.

We find the model with an interaction between communicative and developmental naturalness provides the best predictive fit (elpd=-340.3, SE=8.9) of the data compared to the model with no interaction term (elpd=-346.9, SE=7.7), the communicative fixed effect model (elpd=-354.8, SE=6.5), the developmental fixed effect model (elpd=-361.7, SE=6.6) and the baseline model (elpd=-363.7, SE=4.9). The conditional effects are illustrated in Figure 7. The more communicatively and developmentally natural a system (negative when scaled), the more prevalent the system. Surprisingly, the effect of communicative naturalness ($\beta = -0.94[-1.22 - -0.67]$) was larger than developmental naturalness ($\beta = 0.22[0.10 - 0.35]$), suggesting communicative precision plays an important role in the prevalence of a system.

Discussion

We set out to better understand why some types of linguistic system are more prevalent in the world's languages than others. We hypothesized that systems that better satisfy the constraints of cognitive technologies will be more prevalent. Cognitive technologies are the product of cultural evolution satisfying our need to build adaptive solutions to dynamic

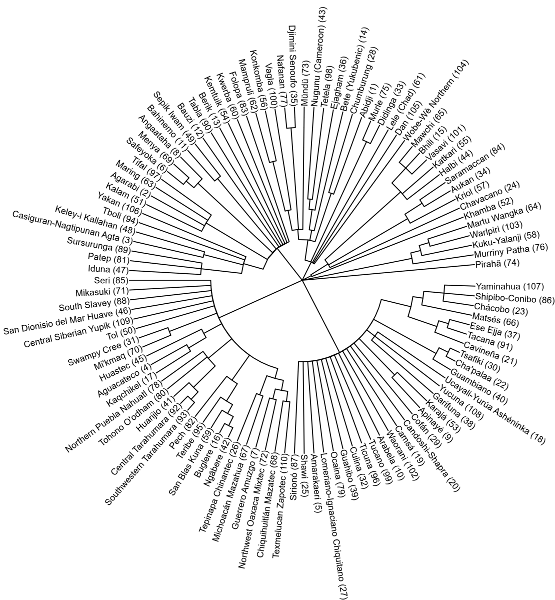

Figure 6: The phylogenetic tree of the WCS languages, computed using phylogenetic data from Glottolog 4.6 (Hammarström et al., 2022). In this visualisation the length of branches is not representative of phylogenetic distance.

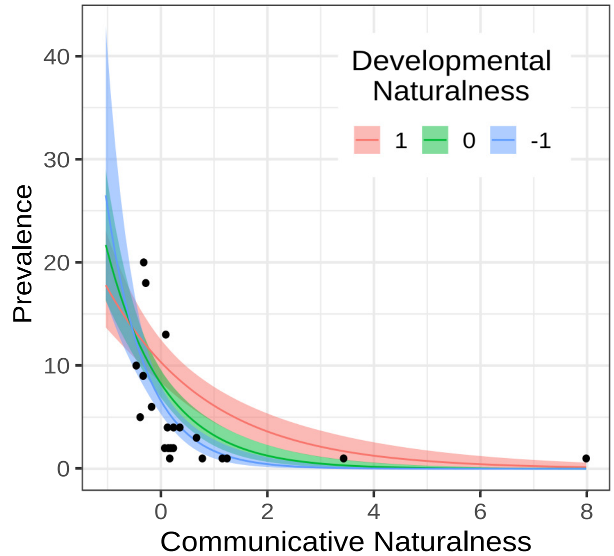

Figure 7: Conditional effects of naturalness on typological prevalence. Both developmental and communicative naturalness are centered and scaled. Developmental naturalness lines reflect mean and $\pm$ one standard deviation. Lines reflect median and shaded regions reflect 95% credible intervals. Points reflect prevalence and mean naturalness of each unique color system.

goals in rapidly changing environments. This cultural evolution requires efficient development and social transmission of cognitive technologies that successfully achieve goals. Developmentally, we don't have the time to wait for the world to deliver effective learning instances [3]. In terms of social transmission, this may be via communication systems or pedagogy (e.g., Shafto, Goodman, & Griffiths, 2014), likely both.

In this work, we formalized developmental and communicative efficiency constraints as measure of naturalness (that will also apply to other constraints on cognitive technologies). Our developmental measure is consistent with previous theorizing on the relation of developmental efficiency and prevalence (Gentner & Bowerman, 2009). We find evidence that color systems that are both developmentally and communicatively natural are more prevalent than less natural systems. Our results are consistent with the Typological Prevalence Hypothesis, but also suggest an important, underexplored role for communicative naturalness in explaining prevalence. Our results further suggest that there are functional pressures influencing prevalence of color systems beyond historical inheritance and "accidents of history" accounts of language evolution.

While we argue for communicative efficiency with regards to speaker intentions, there are alternative hypotheses articulated about how communicative efficiency could influence prevalence. One hypothesis is that how languages trade-off complexity and communicative accuracy is determined by the communicative need of the domain (Kemp et al., 2018). If a conceptual domain is regularly discussed (e.g., snow maintenance in Alaska), then the language will have greater communicative precision (Regier, Carstensen, & Kemp, 2016; Bradford, Thomas, & Xu, 2022) and, thus, be more complex. Under this account, typological prevalence would be predicted by the magnitude and consistency of the need for the conceptual domain across languages. This has yet to be evidenced; however, it would not be inconsistent with our proposal. In fact communicative need and perceptual naturalness (underlying our speaker intentions) has been argued to be particularly important in the domain of color (Zaslavsky, Kemp, Tishby, & Regier, 2019). Therefore, it will be important to look at how constraints on cognitive technologies influence their dynamics across other semantic domains.

Constraints on Generalization

While our framework is intended to generalize, our results are limited to the datasets and assumptions made in the paper. Critical among these are: (1) the WCS only provides data for a small sample of attested languages; (2) there are other plausible methods of clustering color term systems (e.g., Lindsey & Brown, 2009) (3) and our developmental models could not be evaluated against empirical data as none exists for most of the WCS languages. Future research is needed to determine the robustness of our results to these and our other assumptions.

[3] While there is a lot of data in the world, humans often learn from just a few instances (Mollica & Piantadosi, 2017; Koedinger, Carvalho, Liu, & McLaughlin, 2023)

References

Abu-Aisheh, Z., Raveaux, R., Ramel, J.-Y., & Martineau, P. (2015). An exact graph edit distance algorithm for solving pattern recognition problems. In *4th international conference on pattern recognition applications and methods 2015.*

Beekhuizen, B., & Stevenson, S. (2018). More Than the Eye Can See: A Computational Model of Color Term Acquisition and Color Discrimination. *Cognitive Science*, *42*(8), 2699–2734. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12665) doi: 10.1111/cogs.12665

Berger, T. (2003). Rate-distortion theory. *Wiley Encyclopedia of Telecommunications*.

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

Bradford, L., Thomas, G., & Xu, Y. (2022). Communicative need modulates lexical precision across semantic domains: A domain-level account of efficient communication. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44).

Cook, R. S., Kay, P., & Regier, T. (2005). The world color survey database. In *Handbook of categorization in cognitive science* (pp. 223–241). Elsevier.

de Villemereuil, P., & Nakagawa, S. (2014). General quantitative genetic methods for comparative biology. *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*, 287–303.

Garland, T., Jr, & Ives, A. R. (2000). Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *The American Naturalist*, *155*(3), 346–364.

Gentner, D., & Bowerman, M. (2009). Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In *Crosslinguistic approaches to the psychology of language: Research in the tradition of dan isaac slobin* (pp. 465–480). Psychology Press.

Gyevnar, B., Dagan, G., Haley, C., Guo, S., & Mollica, F. (2022). Communicative efficiency or iconic learning: Do acquisition and communicative pressures interact to shape colour- naming systems? *Entropy*, *24*(11), 1542. doi: 10.3390/e24111542

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2022, May). *Glottolog 4.6*. `http://glottolog.org`. (Accessed on 2023-02-01) doi: 10.5281/zenodo.6578297

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*(6084), 1049–1054.

Kemp, C., Xu, Y., & Regier, T. (2018). Semantic typology and efficient communication. *Annual Review of Linguistics*, *4*(1), 109–128.

Koedinger, K. R., Carvalho, P. F., Liu, R., & McLaughlin, E. A. (2023). An astonishing regularity in student learning rate. *Proceedings of the National Academy of Sciences*, *120*(13), e2221311120.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, *78*(9), 1464–1480. (Conference Name: Proceedings of the IEEE) doi: 10.1109/5.58325

Lawless, J. F. (1987). Negative binomial and mixed poisson regression. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 209–225.

Lindsey, D. T., & Brown, A. M. (2009). World color survey color naming reveals universal motifs and their within-language diversity. *Proceedings of the National Academy of Sciences*, *106*(47), 19785–19790.

Markman, A. B. (1998). *Knowledge representation*. Psychology Press: NY.

Mollica, F., Bacon, G., Zaslavsky, N., Xu, Y., Regier, T., & Kemp, C. (2021). The forms and meanings of grammatical markers support efficient communication. *Proceedings of the National Academy of Sciences*, *118*(49), e2025993118.

Mollica, F., & Piantadosi, S. T. (2017). How data drive early word learning: A cross-linguistic waiting time analysis. *Open Mind*, *1*(2), 67–77.

Mollica, F., & Piantadosi, S. T. (2021). Logical word learning: The case of kinship. *Psychonomic Bulletin & Review*, 1–34.

Regier, T., Carstensen, A., & Kemp, C. (2016). Languages support efficient communication about the environment: Words for snow revisited. *PloS one*, *11*(4), e0151138.

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.

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. *The Handbook of Language Emergence*, 237–263.

Rosch, E. (1973). Natural categories. *Cognitive psychology*, *4*(3), 328–350.

Rosch, E. (1978). Principles of categorization. In *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum: New York.

Saratsli, D., Bartell, S., & Papafragou, A. (2020). Cross-linguistic frequency and the learnability of semantics: Artificial language learning studies of evidentiality. *Cognition*, *197*, 104194.

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive psychology*, *71*, 55–89.

Steinert-Threlkeld, S. (2021). Quantifiers in natural language: Efficient communication and degrees of semantic universals. *Entropy*, *23*(10), 1335.

Steinert-Threlkeld, S., & Szymanik, J. (2019). Learnability and semantic universals. *Semantics and Pragmatics*, *12*, 4.

Steinert-Threlkeld, S., & Szymanik, J. (2020). Ease of learning explains semantic universals. *Cognition*, *195*, 104076.

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology*, *2*, 533–558.

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.