Word meanings across languages support efficient communication

Terry Regier, Charles Kemp, and Paul Kay

Introduction

Why do languages have the semantic categories they do? Each language partitions human experience into a system of semantic categories, labeled by words or morphemes, which are used to communicate about experience. These categories often differ widely across languages. Thus, languages do not merely provide different labels for the same universally shared set of categories—instead, both the labels and the categories themselves may be to some extent language-specific. However this cross-language variation is constrained. Words with similar or identical meanings often appear in unrelated languages, and most logically possible meanings are unattested—suggesting that there are universal forces constraining the cross-language diversity. Accounting for this pattern of *wide but constrained variation* is a central theoretical challenge in understanding why languages have the particular forms they do.

Previous approaches to this problem have often pursued a single semantic domain in detail, and have proposed solutions based on principles specific to that domain. For example, color naming has been viewed as constrained by a set of universal focal colors, and kinship categories as shaped by constraints specific to kinship. Here, we advance an account of semantic variation that is instead grounded in a simple functional principle that holds across domains: that language should support *efficient communication*. Specifically, following Rosch and others, we argue that good systems of categories are simple, which minimizes cognitive load, and informative, which maximizes communicative effectiveness. These two constraints compete against each other, and we propose that semantic systems in the world's languages tend to achieve a near-optimal tradeoff between these two constraints.

In what follows, we first briefly review existing work that is relevant to our proposal. We then develop a general-purpose computational framework that instantiates our proposal, and apply it to three domains with qualitatively different structures, shown in Figure 1.
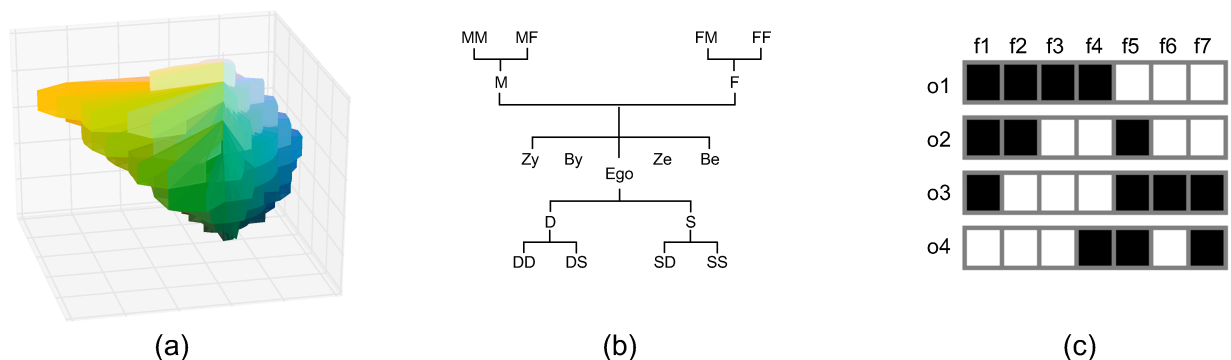

(a) (b) (c)

Figure 1: The three semantic domains we consider in this chapter: (a) color, (b) kinship, and (c) a domain in which objects are represented as binary feature vectors.

Our first analysis considers a continuous perceptual space and shows that our theory accounts for patterns in color naming across more than a hundred languages. Our second analysis considers a hierarchically structured network of relations, and shows that our theory accounts for patterns in kinship naming across

hundreds of languages. Our third analysis shows how our theory applies to domains in which objects are represented using binary feature vectors. We conclude from these three analyses that the tradeoff between simplicity and informativeness may provide a domain-general foundation for variation in category systems across languages. We close by discussing future research directions to further test this proposal.

Informativeness and simplicity as competing principles

We propose that attested semantic systems reflect the competing communicative principles of informativeness and simplicity. An informative system is one that supports precise communication; a simple system is one with a compact cognitive representation, which is therefore easy to learn, remember, and use. These two considerations trade off against each other. A maximally informative system would have a separate word for each object in a given semantic domain—e.g. a separate color term for each distinguishable hue—which would be complex and unwieldy. In contrast, a maximally simple system would have just one word for all objects in a given semantic domain—e.g. a single color term for all colors—and this would not support precise communication. In this sense, the two principles compete, and we propose that attested semantic systems are those that achieve a near-optimal tradeoff between these two competing principles—and thus achieve communicative efficiency.

Our proposal reflects and builds on a number of existing ideas. Most saliently, it reflects a general line of thinking that invokes *competing motivations* to explain why languages assume the forms they do (Du Bois, 1985; Bates & MacWhinney, 1982; Croft, 2003; Haiman, 2010). This line of thinking has a long history, and has often focused on the particular pair of competing motivations we consider here, as Haspelmath (1999: 180) points out:

> There is a long tradition in theoretical linguistics which holds that structural patterns of grammar are determined by highly general preferences or constraints that may come into conflict with each other. Gabelentz (1901: 256) was very clear about the tension between the "striving for ease" (*Bequemlichkeitsstreben*) and the "striving for clarity" (*Deutlichkeitsstreben*).

Simplicity satisfies the "striving for ease", and informativeness satisfies the "striving for clarity". Similar ideas appear in current theorizing. For example, the treatment of phonology in optimality theory (e.g. Prince & Smolensky, 1997) consists in the specification of a set of competing constraints, and a simple principle for resolving the competition (constraints are ranked, and higher-ranked constraints take precedence over all lower-ranked constraints). At an abstract level, the various competing phonological constraints can largely be grouped into two classes: *faithfulness* constraints and *markedness* constraints. Faithfulness constraints require a produced surface form to correspond as closely as possible to an underlying form—this can be seen as loosely analogous to informativeness or clarity, which requires that the information conveyed to a listener correspond as closely as possible to that intended by the speaker. Markedness constraints place structural requirements on the produced surface form, often minimizing the complexity of that form in some respect—this can be seen as broadly analogous to simplicity or ease. Although such ideas have not been widely applied to semantics, Jones (2010a; 2010b) has explored cross-language variation in the structuring of color and kinship in terms of optimality theory.

The general notion of competing motivations also arises in historical linguistics, as do the specific competing motivations of simplicity and informativeness (e.g. Hopper & Traugott, 2003: 71-72). An example from the history of English is illustrated by sentences like those below, from Bever and Langendoen (1972: 51):

a. ? There are lots of vulgar people live in Grosvenor Square. [Wilde]
b. ? It was haste killed the yellow snake. [Kipling]
c. ? Who is this opens the door? [Thackeray]
d. ? I wonder who it was defined man as a rational animal. [Wilde]

Such sentences were unremarkable as recently as the nineteenth century. Bever and Langendoen note that as long as English retained a rich inflectional system, including case marking on nouns (a form of complexity), the post verbal nominals in such examples could not be mistaken for the subject of a finite clause (communication would be unambiguous and thus informative)—and the long-standing practice of optionally dispensing with a complementizer or relative pronoun to introduce the relative clause could be maintained. However, the subsequent loss of nominal inflections seems to have created the necessity, for the sake of communicative clarity, to somewhat complicate the signal by instituting the requirement of a complementizer or relative pronoun to mark the relative clause boundary.

These examples illustrate the increasingly influential proposal that language is shaped by the need for *efficient communication*, which by its nature involves a tradeoff between simplicity and informativeness. Other examples—often formalized in information-theoretic terms—include the suggestions that word frequency distributions (Zipf 1949; Ferrer i Cancho & Solé, 2003), syllable durations (Aylett & Turk, 2004), word lengths (Piantadosi, Tily, & Gibson, 2011), syntactic structures and processing (Hawkins, 2004; Levy, 2008; Jaeger & Tily, 2010; O'Grady, 2005), and the learning of case marking (Fedzechkina, Jaeger, & Newport, 2012) all facilitate efficient communication.

The aspect of language most directly relevant to our proposal is the question why languages have the semantic categories they do. The semantic categories found in languages often seem to be "natural" ones—but natural in what sense? Rosch (1999) proposed that natural-seeming systems of categories are those that "provide maximum information with the least cognitive effort" (p. 190). Conceptually, our proposal is essentially the same as Rosch's. Our contribution is to formalize the proposal, and test it against a wide range of data, across languages and across semantic domains.

Computational formulation

In this section, we present our proposal in formal terms, building on the computational formulation of Kemp & Regier (2012). We discuss in turn each of the two competing principles, informativeness and simplicity, and indicate which aspects of each may be cast in domain-general terms, and which others require domain-specific treatment.

Informativeness

We take a communicative system to be informative to the extent that it supports *accurate mental reconstruction* by a listener of a speaker's intended message (cf. communication accuracy: Lantz & Stefflre, 1964). This idea can be illustrated through a simple communicative scenario, as in Figure 2 (cf. Baddeley & Attewell, 2009).

In this figure, time and causality flow from left to right. First, a speaker sees or otherwise experiences a specific target object t drawn from the universe U of all objects in a given semantic domain. For example, in
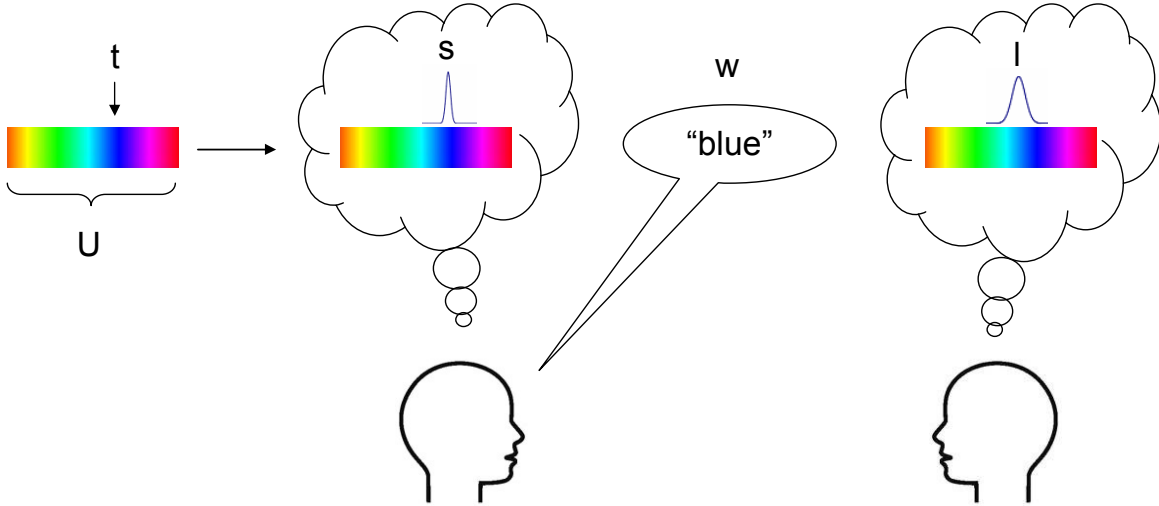
Figure 2: A scenario illustrating informative communication.

Figure 2, the speaker has seen a specific color t drawn from the set U of all colors, shown for simplicity as a 1-dimensional spectrum. The speaker mentally represents this target object as a probability distribution s over U, centered at t.[1] The speaker then attempts to convey this mental representation to a listener, by producing a word w—here, the word "blue". The listener wishes to understand the speaker's intended message, and to that end attempts to *mentally reconstruct* the speaker's representation s from the word the speaker has just used to describe it. The listener's mental reconstruction, l, is also a probability distribution over U, and is intended to approximate the speaker's mental representation. However, the listener has no direct access to the speaker's mental representation; she knows only that that mental representation was expressed using the word w. For this reason the listener's mental reconstruction l corresponds to the entire category named by w—in Figure 2, the category named by the word "blue"—which will in most cases be broader than s. We then define the reconstruction error $e(t)$ for the target object t as the dissimilarity between the speaker's and the listener's representations for that object:

$$e(t) = D(s||l) \tag{1}$$

$D(s||l)$ is the relative entropy, or Kullback-Leibler (KL) divergence, a standard measure of the dissimilarity between two probability distributions s and l:

$$D(s||l) = \sum_{i \in U} s(i) \log_2 \frac{s(i)}{l(i)} \tag{2}$$

The KL divergence $D(s||l)$ assumes that s is the true or actual distribution, and l is an approximation to s. This assumption holds in our communicative scenario above, because l represents the listener's mental reconstruction of the speaker's distribution s. $D(s||l)$ measures the information lost when using l as an approximation to s. In this way, the KL divergence (and therefore $e(t)$) captures the information lost in

---

[1]Modeling the speaker's mental representation as a probability distribution over objects, rather than as a single object, allows us to capture possible uncertainty in the speaker's mental representation of the target object. In our analyses below, we do not explore this potential—we instead assume that the speaker is certain of the target object. However the general formulation provided here enables the future application of the same ideas to analyses in which the speaker may have imperfect memory or perception of the target object.

communication. Smaller values for $e(t)$ denote less information lost in communication, and accordingly a closer approximation by the listener's representation to the speaker's representation.

For the analyses presented in this chapter, all probability distributions are discrete—although in some cases these are discrete approximations to continuous distributions. For this reason we use sums in these equations; analogous treatment of continuous distributions is also possible, in which case the sums would be replaced by integrals. For all of our analyses, we also assume that the speaker is certain of the specific target object $t \in U$ that is to be conveyed to the listener. In the case of speaker certainty with a discrete distribution, the speaker distribution is defined simply as $s(t) = 1$ for the target object t and $s(i) = 0$ for all other objects $i \neq t$ in the universe U. In this case, the Kullback-Leibler divergence $D(s||l)$ reduces to surprisal:

$$D(s||l) = \sum_{i \in U} s(i) \log_2 \frac{s(i)}{l(i)} = \log_2 \frac{1}{l(t)} = -\log_2(l(t)) \qquad (3)$$

This quantity measures how much information is lost by approximating the certain knowledge that the target object is t by the listener distribution l—and thus, how surprised the listener should be to find out that the target object is in fact t.

Finally, we need an aggregate measure of reconstruction error over all objects in the domain universe U. We define $n(t)$ as the need probability for target object t—that is, the probability that the speaker will wish to communicate about that object, rather than any other in the domain universe U. We then define the expected reconstruction error for the domain as a whole as the reconstruction error for all objects in the domain, weighted by the need probabilities for those objects:

$$E = \sum_{t \in U} n(t)e(t) \qquad (4)$$

E is a measure of the error, or information loss, incurred by using a given communicative system to communicate about a given semantic domain. We take a communicative system to be informative to the extent that it exhibits low E, and thus supports accurate mental reconstruction by the listener of the speaker's intended message.

The formulation presented above is domain-general, but applying it to a specific domain will require domain-specific instantiations of the speaker's and listener's distributions, and need probabilities. In each of our domain-specific case studies below—concerning color, kinship, and a domain represented in terms of binary feature vectors—these distributions are grounded in a characterization of the domain in question.

Simplicity

The core intuition behind the principle of simplicity is that simple communicative systems are advantageous because they require comparatively little cognitive effort to learn, remember, and use. This rationale is based on the simplicity of a *mental representation* of the communicative system (e.g. Chater & Vitányi, 2003). Thus, it is natural to formalize the simplicity of a given communicative system as the length or size of the mental representation that supports it. For example, in our kinship analyses below, we take the complexity of a given kinship system to be the smallest number of symbolic rules needed to define it, and we take simple systems to be those that can be defined using a small number of rules. In this approach, the general idea of defining simplicity in terms of the size of a representation is domain-general; however the specifics of that

representation and the means by which size is assessed will be domain-specific. This approach makes direct contact with the central intuition that motivates the simplicity principle.

However, there is also a limitation to this approach: it assumes a prior theoretical commitment to a particular representation language. In order to test the proposal while making as few assumptions as possible, one may wish to cast the idea of simplicity in pre-theoretical terms that do not assume a representational commitment of any sort. In this case, a natural choice for a complexity metric is the number of different semantic categories in a given communicative system: the number of distinct words or morphemes it contains within a given domain. This approach is entirely domain-general precisely because it makes no assumptions about the character of the domain being modeled. We use this approach in our color and binary features case studies below.

Case study 1: Color, a continuous domain

The study of color terms across languages has long been central to the debate over the relation of language and thought. It has been known for decades that there are universal tendencies in color naming across languages (e.g. Berlin & Kay, 1969; Kay & Regier, 2003). Recently, however, an opposed view has also gained prominence, arguing that color naming varies far more across languages than had been suspected (e.g. Lucy, 1997; Roberson, Davies, & Davidoff, 2000)—and that even languages with similar color naming systems differ in the placement of boundaries between categories (Roberson, Davidoff, Davies, & Shapiro, 2005). Overall, we have an empirically mixed picture, with both universal tendencies in color naming and some deviation from those tendencies (Regier, Kay, & Khetarpal, 2007)—mirroring the general "cluster and outlier" pattern found when considering other aspects of language in cross-language perspective (Evans & Levinson, 2009: 445).

In the case of color, this empirically complex picture may be explainable in theoretically straightforward terms originally advanced by Jameson and D'Andrade (1997):

> One possible explanation [for patterns of color naming] is ... the irregular shape of the color space... Hue interacts with saturation and lightness to produce several large "bumps" ... We assume that the names that get assigned to the color space ... are likely to be those names which are most informative about color (p. 312).

The bumps to which Jameson and D'Andrade (1997) referred can be seen on the color solid in Figure 1(a) above. For example, the yellowish area protrudes more from the solid than does the blue region opposite it, indicating that colors in the yellow region can be more saturated (colorful, or un-gray) than colors in the blue region. Jameson and D'Andrade's (1997) general point is that this perceptual irregularity of the color domain has potential consequences for naming, since the irregularity or bumpiness helps to determine what will be an informative partition of the domain. Given that observation, a natural possibility is that both universal tendencies and the observed deviation from them may correspond to partitions of color space that are highly informative, with different color naming systems representing different means to this same end. Moreover, informativeness can be expected to trade off with simplicity: systems with many color terms will allow relatively fine-grained and thus highly informative partitions of color space, whereas systems with only a few color terms will allow only fairly coarse-grained partitions. We explore the hypothesis that attested color naming systems achieve a near-optimal tradeoff between informativeness and complexity—that is, that they are nearly as informative as is theoretically possible for their level of complexity (i.e. for their number of color terms).

Regier, Kay, & Khetarpal (2007) formalized this proposal and tested it against the color naming data of the World Color Survey, or WCS; this dataset is described in detail by Kay, Berlin, Maffi, Merrifield, & Cook (2009). The WCS contains color naming data from 110 languages of non-industrialized societies, collected relative to a standard color naming stimulus grid, approximated in the upper panel of Figure 3. The colors in this grid represent a Mercator projection of the outermost surface of the color solid shown in Figure 1(a).[2] Each speaker of each language in the WCS was shown each color chip in this grid one at a time, in a fixed random order, and asked to provide their language's name for that color. In the analyses reported here and by Regier et al. (2007), these data were treated as follows. For each language, each chip in the stimulus grid was labeled by the modal color term for that chip—that is, the color term that was used to label that chip by the greatest number of speakers of that language. We refer to the resulting labeling of each chip in the grid as that language's *mode map*. An example mode map is shown in the lower panel of Figure 3.
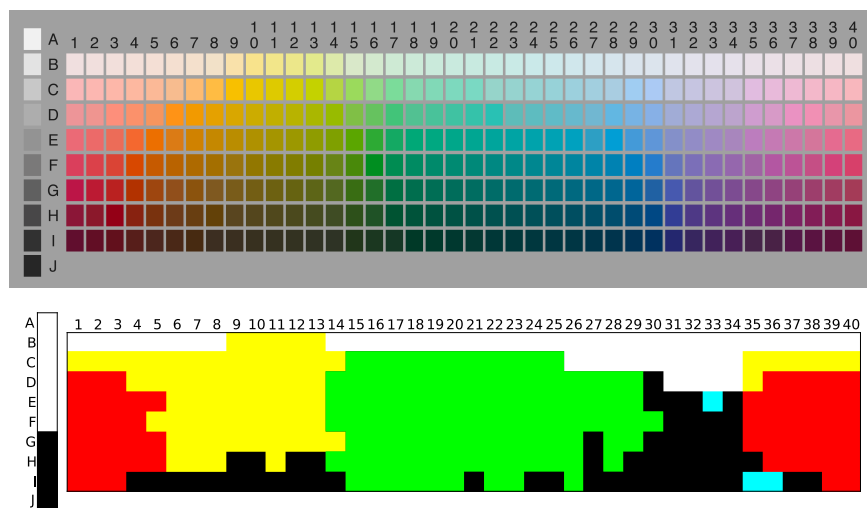

Figure 3: (Upper panel) Color naming stimulus grid. (Lower panel) Mode map for the Iduna language (Austronesian, Papua New Guinea), mapped against the stimulus grid. Each colored blob indicates the extension of a color term in this language. There are 5 major color terms, corresponding roughly to black, white, red, yellow/orange, and green/blue. There are also a small number of chips for which the modal response was another category (shown in light blue).

At the core of Regier et al.'s (2007) formalization was the intuition that categories should be constructed so as to maximize similarity within categories and minimize similarity across categories (e.g. Garner, 1974). They found that this idea accounted well for the observed variation in color naming found in the WCS. However, their formalization was not cast in terms that map directly onto the communicative scenario sketched above. Here, we cast the same ideas in terms that do map directly onto that communicative scenario. This work was undertaken as part of an ongoing collaboration with Joshua Abbott at UC Berkeley.

Following Regier et al. (2007), we treat the continuous color domain as discretized into the 330 color chips

---

[2]The Munsell color chips of maximum chroma (saturation) represented in this research are approximately as saturated as can be achieved using durable, non-fluorescent pigments. Some higher chroma values can be produced with other pigments, especially in the blue-green and green regions; in the focal green region perhaps an additional 2-4 Munsell chroma steps can be achieved. Some plants and flowers also display chroma values that somewhat exceed the Munsell atlas maxima, and television and computer screens can produce even higher apparent chroma values, in some cases approaching the theoretical maxima (Luther, 1927; MacAdam, 1935). These are, however, luminous rather than reflective colors and—significantly for our purposes—historically recent and mostly absent from the cultures in which the unwritten languages of the WCS were spoken in the 1970s. We are indebted to Rolf Kuehni for much of the content of this footnote. He is not responsible for any errors or misstatements it may contain.

of the stimulus grid shown in the upper panel of Figure 3. We assume the speaker is certain about which color chip t is to be conveyed: thus in the speaker distribution, $s(t) = 1$ and $s(i) = 0$ for all other color chips $i \neq t$. As in the scenario above, the speaker produces a color term w conveying t; this is the name for t under the color naming system shared by the speaker and the listener. The listener then reconstructs the speaker's intended meaning by constructing a probability distribution l based on the category named by w. Specifically, the listener's distribution l is constructed following a simple variant of the exemplar-based categorization model of Nosofsky (1986). In this model, a category is represented as a mixture of equally-sized Gaussian blobs, each centered at a known exemplar of the category, and each providing a measure of similarity to that exemplar. As a result, colors that are perceptually similar to many exemplars of the category will tend to be assigned higher probability, and will thus be taken to be better examples of the category. If we let $j \in cat(w)$ represent the set of colors that are named by w, then:

$$l(i) \propto \sum_{j \in cat(w)} sim(i, j) \tag{5}$$

We take the perceptual similarity of two colors x and y to be a Gaussian function of the perceptual distance between them (Nosofsky, 1986; Regier et al., 2007):

$$sim(x, y) = \exp(-c \times dist(x, y)^2) \tag{6}$$

where c = .001 for all analyses reported here and in Regier et al. (2007), and $dist(x, y)$ is the distance between colors x and y in the CIELAB color space (Brainard, 2003). In this model, the irregular shape of the color solid is captured by these similarity values.

Given these definitions of the speaker and listener distributions, the expected reconstruction error of a color naming system as a whole is provided by the general formulation of our proposal in Equations 1 through 4 above. For now, we assume that the need probability $n(t)$ in Equation 4 is uniform across color chips t; in future work we intend to investigate other need probability distributions.

We tested this formalization against the color naming data of the WCS as follows. First, for each language in the WCS, we *regularized* that language's mode map to remove minor color terms that could be considered empirical noise, such as the color term shown in light blue in the Iduna mode map in Figure 3. We considered a color term in a mode map to be a minor color term if it labeled fewer than 10 chips; otherwise we considered it a major color term. For each chip in the original mode map that was labeled by a minor color term, we relabeled that chip with the label of the most perceptually similar (following Equation 6) color chip that was labeled by a major color term. This process relabeled, on average, 1.6% of the chips across the WCS languages' mode maps. The analyses we describe below are based on the resulting regularized WCS mode maps.

We found that the languages of the WCS all had between 3 and 11 major color terms in their regularized mode maps. We wished to know whether these attested color naming systems were near-optimally informative given their level of complexity. To that end, we created a large set of hypothetical color naming systems and asked whether the attested systems were near-optimally informative relative to the hypothetical ones. Specifically, for $n = 3, 4, ... 11$, we began with a random assignment of n categories to the color chips of the naming grid, and then searched for the optimal system with that number of categories through steepest descent in expected reconstruction error E—that is, by repeatedly changing category labels on chips so as to reduce E as much as possible at each step. This optimization process was repeated 20 times from different initial random configurations of the n categories, and we recorded all values of E encountered during this

search. We considered the final configuration with lowest E over the 20 runs to be the theoretically optimal system with that number of categories.

Following Regier et al. (2007), Figure 4 shows theoretically optimal color naming systems with $n = 3,4,5$, and 6 categories, found through the optimization process just described, compared with color naming systems of selected languages in the WCS.[3]
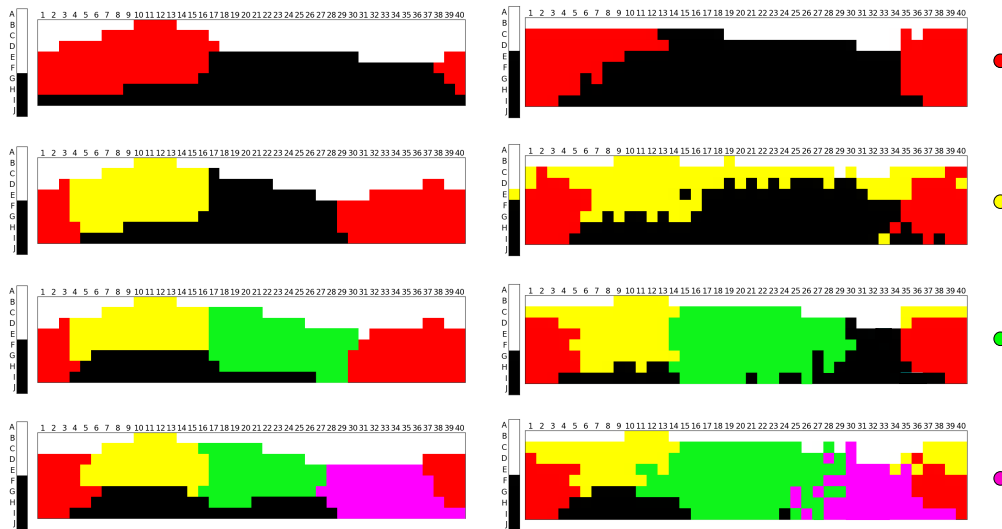

Figure 4: Theoretical optima (left) for n=3,4,5,6 categories, compared with color naming systems (right, top to bottom) of Ejagam (Bantoid, Nigeria/Cameroon), Culina (Arawan, Peru/Brazil), Iduna (Austronesian, Papua New Guinea), and Buglere (Chibchan, Panama). The colored dot next to each empirical system indexes that system in the plot of Figure 5.

As can be seen, there are languages in the WCS with color naming systems that are similar to these theoretically optimally informative systems—in fact, there are many more such languages than are shown here. These results provide an explanation for established generalizations about color naming across languages. The progression of theoretical optima shown above, and the languages that match them, correspond to stages of the implicational hierarchy proposed by Berlin and Kay (1969). That hierarchy holds that if a language has three terms, those terms will tend to correspond roughly to black, white and red, as is the case here; if there is a fourth term, it will tend to be yellow, and so on. Thus, our theory explains these known generalizations about color naming in terms of general communicative principles that may also apply to other domains.

These results support the claim that color naming across languages reflects a drive for efficient, informative communication. However, these results provide only a partial picture of our results, against a selected subset of the empirical data. A more comprehensive picture is provided by Figure 5.

In Figure 5, for each level of complexity (i.e. for each number of color terms), the solid gray vertical bar represents the full range of costs (expected reconstruction error E) found for systems of that complexity, across all optimization runs (described above) for that complexity. In general, optimization will cause a system to move downward along this bar. The top of the gray bar thus represents the highest cost ever encountered across all optimization runs for that complexity—corresponding to one of the initial random

---

[3]Comparison of this regularized mode map for Iduna with the non-regularized version shown in Figure 3 reveals the effects of the regularization process we use. The mode maps of the other three languages in this figure were unchanged by the regularization process.
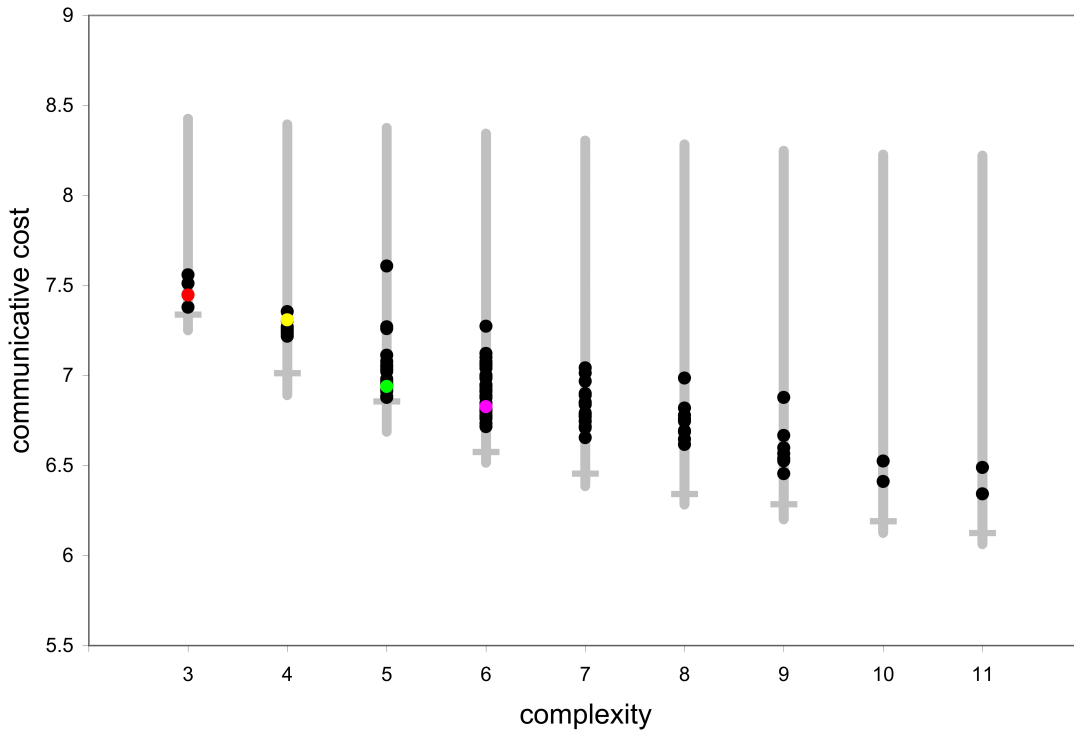
Figure 5: Communicative cost (expected reconstruction error E) vs. complexity (number of color terms) for hypothetical (range shown by gray vertical bars) and all WCS-attested (black and colored dots) color naming systems. Hypothetical systems were those encountered during optimization. Crossbars show the highest local minimum encountered at each level of complexity. Colored dots show specific WCS languages referenced above in Figure 4.

assignments of chips to categories. The bottom of the gray bar represents the lowest cost end-result of optimization across all runs for that complexity—which we take to be the theoretically optimally informative system with that number of categories. The black and colored dots represent the 110 actual color naming systems of languages in the WCS; the colored dots correspond to those languages shown in Figure 4, with color naming systems that are similar to theoretically optimal systems.

We also recorded the worst (highest cost) local optimum found over all optimization runs for a given level of complexity, to give a sense for the range of final E values that might be reached through such an optimization process. The small gray crossbars in Figure 5 show these local optima.

These results show that the color naming systems attested in the WCS tend to be fairly informative for a given level of complexity. While there exist hypothetical systems that are more informative (lower along the corresponding vertical gray bar), attested systems show expected reconstruction error near the lower end of the theoretically possible range. In some cases, attested systems show reconstruction error that lies near a local minimum, suggesting that these systems may result from an evolutionary process that has itself encountered a local optimum. A related possibility is that other systems may constitute points on an evolutionary trajectory from one optimum to another (Regier et al., 2007). Finally, the results above also show, as expected, that systems of greater complexity can in principle—and do in actuality—support more informative (less costly) communication.

These results, taken together with those of Figure 4, show that a number of languages in the WCS have highly informative color naming systems, and that in some cases these are similar to the theoretically optimal

10

system for that number of color terms. At the same time, there are a substantial number of other languages in the WCS that do not match these optimal templates especially closely—these may be thought of as the outliers in Evans and Levinson's (2009: 445) "cluster and outlier" formulation. For such languages, a natural question is whether their color naming systems are nonetheless more informative than a natural set of comparable hypothetical alternatives. To test this, following Regier et al. (2007), we created a set of hypothetical variants of each language in the WCS, by rotating that language's color naming system in the hue dimension by 0,1,2,3,... hue columns, as shown in the left panel of Figure 6. We then determined which of these variants of that language's color naming system was most informative (exhibited lowest E). The right panel of Figure 6 shows that for most WCS languages, the most informative variant of that language is the attested system itself (unrotated, i.e. 0 columns rotation). We conclude that while not all color naming systems in the WCS closely match optimal templates, they strongly tend to be more informative than simple variants of those systems.
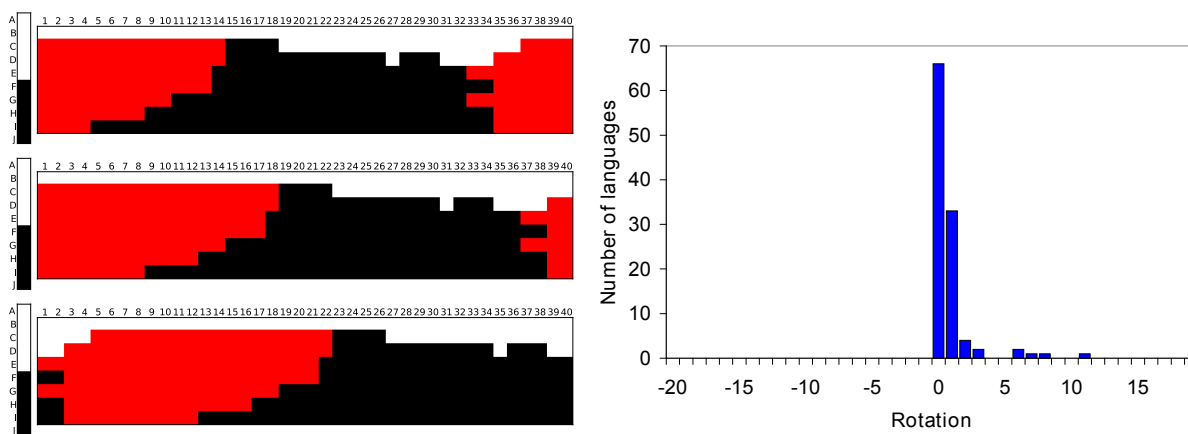

Figure 6: Left: The color naming system of Wobé (Niger-Congo, Côte d'Ivoire), shown unrotated (top) and rotated 4 (middle) and 8 (bottom) hue columns. Right: For each amount of rotation, the number of WCS languages exhibiting maximum informativeness (minimum cost) at that rotation. For most languages, the unrotated variant (0 columns rotation) is most informative.

In general, we find that languages in the WCS tend to have color naming systems with expected reconstruction error in the lower range of possible values for that number of categories—that is, these systems tend to be highly informative, for a given level of complexity. We have also seen that this fact explains known generalizations concerning color naming across languages. Thus, color naming in the world's languages does appear to reflect the general principles of simplicity and informativeness, principles that may also extend to other domains.

Case study 2: Kinship, a discrete and hierarchically structured domain

Like color, kinship is a semantic domain that has fascinated scholars for decades (e.g. Murdock, 1949; Greenberg, 1966; Nerlove & Romney, 1967; Greenberg, 1990; Jones, 2010a). It is also a domain that, like color, exhibits wide but constrained cross-language variation. This variation is exemplified in Figure 7, which compares the kin naming system of English with that of Northern Paiute (Kroeber, 1917), an indigenous language of northeastern California and adjoining areas in Nevada and Oregon, mapped against the same subset of a family tree.
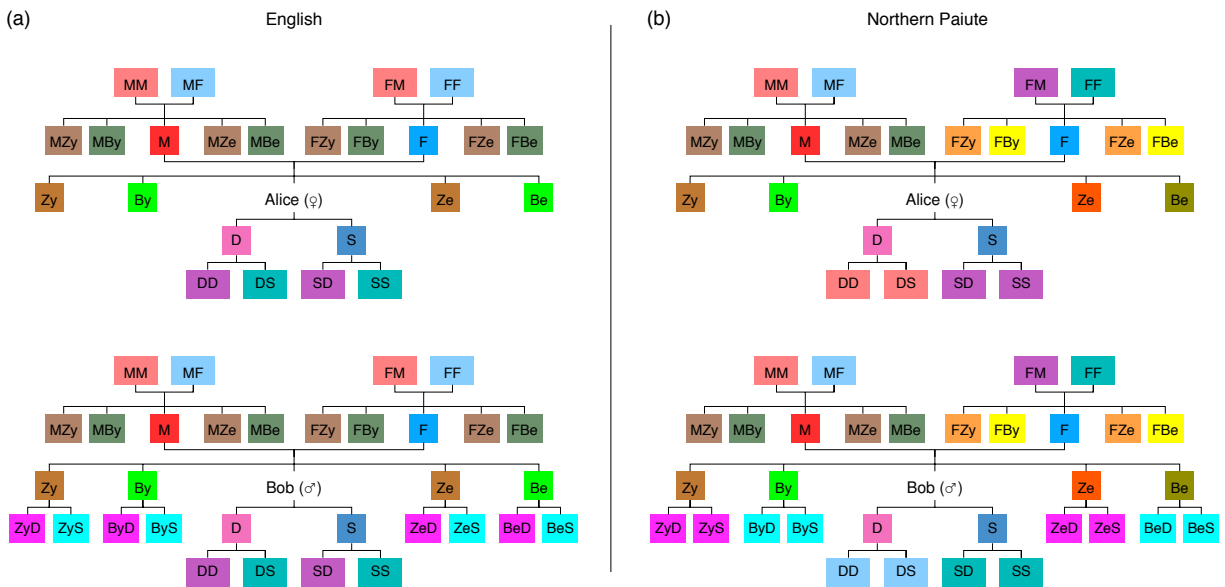
Figure 7: The kin naming systems of (a) English and (b) Northern Paiute. Color codes denote the extensions of kin terms in these languages. Adapted from Kemp & Regier (2012).

In this figure, kin types are shown relative to a female ego named Alice, and relative to a male ego named Bob. The names for these kin types are color-coded, as was the case for color. As can be seen, in Northern Paiute, unlike English, women and men use different kin terms for their grandchildren. Specifically, in Northern Paiute, Alice refers to her daughter's children (DD = daughter's daughter and DS = daughter's son) using the same kin term she uses for her maternal grandmother (MM = mother's mother). This pattern can be motivated by the observation that Alice is the maternal grandmother of DD and DS, so they refer to her using the word for maternal grandmother—and she simply uses the same word in referring to them. Similarly, Alice refers to her son's children (SD and SS) using the same kin term she uses for her paternal grandmother (FM), with analogous motivation. While this is a comparison of only two languages, it is intended to convey a preliminary sense for the cross-language variation that is observed in kin naming, as in color naming.

At the same time, the domains of color and kinship differ in many important respects. Colors are embedded within a continuous space, whereas kin types are elements of a discrete combinatorial system arranged in a structured hierarchy. Moreover colors are intrinsically perceptual, whereas the meanings of kin terms are relational and conceptual. Despite these differences, Kemp and Regier (2012) showed that the kind of analysis just carried out for color terms can also be conducted for kin terms.

Kemp and Regier (2012) assumed that: (1) kin naming systems are mentally encoded in a symbolic representation language, (2) this representation language is universal, and (3) kin naming systems for different languages can be created by combining elements of this representation language in different ways. Figure 8 shows the shortest possible description of the English kin naming system in this representation language.

Kemp and Regier (2012) took the complexity of a kin naming system to be the minimum number of rules[4] needed to define it; thus the complexity of the English kin naming system is 15. They computed the informational cost of communicating about kin in terms of the communicative scenario sketched earlier, but now instantiated in the domain of kinship, as shown in Figure 9.

---

[4]Each rule is based on one of a small set of rule templates, and each template supports a single simple operation such as conjunction, disjunction, relative product, and the like. As a result, rules are not arbitrarily complex, and the number of rules needed to describe a system is often greater than the number of kin terms in that system. Details are provided by Kemp and Regier (2012).

🟥	$mother(x, y)$	$\leftrightarrow PARENT(x, y) \wedge FEMALE(x)$
🟦	$father(x, y)$	$\leftrightarrow PARENT(x, y) \wedge MALE(x)$
🩷	$daughter(x, y)$	$\leftrightarrow CHILD(x, y) \wedge FEMALE(x)$
🟦	$son(x, y)$	$\leftrightarrow CHILD(x, y) \wedge MALE(x)$
🟫	$sister(x, y)$	$\leftrightarrow \exists z\, daughter(x, z) \wedge PARENT(z, y)$
🟩	$brother(x, y)$	$\leftrightarrow \exists z\, son(x, z) \wedge PARENT(z, y)$
	$sibling(x, y)$	$\leftrightarrow \exists z\, CHILD(x, z) \wedge PARENT(z, y)$
🟫	$aunt(x, y)$	$\leftrightarrow \exists z\, sister(x, z) \wedge PARENT(z, y)$
🟩	$uncle(x, y)$	$\leftrightarrow \exists z\, brother(x, z) \wedge PARENT(z, y)$
🟪	$niece(x, y)$	$\leftrightarrow \exists z\, daughter(x, z) \wedge sibling(z, y)$
🟦	$nephew(x, y)$	$\leftrightarrow \exists z\, son(x, z) \wedge sibling(z, y)$
🟥	$grandmother(x, y)$	$\leftrightarrow \exists z\, mother(x, z) \wedge PARENT(z, y)$
🟦	$grandfather(x, y)$	$\leftrightarrow \exists z\, father(x, z) \wedge PARENT(z, y)$
🟪	$granddaughter(x, y)$	$\leftrightarrow \exists z\, daughter(x, z) \wedge CHILD(z, y)$
🟩	$grandson(x, y)$	$\leftrightarrow \exists z\, son(x, z) \wedge CHILD(z, y)$

Figure 8: The shortest description of the English kin naming system in the representation language of Kemp & Regier (2012). Adapted from Kemp & Regier (2012).

Here, the speaker is certain that the intended kin type t is her older brother, and for this reason the speaker distribution has probability mass 1.0 allocated to older brother, and 0.0 for all other kin types. The speaker uses the kin term w ("brother") to convey this meaning, and the listener must mentally reconstruct what kin type was intended. Kemp and Regier (2012) assumed that the listener's distribution is constructed by allocating probability mass only to kin types that are in the extension of the category named by w (here, younger brother and older brother), in proportion to their need probability. Thus, if we let $cat(w)$ represent the set of kin types that are in the category named by w, then:

$$l(i) = \frac{n(i)}{\sum_{j \in cat(w)} n(j)} \tag{7}$$

if $i \in cat(w)$, and 0 otherwise. The need probability $n(i)$ for referring to a given kin type i was estimated
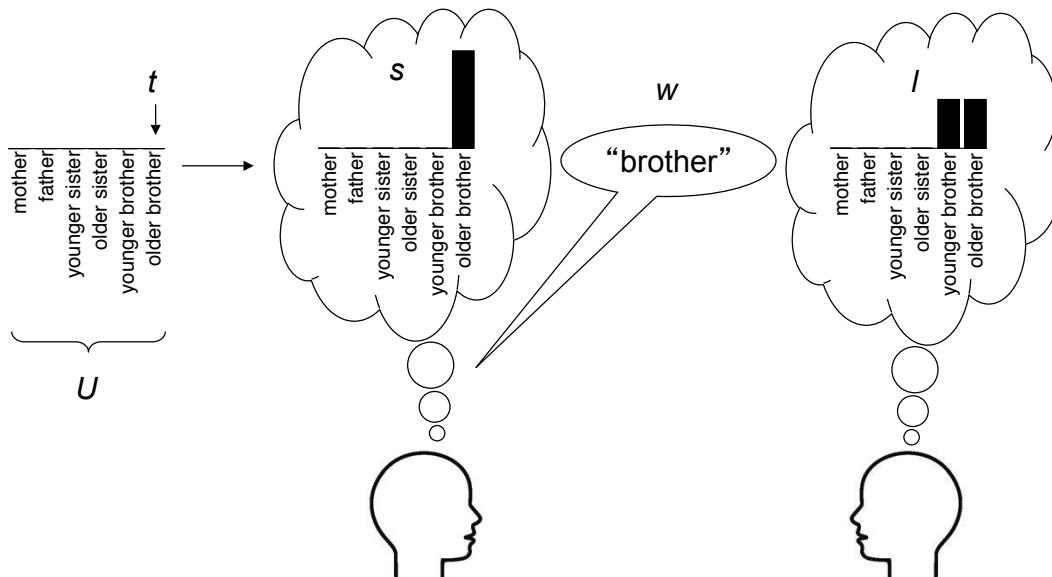

Figure 9: A scenario illustrating communication about a kin type.

through corpus counts for kin terms in English and German.[5] Given these domain-specific definitions, the communicative cost of a kin naming system as a whole is provided by the general formulation of our proposal in Equations 1 through 4 above.

Kemp and Regier (2012) used these kinship-specific implementations of complexity and cost to analyze a cross-language kin classification dataset compiled by Murdock (1970). This dataset contains kin classification systems for over 500 languages and constitutes an attempt to document the full range of attested variation in kinship systems worldwide. Some of these systems are incompletely specified in Murdock's dataset, so Kemp and Regier (2012) focused on the 487 languages for which full specifications were provided: specifications that provided a kin name for each of the 56 kin types (24 relative to female ego and 32 relative to male ego) shown in Figure 7 above.

By analogy with the color analyses above, Kemp and Regier (2012) created a large set of hypothetical kinship systems, and compared these hypothetical systems to attested systems with respect to both complexity and cost. As before, the prediction was that most attested systems will be nearly as informative (uncostly) as is theoretically possible for their level of complexity. The results of this analysis are shown in Figure 10.
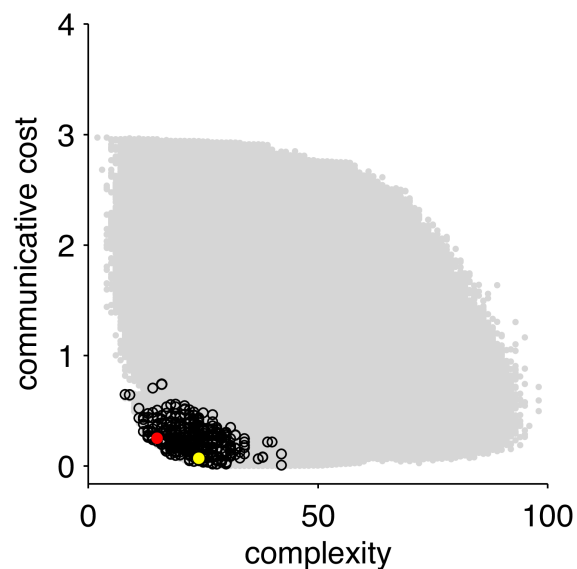

Figure 10: Communicative cost (expected reconstruction error) vs. complexity for hypothetical (gray circles) and attested (black and colored circles) kin naming systems. Colored circles show the systems of English (red) and Northern Paiute (yellow). Adapted from Kemp & Regier (2012).

In this figure, as in the figure of the corresponding color results, gray circles represent hypothetical systems and black circles represent attested ones. As predicted, the attested kinship systems in the Murdock dataset tend to exhibit fairly low communicative cost for their level of complexity—in line with the results found for color under the same general formulation. The red circle represents the system of English, and the yellow circle represents that of Northern Paiute. It can be seen that despite their differences, each of these two systems represents a near-optimal tradeoff of complexity and cost.

At the same time, as in the case of color, many other systems exhibit communicative cost that, while relatively

---

[5]Different cultures may impose different need probability distributions, a possibility we discuss later. In the analyses reported here, Kemp and Regier (2012) provisionally assumed a universal distribution, and approximated it by English and German counts for reasons of convenience: large corpora for these languages were readily available. Kin term counts were converted to a distribution over kin types by evenly spreading the count for a given kin term over those kin types named by that kin term: for example, the count for the English kin term "aunt" was evenly distributed over the kin types MZy, MZe, FZy, FZe.

low, is still clearly higher than the theoretical optimum for that level of complexity. Such systems might have been expected to be more informative for that level of complexity, or simpler (less complex) for that level of informativeness. A natural question—again as in the case of color—is whether such systems are nonetheless superior to hypothetical variants of these systems. In the case of color, we used rotation in hue as a (domain-specific) means of systematically generating hypothetical variants of an existing system, in order to compare the original system to those variants. In the domain of kinship, rotation does not seem a natural way to systematically generate variants of an existing system, so Kemp & Regier (2012) pursued a different (again domain-specific) means to that same end. They first divided the family tree into "chunks", each containing two male and two female kin types, as shown in Figure 11(a). They then considered permutations of the existing system in which category labels were exchanged across chunks—for example, exchanging the kin terms used to label kin types in the grandparent chunk (MM, MF, FM, FF) with those used to label kin types in the grandchildren chunk (DD, DS, SD, SS). They considered all such permutations that respected category boundaries: that is, permutations that moved entire categories and not just parts of categories.
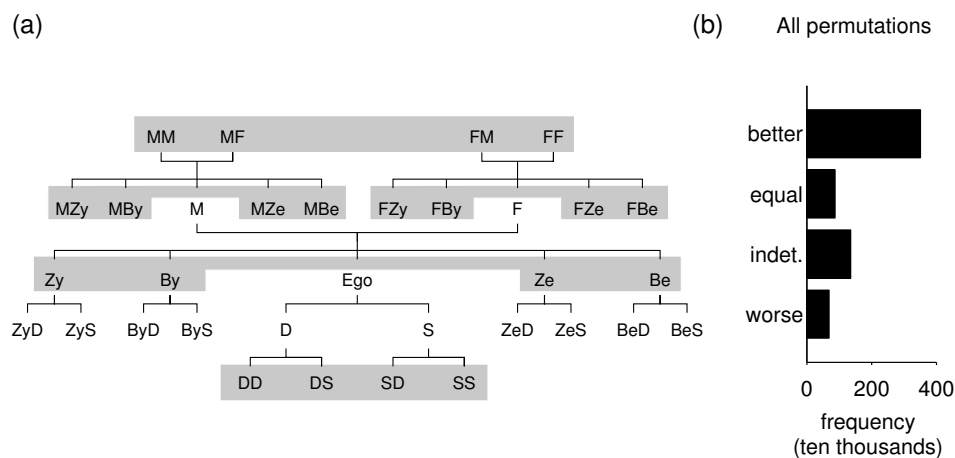

Figure 11: Permutation analysis for kinship. (a) The gray regions indicate chunks of the family tree used for the permutation analysis. (b) Attested systems typically score better than permuted versions of these systems and rarely score worse. Adapted from Kemp & Regier (2012).

The color rotation analyses preserved the complexity of a system while potentially varying its informativeness. The kinship permutation analyses, in contrast, could potentially vary both the informativeness and the simplicity of a given kin naming system: the shortest description of the permuted system need not have the same length as that of the original system. For this reason, Kemp and Regier (2012) compared the original and permuted systems with respect to both communication cost and complexity, and considered 4 possible outcomes of each permutation: (1) the attested system might score *better* than the permuted system along one dimension and no worse along the other, (2) the attested system might score *the same as* the permuted system along both dimensions, (3) the attested system might score *worse* than the permuted system along one dimension and no better along the other, and (4) the attested system might score better than the permuted system along one dimension and worse along the other, in which case the comparison is *indeterminate*. Figure 11(b) shows that when the full set of permutations is applied to the kin naming systems of the Murdock dataset, attested systems tend to score better than permutations of those systems, and rarely score worse. This finding suggests that kin naming systems in the Murdock dataset tend to trade off communicative cost and complexity more efficiently than do comparable hypothetical systems derived from them.

Finally, Kemp and Regier (2012) demonstrated that these ideas of efficient communication account for a number of descriptive generalizations about kin classification that have been proposed—by analogy with the

demonstration we saw earlier in the case of color. An example of such a generalization about kinship is the markedness constraint that near relatives (e.g. siblings) are more likely than distant relatives (e.g. parent's siblings) to be split into multiple categories (Greenberg, 1966; 1990). Kemp and Regier (2012) showed that on the present theory, this constraint follows from the distribution of need probabilities across the family tree. Another example is the observation (e.g. Nerlove & Romney, 1967: 181; Greenberg, 1990: 320) that kinship categories are more likely to have conjunctive definitions (e.g. parent AND female) than disjunctive definitions (e.g. parent OR female). On the present theory, this constraint follows from the observation that conjunctive definitions are more informative: they narrow the scope of possible reference, whereas disjunctive definitions broaden it, relative to the elements (e.g. parent, female) being combined.

We have seen that kin naming systems across languages, like color naming systems, tend to achieve a near-optimal tradeoff between informativeness and simplicity—and that this fact accounts for some known descriptive generalizations concerning color and kin naming across cultures. Thus, the account we have proposed here provides a unified explanation for cross-language variation in two semantic domains that are representationally quite different. We turn next to examine the possible extension of these ideas to a domain defined in yet another way.

Case study 3: Binary feature vectors

Although the domains of color and kinship are different in many ways, the two share some similarities. In particular, the objects within both domains vary with respect to a relatively small set of dimensions. Colors vary along the dimensions of hue, saturation, and lightness, and kin-types can be characterized by combining a relatively small set of genealogical relationships. Other semantic domains, including animals and plants, can not be characterized as simply. For this reason, Jones (2010b: 210) suggests that his optimality-theoretic account of color and kinship categories is unlikely to extend to the domain of folk biology.

In contrast, we believe that our information-theoretic approach may help to explain how animals and plants are organized into categories. Several researchers have used feature-based representations to capture knowledge about animals (Boster & D'Andrade, 1989) and plants (Boster, 1984), and this section describes how our approach could be applied to domains in which objects are represented using binary feature vectors. As a concrete example, we will analyze a toy dataset that includes six fruits that are defined in terms of 25 features (Rosch et al., 1976; Tversky and Hemenway, 1984; Corter and Gluck, 1992). This analysis is only a proof of concept, but we hope that it will suggest how our theory could be tested by collecting and analyzing large feature-based data sets.

Figure 12 shows a communicative scenario about fruit represented as feature vectors—here simplified for exposition with only 2 categories defined over only 3 features. In this figure, the universe U includes all 8 possible binary feature vectors of length 3. Just as the Rosch et al. data set includes labels for only 6 of the 2^{25} possible feature vectors of length 25, here we assume that only 5 of the $2^3 = 8$ feature vectors have been labeled, leaving several feature vectors unlabeled. Each feature vector represents a fruit, and the labeled fruits are 2 peaches and 3 apples.

We formalize this scenario as we did the analogous scenarios for color and kinship, to assess the communicative cost and complexity of different ways of assigning labels to objects. As for our previous analyses, we assume that the speaker is certain about the identity of the target object. We assume that the listener distribution l is defined as:
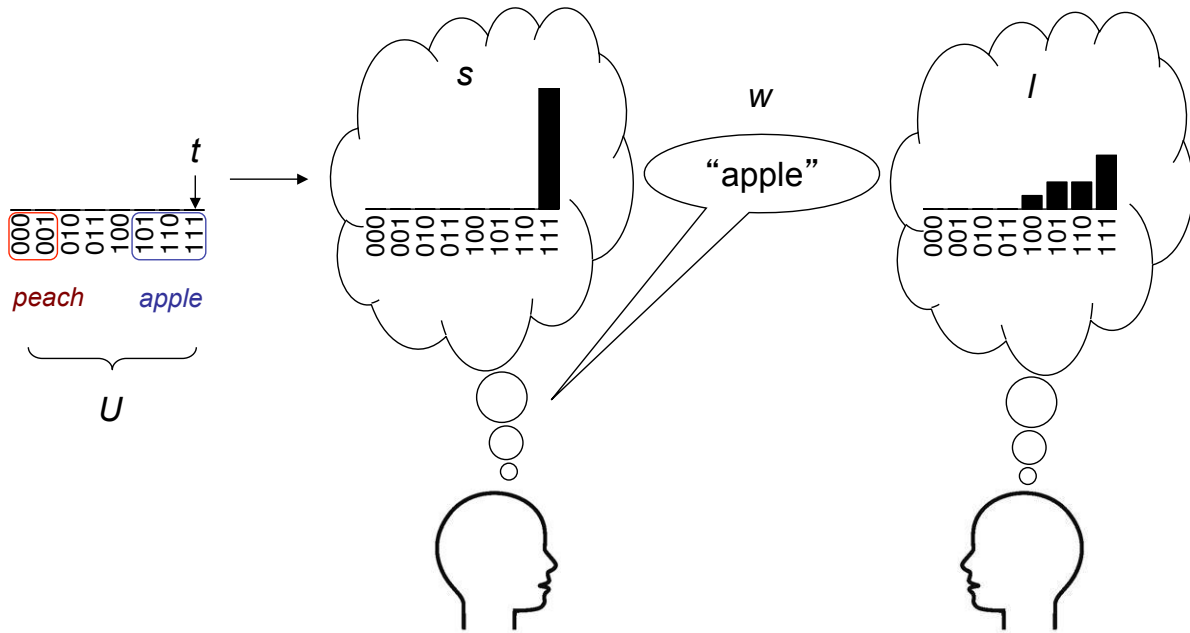
Figure 12: A scenario illustrating communication about an object represented as a vector of three binary features. Of the 8 possible feature vectors, 2 have been assigned to the category "peach" and 3 have been assigned to the category "apple".

$$l(i) = p(i|w) = p(f_1...f_n|w) = p(f_1|w)...p(f_n|w) \qquad (8)$$

Equation 8 assumes that the n features f_1 through f_n are conditionally independent given the category label w. The same independence assumption is made by previous models of categorization, including the model of Anderson (1990). The distribution for each feature $p(f_i|w)$ records the relative frequency of feature f_i across those objects that are labeled w. For example, in Figure 12, 2 of the 3 labeled apples exhibit feature 2; therefore $p(f_2 = 1|\text{apple}) = 2/3$. Those objects whose feature vectors share features with many exemplars of the category will be assigned higher probability, and thus be taken as better examples of the category—by analogy with the listener distribution we used earlier for color. For the same reason, this listener distribution also assigns non-zero probability to an object (represented by feature vector 100) that has never previously been labeled, but that shares features with others that have been labeled "apple".

We assume that the need probability distribution $n(i)$ is uniform over the feature vectors i in the dataset, and that $n(i)$ for any other feature vector is zero. As with our color analysis, we take the complexity of a system to be the number of categories that it contains. We analyzed the Rosch et al. fruit dataset in terms of our general formal proposal, instantiated for binary feature vectors as just described.

There are 203 ways to partition the 6 fruits into a system of categories, and Figure 13(a) shows the three systems of categories that are commonly used in English. The middle system organizes the objects into 3 basic level categories (apples, peaches, and grapes), and the other two systems organize the objects into superordinate (fruit) and subordinate categories. Figure 13(b) plots the informativeness of each possible system against its complexity. If the English superordinate, basic level and subordinate categories are "natural" categories in the same sense as color terms and kinship names across languages are, they should optimize the tradeoff between informativeness and complexity just as the color and kinship names do—that is, they should show the minimum reconstruction error possible for their level of complexity. The three black circles in Figure 13(b) correspond to these three systems, and confirm that this is indeed the case. Our analysis
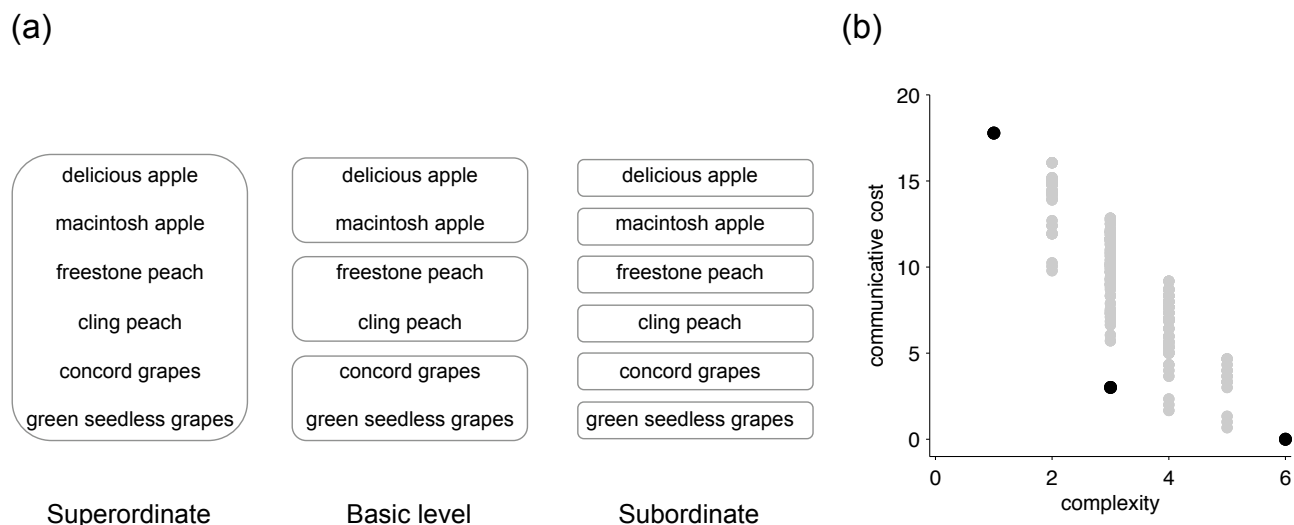
Figure 13: (a) The 6 fruits in the Rosch et al. data set, here organized into superordinate (fruit), basic level (apple, peach, grapes), and subordinate categories. (b) Reconstruction error versus complexity for all possible systems that organize the 6 fruits into categories. Black dots show the superordinate, basic level, and subordinate categories of (a).

does not predict where along the optimal frontier the basic (or psychologically-privileged) categories should lie, but Corter and Gluck (1992) showed how an information-theoretic approach similar to ours can pick out one level of categorization as psychologically-privileged.

The toy example in this section did not consider cross-linguistic data, but the approach described here could potentially be used to analyze such data in a domain that is represented in terms of binary feature vectors. Extending our work in that direction is a priority for future research.

Conclusions

In this chapter, we have proposed that systems of semantic categories in the world's languages reflect the need for efficient communication, in that they near-optimally balance the competing principles of simplicity and informativeness. Our proposal echoes ideas from Gabelentz (1901), to Rosch (1999), to recent studies of efficient communication in other aspects of language (e.g. Piantadosi, Tily, & Gibson, 2011; Fedzechkina, Jaeger, & Newport, 2012). Our contribution has been to formalize this proposal in terms of a general computational framework, and to test it against a broad range of empirical data from hundreds of languages, across qualitatively different semantic domains. Our analyses of color and kinship have shown that our framework accounts for cross-language data in both of these domains, and also provides a more general explanation for known patterns in the data that had previously been explained in color-specific or kinship-specific terms. Our analysis of a domain defined in terms of binary feature vectors has provided further evidence for the generality of our framework.

These results open up a number of questions for future research. Perhaps the most salient is whether the same ideas also apply to semantic domains other than those we have investigated here. Linguists, anthropologists, and psychologists have collected cross-linguistic data for many other semantic domains, including spatial relations (e.g. Levinson, Meira, & the Language and Cognition Group, 2003), plants and animals (e.g. Berlin, 1992; Atran, 1995), emotions (e.g. Boster, 2005), artifacts (e.g. Malt, Sloman, Gennari,

Shi, & Wang, 1999), and body parts (e.g. Majid, Enfield, & van Staden, 2006), among others. Khetarpal et al. (2009; 2013) have presented analyses of spatial terms that closely parallel the analysis of color given in this chapter, but it remains to be seen whether our theory can be productively applied to the many other semantic domains that have been discussed in the literature.

Another open question concerns historical language change. Our theory holds that semantic systems will tend to support efficient communication, but it makes no claims concerning the historical processes that produce this state of affairs. Still, our analysis of color hinted that systems of categories may be understood as local optima with respect to a historical optimization process. Other researchers have developed formal accounts of linguistic and cultural evolution (e.g. Kirby, 2002; Jäger, 2007; Kalish, Griffiths, & Lewandowsky, 2007), and such evolutionary accounts have helped to explain why color categories (Steels & Belpaeme, 2005; Dowman 2007; Komarova, Jameson, & Narens, 2007; Xu, Dowman, & Griffiths, 2013) and kinship categories (Epling, Kirk, & Boyd, 1973; Jordan, 2011) take the forms that they do. A natural direction for future work is to integrate our account with such an evolutionary perspective.

Finally, an important open question concerns possible determinants of cross-language and cross-culture variation. We have emphasized that there are many "good" (communicatively efficient) ways to categorize a given domain. Implicit in this picture is the notion that cross-language variation arises because different languages settle on different solutions among the set of candidates that are rated as highly efficient by our analyses. However there is another potential source of variation that we have not yet explored in detail, and that affects which candidates will be highly rated. In both our color analyses and our kinship analyses, we assumed—conveniently and we hope reasonably, but almost certainly falsely—that the distribution of need probabilities over objects in the domain was the same distribution for all languages. A more targeted test of our ideas is possible in which need probabilities are determined on a per-language basis, and the optimality of a given semantic system is assessed given those language-specific need probabilities. For example, languages spoken in environments of lush vegetation might have greater need to refer to greenish colors—or less need, if this vegetation is generally taken as background. Thus, different languages may represent efficient solutions to communicating in the context of rather different culturally or environmentally driven referential habits—a possibility we have not yet examined.

Our theory is based on general principles that are instantiated differently within different domains—and it therefore incorporates both domain-general and domain-specific elements. Determining which aspects of categorization are grounded in general vs. domain-specific principles is itself a major question that our results only begin to answer. We do not claim that all major features of categorization across languages and across domains can be attributed to a tradeoff between the domain-general principles of simplicity and informativeness. Our results do suggest, however, that these two broad domain-general principles shape categorization to a greater extent than has previously been recognized—and that many important aspects of categorization across cultures may thus reflect the functional need for efficient communication.

Acknowledgments

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Atran, S. (1995). Classifying nature across cultures. In E. E. Smith & D. N. Osherson (Eds.), *Thinking: An invitation to cognitive science, vol. 3* (pp. 131-174). Cambridge, MA: MIT Press.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech, 47*, 31-56.

Baddeley, R., & Attewell, D. (2009). The relationship between language and the environment: Information theory shows why we have only three lightness terms. *Psychological Science, 20*, 1100-1107.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173-218). New York: Cambridge University Press.

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

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

Bever, T.G. & Langendoen, D.T. (1972). The interaction of speech perception and grammatical structure in the evolution of language. In R. Stockwell & R. Macaulay (Eds.), *Linguistic change and generative theory.* Bloomington: Indiana University Press.

Boster, J. S. (1984). Classification, cultivation, and selection of Aguaruna cultivars of Manihot esculenta (Euphorbiaceae). *Advances in Economic Botany, 1*, 34-47.

Boster, J. (2005). Emotion categories across languages. In H. Cohen & C. LeFebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 187-222). Amsterdam: Elsevier.

Boster, J. & D'Andrade, R. (1989). Natural and human sources of cross-cultural agreement in ornithological classification. *American Anthropologist, 91(1)*, 132-142.

Brainard, D. H. (2003). Color appearance and color difference specification. In S. K. Shevell (Ed.), *The science of color: Second edition* (pp. 191-216). Oxford, UK: Elsevier.

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences, 7*, 19-22.

Corter, J. & Gluck, M. (1992). Explaining basic categories: Feature predictability and information. *Psychological Bulletin, 111*, 291-303.

Croft, W. (2003). *Typology and universals: Second edition.* Cambridge, UK: Cambridge University Press.

D'Andrade, R. G., Quinn, N. R., Nerlove, S. B., & Romney, A. K. (1972). Categories of disease in American-English and Mexican-Spanish. In A.K. Romney, R.N. Shepard, and S.B. Nerlove, (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences II* (pp. 9-54). New York: Seminar Press.

Dowman, M. (2007). Explaining color term typology with an evolutionary model. *Cognitive Science, 31*, 99-132.

Du Bois, J.A. (1985). Competing motivations. In Haiman, J. (Ed.), *Iconicity in syntax* (pp. 343-366). Amsterdam: Benjamins.

Epling, P., Kirk, J., & Boyd, J. (1973). Genetic relations of Polynesian sibling terminologies. *American Anthropologist, 75*, 1596-1625.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429-492.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *PNAS, 109*, 17897-17902.

Ferrer i Cancho, R. & Solé, R. (2003). Least effort and the origins of scaling in human language. *PNAS, 100*, 788-791.

Gabelentz, G. von der (1901). *Die Sprachwissenschaft, ihre Aufgaben, Methoden, und bisherige Ergebnisse.*

Leipzig: Weigel.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Greenberg, J. (1966). *Language univerals*. The Hague: Mouton de Gruyter.

Greenberg, J. (1990). Universals of kinship terminology: Their nature and the problem of their explanation. In K. Denning & S. Kemmer (Eds.), *On language: Selected writings of Joseph Greenberg* (pp. 310-327). Stanford: Stanford University Press.

Haiman, J. (2010). Competing motivations. In J. J. Song (Ed.), *The Oxford handbook of linguistic typology* (pp. 148-165). Oxford: Oxford University Press.

Haspelmath, M. (1999). Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft, 18*, 180-205.

Hawkins, J. (2004). *Efficiency and complexity in grammars*. New York: Oxford University Press.

Hopper, P., & Traugott, E. (2003). *Grammaticalization. Second edition*. Cambridge, UK: Cambridge University Press.

Jaeger, T. F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*, 323-335.

Jäger, G. (2007). Evolutionary game theory and typology: A case study. *Language, 83*, 74-109.

Jameson, K. A., & D'Andrade, R. G. (1997). It's not really red, green, yellow, blue. In C. L. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 295-319). Cambridge: Cambridge University Press.

Jones, D. (2010a). Human kinship, from conceptual structure to grammar. *Behavioral and Brain Sciences*, 33, 367-416.

Jones, D. (2010b). Grammars of kinship and color: Cognitive universals and optimal communication. In D. Jones and B. Milicic (Eds.), *Kinship, language, and prehistory: Per Hage and the renaissance in kinship studies* (pp. 196-211). Salt Lake City: University of Utah Press.

Jordan, F. (2011). A phylogenetic analysis of the evolution of Austronesian sibling terminologies. *Human Biology*, 83, 297-321.

Kalish, M., Griffiths, T., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin and Review, 14*, 288-294.

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The World Color Survey*. Stanford: Center for the Study of Language and Information.

Kay, P. & Regier, T. (2003). Resolving the question of color naming universals. *PNAS, 100*, 9085-9089.

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

Khetarpal, N., Majid, A., & Regier, T. (2009). Spatial terms reflect near-optimal spatial categories. In N. Taatgen et al. (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*.

Khetarpal, N., Neveu, G., Majid, A., Michael, L., & Regier, T. (2013). Spatial terms across languages support near-optimal communication: Evidence from Peruvian Amazonia, and computational analyses. In M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*.

Kirby, S. (2002). Learning, bottlenecks, and the evolution of recursive syntax. In Ted Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models*. Cambridge: Cambridge University Press.

Komarova, N., Jameson, K. & Narens, L. (2007). Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology, 51*, 359-382.

Kroeber, A. (1917). *California kinship systems*. Berkeley: University of California Press.

Lantz, D., & Stefflre, V. (1964). Language and cognition revisited. *Journal of Abnormal and Social Psychology, 69*, 472-481.

Levinson, S., Meira, S., & the Language and Cognition Group (2003). 'Natural concepts' in the spatial topo-

logical domain—adpositional meanings in crosslinguistic perspective: An exercise in semantic typology, *Language, 79*, 485-516.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126-1177.

Lucy, J. (1997). The linguistics of "color". In C. L. Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 320-346). Cambridge: Cambridge University Press.

Luther, R. (1927). Aus dem Gebiet der Farbreiz-Metrik. *Zeitschrift für Technische Physik, 8*, 540-558.

MacAdam, D. (1935). The theory of the maximum visual efficiency of colored materials. *Journal of the Optical Society of America, 25*, 249-249.

Majid, A., Enfield, N.J., & van Staden, M. (2006). Parts of the body: Cross-linguistic categorisation (Special Issue). *Language Sciences, 28*.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language, 40*, 230-262.

Murdock, G. (1949). *Social structure*. New York: MacMillan.

Murdock, G. (1970). Kin term patterns and their distribution, *Ethnology 9*, 165-208.

Nerlove, S., & Romney, A. (1967). Sibling terminology and cross-sex behavior. *American Anthropologist 69*, 179-187.

Nosofsky, R. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Erlbaum.

Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *PNAS, 108*, 3526-3529.

Prince, A., & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science, 275*, 1604-1610.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *PNAS, 104*, 1436-1441.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50*, 378-411.

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*, 369-398.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8(3)*, 382-439.

Rosch, E. (1999). Principles of categorization. In Margolis, E. and Laurence, S. (Eds.), *Concepts: Core readings* (pp. 189-206). Cambridge, MA: MIT Press. Reprinted from Rosch, E. & Lloyd, B.B. (Eds.) (1978), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

Steels, L. & Belpaeme, T. (2005) Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28*, 469-489.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113(2)*, 169-193.

Xu, J., Dowman, M., & Griffiths, T. (2013). Cultural transmission results in convergence towards colour term universals. *Proceedings of the Royal Society B, 280*, 20123073.

Zipf, G. (1949). *Human behavior and the principle of least effort*. Cambridge,MA: Addison-Wesley.