Using statistics for cross-linguistic semantics: a quantitative investigation of the typology of color naming systems

Gerhard Jäger
gerhard.jaeger@uni-tuebingen.de

Abstract

The paper presents a statistical evaluation of the typological data about color naming systems across the languages of the world that have been obtained by the World Color Survey. In a first step, we discuss a principal component analysis of the categorization data. This leads to a small set of easily interpretable features that are dominant in color categorization. These features were used for a dimensionality reduction of the categorization data.

Based on the thus pre-processed data, it is investigated how the participants of the World Color Survey partition the six primary colors *black, white, red, green, yellow*, and *blue* into semantic categories. We find a substantial number of counterexamples to the implicative semantic universals that have been suggested in the literature. Finally, an alternative system of semantic universals pertaining to color naming systems is proposed that provides a better fit of the data.

1 Introduction: Berlin and Kay's study

In their ground-breaking study (Berlin and Kay, 1969), Berlin and Kay investigated the cross-linguistic patterns of color naming systems in typologically diverse languages. Even though they only used a small sample of twenty languages in their initial pilot study, they were able to make a convincing case that there are strong universal tendencies, despite all cross-linguistic variation. On the basis of this pilot study and published data from 78 more languages, they came up with the following empirical generalizations (cited from Kay and Maffi 1999, p. 744):

I "There exists a *small set of perceptual landmarks* (that we can now identify with the Hering primary colors: black, white, red, yellow, green, blue) that individually or in combination *form the basis of the denotations of most of the major color terms* of most of the languages of the world.

II Languages are frequently observed to *gain basic color terms in a partially fixed order.* Languages are infrequently or never observed to lose basic color terms."
(emphasis in original)

More specifically, they proposed a hierarchy of colors that forms the basis of a series of implicative universals. This hierarchy (Berlin and Kay 1969, p. 4) is given in Figure 1.

$$\begin{bmatrix}\text{white}\\ \text{black}\end{bmatrix} < \text{red} < \begin{bmatrix}\text{green}\\ \text{yellow}\end{bmatrix} < \text{blue} < \text{brown} < \begin{bmatrix}\text{purple}\\ \text{pink}\\ \text{orange}\\ \text{gray}\end{bmatrix}$$

Figure 1: Hierarchy of colors according to Berlin and Kay (1969)

If a color a occurs to the left of color b in this hierarchy, then every language that has a word for b will also have a word for a. So languages with four color terms always have words for *black, white* and *red*, while the fourth word may denote *yellow* or *green.* Languages with five color terms will have words both for *yellow* and *green* etc.

2 The World Color Survey

Berlin and Kay's work sparked a controversial discussion. To counter the methodological criticism that has been raised in this context, Kay and several co-workers started the **World Color Survey** project (WCS, see Cook et al., 2005 for details), a systematic large-scale collection of color categorization data from a sizable amount of typologically distinct languages across the world.

To be more precise, the WCS researchers collected field research data for 110 unwritten languages, working with an average of 24 native speakers for each of these languages. During this investigation, the Munsell chips were used, a set of 330 chips of different colors that cover 322 colors of maximal saturation plus eight shades of gray. Figure 13 (see Appendix) displays them in form of the Munsell chart.

The main chart is a 8×40 grid, with eight rows for different levels of lightness, and 40 columns for different hues. Additionally there is a ten-level column of achromatic colors, ranging from white via different shades of gray to black.[1] The level of granularity is chosen such that the difference between two neighboring chips is minimally perceivable.

Research in the psychology of color perception has revealed that a more faithful spatial representation of the psychological color space must actually be three-dimensional. The standard model here is the CIELab space. It is a 3d space with the dimensions L* (for lightness), a* (the green-red axis) and b* (the yellow-blue axis). The set of perceivable colors forms a three-dimensional solid with approximately spherical shape. Figuratively

[1]The chips are referenced with their horizontal coordinate (A...J) for lightness and their vertical coordinate (1...40) for hue. The gray scale has the x-coordinate 0.

speaking, white is at the north pole, black at the south pole, the rainbow colors form the equator, and the gray axis cuts through the center of the sphere. The CIELab space has been standardized by the "Commission Internationale d'Eclairage" such that Euclidean distances between pairs of colors are monotonically related to their perceived dissimilarity.

The 320 chromatic Munsell colors cover the surface of the color solid, while the ten achromatic chips are located at the vertical axis. The color solid is actually not completely spherical but irregularly shaped. Figure 14 (see Appendix) shows the color solid from several perspectives. (The CIELab coordinates for the 330 Munsell colors that I used are available from the WCS web site.)

For the WCS, each test person was "asked (1) to name each of 330 Munsell chips, shown in a constant, random order, and (2), exposed to a palette of these chips and asked to pick out the best example(s) ('foci') of the major terms elicited in the naming task" (quoted from the WCS homepage). The data from this survey are freely available from the WCS homepage http://www.icsi.berkeley.edu/wcs/data.html.

This invaluable source of empirical data has been used in a series of subsequent evaluations that largely confirm Berlin and Kay's hypothesis that there are universal tendencies in color naming systems across languages (see for instance Kay and Maffi, 1999; Kay and Regier, 2003; Regier et al., 2005; Lindsey and Brown, 2006), even though the controversy about universality vs. relativism continues (see for instance Roberson et al., 2000).

3 Partition models

Kay and McDaniel (1978) proposed a reconceptualization of Berlin and Kay's framework. They assume that a system of basic color terms *partitions* the color space into disjoint but jointly exhaustive regions. For instance, in a two-term systems, the terms do not just denote *white* and *black*, but they split the color space into two regions, the first one comprising *white, red, yellow* and everything in between, and the second one *black, blue, green* and everything in between. Adding a third term amounts to splitting the first category into parts, one comprising *white* and neighboring colors, and the other one covering *red* and *yellow*. Kay and McDaniel maintain Berlin and Kay's assumptions that color categories are always organized around universal perceptual landmarks, and that there is a fixed hierarchy of possible color term systems.

The hierarchy that was proposed by them underwent various revisions by other authors in the light of new data, mainly from the WCS. There are several competing proposals in the literature, see for instance (Kay and Maffi, 1999; Kay et al., 1997). There is no real consensus in the literature about the correct partition hierarchy; the most fine-grained proposal is perhaps the one from Figure 2.4 of (Kay et al., 1997) (reproduced here as Table 1). The proposal only deals with the six primary colors *white, black, red, green, blue* and *yellow*. It is claimed that almost all languages partition the primary colors into one of the systems given in the table.

I	II	III	IV	V
		white red/yellow green/blue black	white red yellow green/blue black	
white/red/yellow black/green/blue	white red/yellow black/green/blue	white red/yellow green black/blue		white red yellow green blue black
		white red yellow black/green/blue	white red yellow green black/blue	
		white red yellow/green/blue black	white red yellow/green blue black	
		white red yellow/green black/blue		

Table 1: Partition hierarchy according to (Kay et al., 1997)

4 Statistical feature extraction

In this paper, I will present a statistical evaluation of the data of the WCS. My results largely confirm the generalizations that the mentioned authors achieved with non-statistical techniques, even though not all of the proposed universals could be confirmed.

As empirical basis for my investigation I only used the data from the WCS concerning the *extension* of color terms. The database contains data from 2,616 speakers from 110 languages. Each of these persons was asked to assign a term from their native language to each of the 330 Munsell chips, thus inducing a partition of this rendering of the color space. However, due to various problems in the elicitation of the data, some of these values are missing. I confined my analysis to the data from those speakers for whom all 330 data points are available. There are 1,771 such speakers from 102 languages. This number is still large enough for a meaningful statistical evaluation, so that an attempt to extrapolate missing values seemed unnecessary.

These data were organized in a contingency table. It has 1,601 rows—one for each term that was used by at least three[2] of the 1,771 test persons, and 330 columns—one for each Munsell chip. Each cell contains the number of test persons that used the term corresponding to the row to name the chip corresponding to the column. So schematically this table has the structure as indicated in Table 2 (of course English, German and French are not among the WCS languages).

	A0	B0	B1	B2	···	I38	I39	I40	J0
red	0	0	0	0	···	0	0	2	0
green	0	0	0	0	···	0	0	0	0
blue	0	0	0	0	···	0	0	0	0
black	0	0	0	0	···	18	23	21	25
white	25	25	22	23	···	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
rot	0	0	0	0	···	1	0	0	0
grün	0	0	0	0	···	0	0	0	0
gelb	0	0	0	1	···	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
rouge	0	0	0	0	···	0	0	0	0
vert	0	0	0	0	···	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Contingency table

Of the 330 Munsell chips, 322 are located at the surface of the color solid, while only the eight gray chips are from the interior. As a result, the interior is statistically under-represented in the WCS data. To partially remedy this problem, I copied the columns for the gray chips E0 and F0 three and for D0 and G0 twice in the contingency table. It thus has 340 columns.

Every row holds a 340-dimensional vector. The lengths of these vectors depend on the number of speakers that used the corresponding term. So even if two terms from different languages express the very same color category, they may be represented by substantially different vectors if the number of test persons from the two languages differs. I therefore divided each row by the number of speakers that used the corresponding term at least once. To ensure that frequently used terms have a higher statistical weight than terms that are used by few people, each row was afterwards copied as many times as there were speakers that used the corresponding term. The result of this transformation is a table with 12,415 rows, one for each speaker/term pair.

Each row represents a measurement of a naturally occurring color category. Variation between the rows thus represents variation between possible categories. However, the way the data are collected is prone to error. To illustrate this point, let us compare the term

[2] This decision was made to exclude idiosyncratic terms.

“enjaga” from language 36 (the Niger-Congo language Ejagam) with the term “tii” from language 105 (Yacoube, also a Niger-Congo language). The corresponding vectors are visualized in Figure 2. Elements of the vectors are represented by shades of gray here,

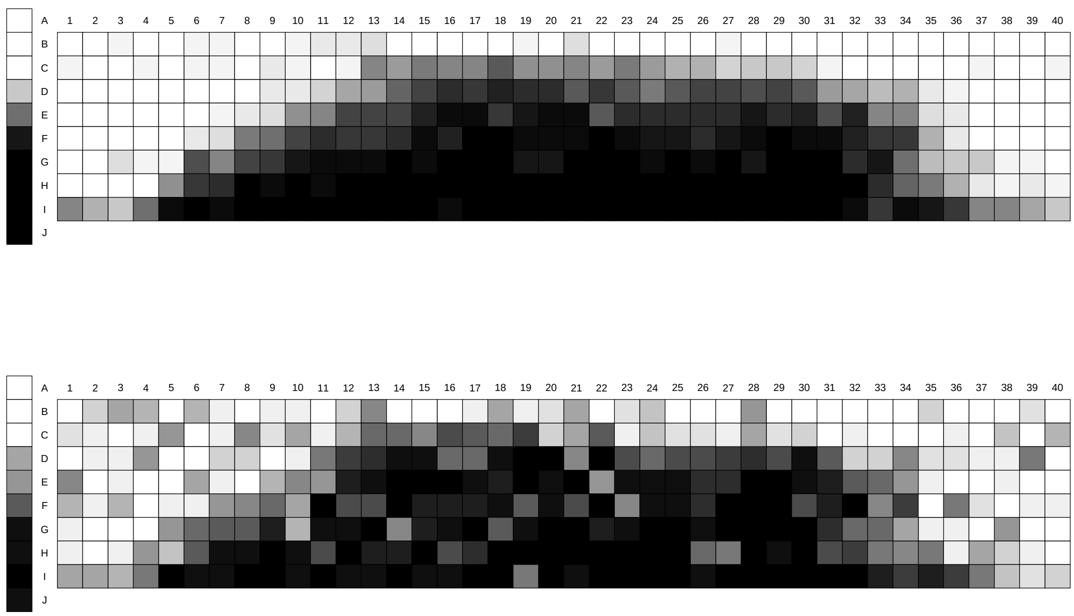

Figure 2: Extensions of “enjaga” (upper graphics) and “tii” (lower graphics)

where “white” means 0 and “black” 1. The vectors are scaled such that the maximal value is always 1.

It seems obvious that both terms denote the same concept — the category that comprises *black* and the cold colors, i.e. *green* and *blue*. Still, there are considerable differences in detail.

To separate genuine variation between categories on the one hand from random variation that is due to the method of data collection on the other hand, I employed **principal component analysis** (PCA), a standard technique for feature extraction and dimensionality reduction that is widely used in pattern recognition and machine learning (see for instance Jolliffe, 2002 for a comprehensive overview).

PCA is applicable to a collection of observations that can be represented as points in an n-dimensional space. For instance, each point could be a measurement and each dimension a feature that is being measured. Another application would be computer vision, where each observation is an image (i.e. intensity values for pixels) and each dimension represents one pixel.

PCA is a technique to reduce the dimensionality of such a data set. The goal is to project a high-dimensional data set onto a low-dimensional space in such a way that as

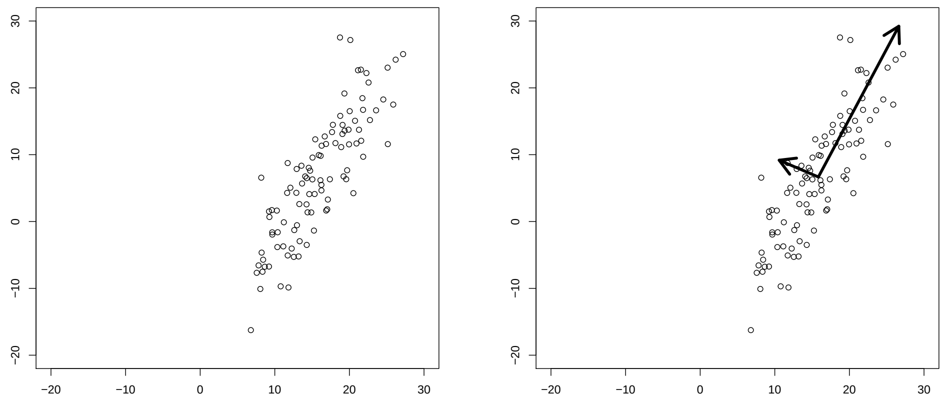

Figure 3: Principal Component Analysis: Illustration

much information that is present in the original data is preserved, while irrelevant noise is removed. The information within the data is conceived as its *variance*, i.e. the average squared pairwise distances between the data points. PCA shifts and rotates the coordinate system in such a way that it becomes easier to separate relavant from irrelevant dimensions.

Technically, PCA proceeds in the following steps (starting from a collection of points in an n-dimensional Euclidean space):

- The origin of the coordinate system is relocated such that it lies at the center of gravity of the data points.
- The direction is chosen where the variance of the data points is maximized. This means that all data points are projected onto some straight line going through the new origin, and the variance of the images of these projection is computed. The line which maximizes this value is the first **principal component** (PC1).
- All data points are projected onto the $n-1$-dimensional hyperplane that cuts through PC1 perpendicularly at the origin.
- The last two steps are repeated in this $n-1$-dimensional space, resulting in PC2 and a remaining $n-2$-dimensional hyperplane.
- This is repeated n times. PC1–PCn become the coordinates of the new coordinate system.

This is schematically illustrated with a two-dimensional space in Figure 3. The left panel represents the raw data in some arbitrary coordinate system. The right panel shows the new coordinate system, with the long arrow representing PC1 and the short arrow PC2.

Dimensionality reduction means that the first m principal components are considered important, while the variation along the remaining dimensions is considered as noise. There is no fool-proof way for picking an appropriate value of m; the statistical literature contains many useful heuristics though (see for instance Bryant and Yarnold 1995; details are discussed below).

This procedure was applied to the contingency table that was derived from the WCS data. It can be seen as a collection of 12,415 vectors in a 340-dimensional space. So each participant/term pair was treated as one point, and each Munsell chip as one dimension (with the caveat that some gray levels were counted twice or three times).

The proportion of the total variance of the data that is explained by the first 30 principal components (of 340 in total) is given in Figure 4. The first PC alone explains 31.3% of

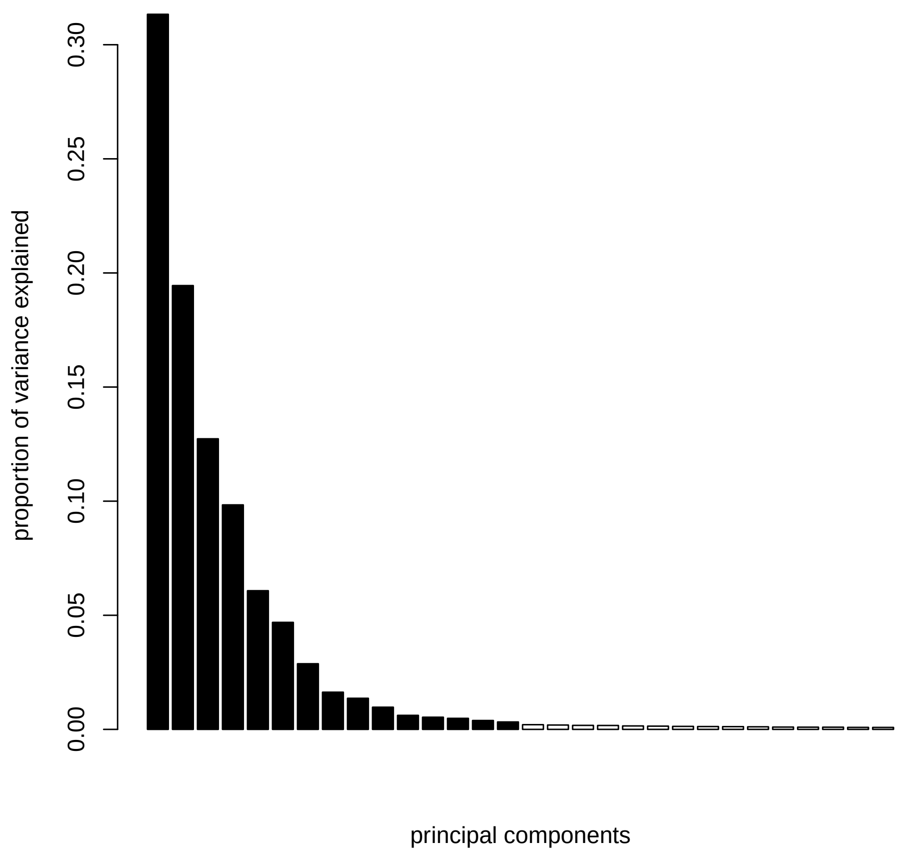

Figure 4: Proportion of total variance explained by principal components

the total variance, the second one 19.5%, and the third one 12.7%. These numbers quickly trail off. The 15th PC for instance only explains 0.3% of total variance.

As said above, it may be tricky to decide on a value of m if there is no *a priori* information how many degrees of freedom the "genuine" variation has. One criterion that is frequently recommended in the literature is "Kaiser's stopping rule" (see for instance Bryant and Yarnold 1995, p. 103). Applied to our data, this suggests a choice of $m = 15$. The first 15 PCs jointly explain 93.3% of the total variance.

To simplify the interpretability of the features (i.e. coordinates) thus extracted, I applied the **Varimax** algorithm — again a standard technique in feature extraction applications. This algorithm rotates the new low-dimensional coordinate system in such a way that the correlation of the new coordinates with the original coordinates is maximized.

The 15 new dimensions that are thus obtained are visualized in Figure 5. Each feature clearly defines a continuous region of the color space. Most of them can readily be named in English, even though not all of these names are morphologically simple: *green, white, red, yellow, black, blue, purple, pink, brown, sky blue, olive green, gray, orange, indigo,* and *pastel green*.

Figure 15 (see Appendix) gives another visualization. Here every original coordinate, i.e. every Munsell chip, is assigned to the new coordinate at which it has the highest value.

From now on we assume that people use these 15 features when they categorize colors. Variation along these 15 dimensions is thus relevant, while variation along the remaining dimensions is considered as noise.

To illustrate the effect of this dimensionality reduction, I projected the two vectors from Figure 2 to the new 15-dimensional space. Projecting the resulting two points on the 15-dimensional hyperplane back into the original 340-dimensional space yields the two vectors in Figure 6. The resulting vectors are in fact almost identical. To make this numerically precise: The cosine between the two original vectors (which is a common measure for the similarity of two vectors) is 0.965, while the cosine between their projections is 0.993.

5 The statistics of partition types

In this section I will use this dimensionality reduction technique to gain information about the partitions of the color space that can be found in the WCS data.

Let me start with a methodological remark. The evaluation of the WCS data by Kay et al. (1997); Kay and Maffi (1999) and others focused on how *languages* partition the color space.[3] Intra-linguistic variation between different speakers of the same language is not systematically taken into account. However, we sometimes find substantial and theoretically noteworthy differences between speakers of the same language. To make this point clear, consider the data from the speakers 1 and 17 from language 81 (Patep, an Austronesian language spoken in Papua New Guinea). They are visualized in Figure 16 (see Appendix). Here the color of the Munsell chips represent Patep terms. There is by and large agreement between the two speakers regarding the extension of the terms that could be translated as "white", "yellow", "red" and "green". However, the first speaker

[3] A noteworthy exception are Lindsey and Brown (2006), who perform a cluster analysis of the WCS data.

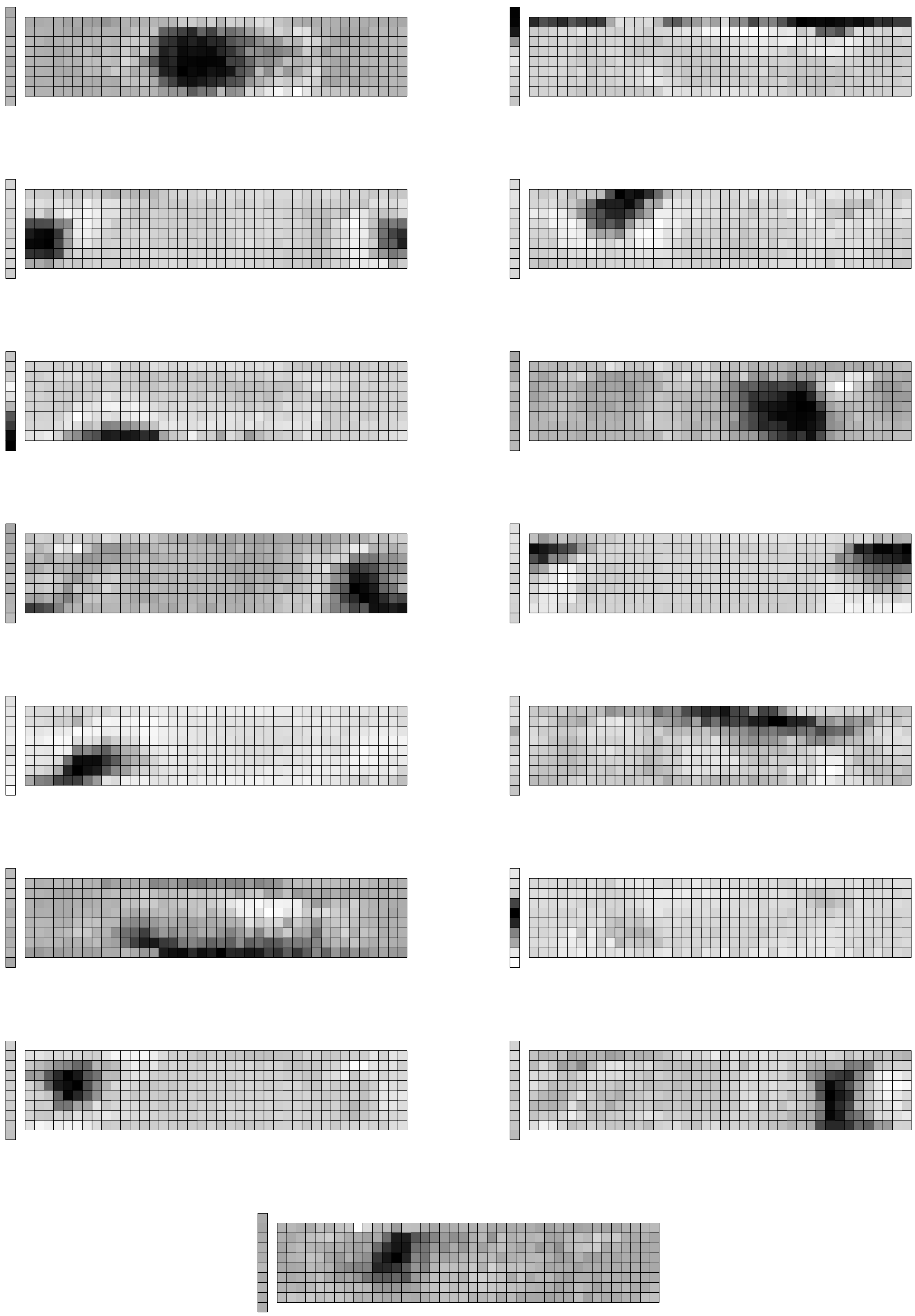

Figure 5: The 15 extracted features

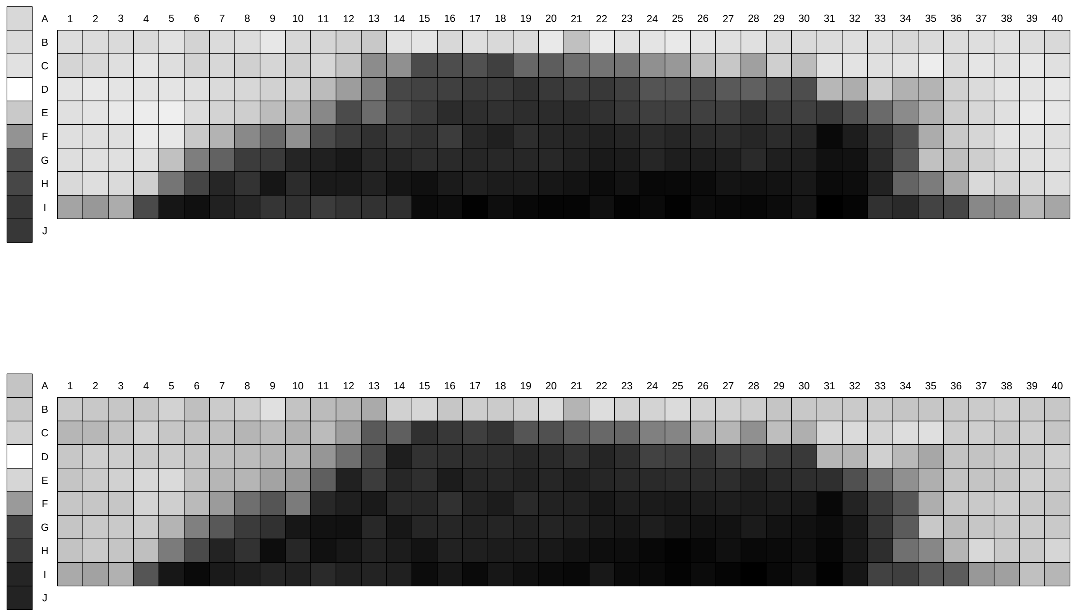

Figure 6: Extensions of "enjaga" and "tii" after dimensionality reduction

uses the same term for *black* and *blue*, while the second speaker reserved this term for *black* and uses another term for *blue* (which was not used at all by the first speaker). Careful inspection of the data revealed a great deal of such interesting intra-linguistic variation. Therefore I decided to investigate partition types on the level of individual speakers, not on the level of languages.

On the level of individual speakers, the extension of a term is a subset of the Munsell chips, i.e. a 340-dimensional vector containing just 0 and 1. A partition of the Munsell color space into k categories can thus be represented as a $340 \times k$-matrix P containing only 0 and 1, such that each row sums up to 1. Each column represents one cell of the partition.

The 15-dimensional coordinate system that was extracted from the data via PCA and Varimax can be represented as a 340×15-matrix $\mathbf{F}$. The i-th column of $\mathbf{F}$ represents the i-th dimension of the dimensionality reduced space.

The columns of P can be projected to the 15-dimensional space, yielding a k-tuple of 15-dimensional vectors. In the language of linear algebra, this can be represented as a $15 \times k$-matrix P', which is defined as

$$P' = \mathbf{F}^\top P,$$

i.e.

$$P'_{f,t} = \sum_{c=1}^{340} \mathbf{F}_{c,f} \times P_{c,t}$$

The rows of P' are the features and the columns are the partition cells, i.e. color terms. The entry of a cell is a number that intuitively represents how strong the partition cell overlaps with the feature.

From P' we can derive a partition over the 15 features. Each feature is assigned to the color term for which it has the highest entry in P'. Two features are considered equivalent iff they are assigned to the same color term. Formally, we define the equivalence relation between features as follows:

$$f_1 \equiv f_2 \text{ iff } \arg_t \max P'_{f_1,t} = \arg_t \max P'_{f_2,t}$$

For simplicity's sake, I will restrict the investigation to the first six features, that nicely correspond to the six primary colors *green, white, red, yellow, black*, and *blue*.

For each of the 1,771 speakers considered, we can extract a partition of the Munsell space P, which in turn leads to a partition of the six primary colors. The frequencies of the different partition types are distributed approximately according to a power law, as can be seen from Figure 7. The partition types are ordered according to their frequency.

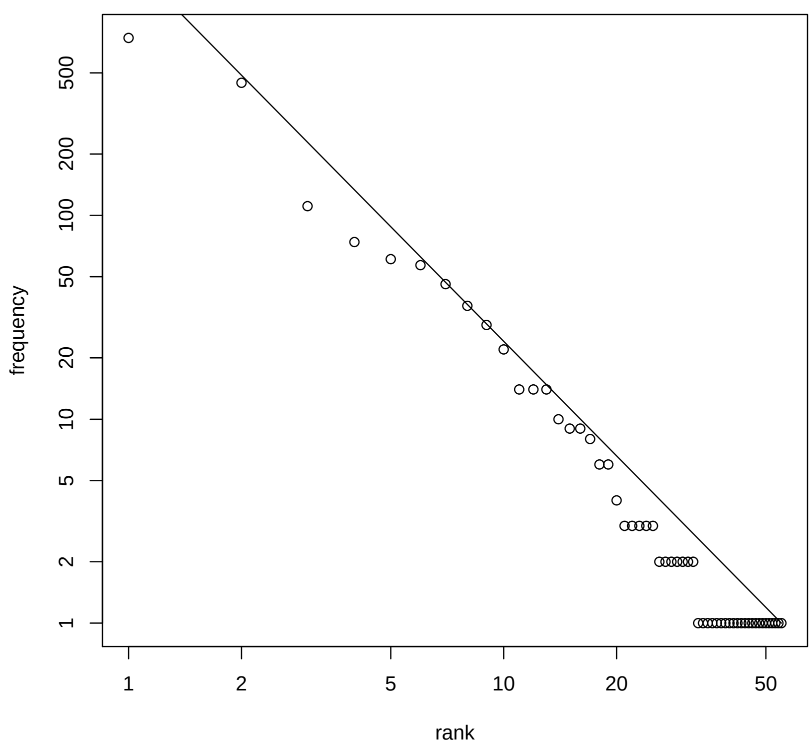

Figure 7: Frequency distribution of partition types over the primary colors

The figure gives the frequency of a partition type as a function of its rank in a log-log plot.

The distribution almost follows a straight line. This means that it can be approximated by a power law

$$fr \sim r^c,$$

where fr is the frequency, r is the rank, and c is a constant exponent. We have $c \approx -1.96$.

The three most frequent types are

- {white}, {red}, {yellow}, {green, blue}, {black} (742 occurrences)
- {white}, {red}, {yellow}, {green}, {blue}, {black} (447 occurrences)
- {white}, {red, yellow}, {green, blue, black} (111 occurrences)

All of the partition types from Table 1 except the leftmost one are attested. Additionally, there are two partition types not mentioned by Kay et al. (1997) which occur in substantial numbers:

- {white, yellow}, {red}, {green, blue}, {black} (61 occurrences; already noted in Kay et al. 1991)
- {white, yellow}, {red}, {green, blue, black} (36 occurrences)

Is mostly assumed in the literature (Kay et al. 1991 being an exception) that any composite color category that comprises *white* and *yellow* also has to include *red*. Our results seem to indicate that this generalization does not hold.

Manual inspection of the data revealed that there are indeed a few clear-cut instances of these types. Two examples are depicted in Figure 8. The upper figure represents the data from speaker 25 of language 60 (the Afro-Asiatic language Lele) and the lower figure the data from speaker 7 of language 17 (the Indo-Aryan language Bhili).

So while there are a few speakers that undeniably use a *white/yellow* category, most instances that were classified as belonging to such a type by the statistical procedure are more fuzzy. The data from speaker 25 of language 74 (the famous Amazonian language Pirahã) are perhaps more representative here (see Figure 17 in the Appendix). Here it is not possible to clearly assign *yellow* to any of the four terms. The various shaded of *yellow* are divided between the categories for *white*, *red* and *green/blue*.

The repercussions of this observations deserve further investigation. Quite possibly, *yellow* is not as clearly a landmark color as *red* or *green*.

These considerations notwithstanding, let us return to the outcome of the statistical evaluation of the various language types. Table 3 gives the frequencies of the types proposed by Kay et al. (1997), plus the two types including a *white/yellow* category mentioned above.

92.6% of all partitions considered fit into one of these 12 types. The remaining 7.4% are distributed over 43 partition types, each of which occurs at most 15 times, i.e. represents less than 1% of the data.

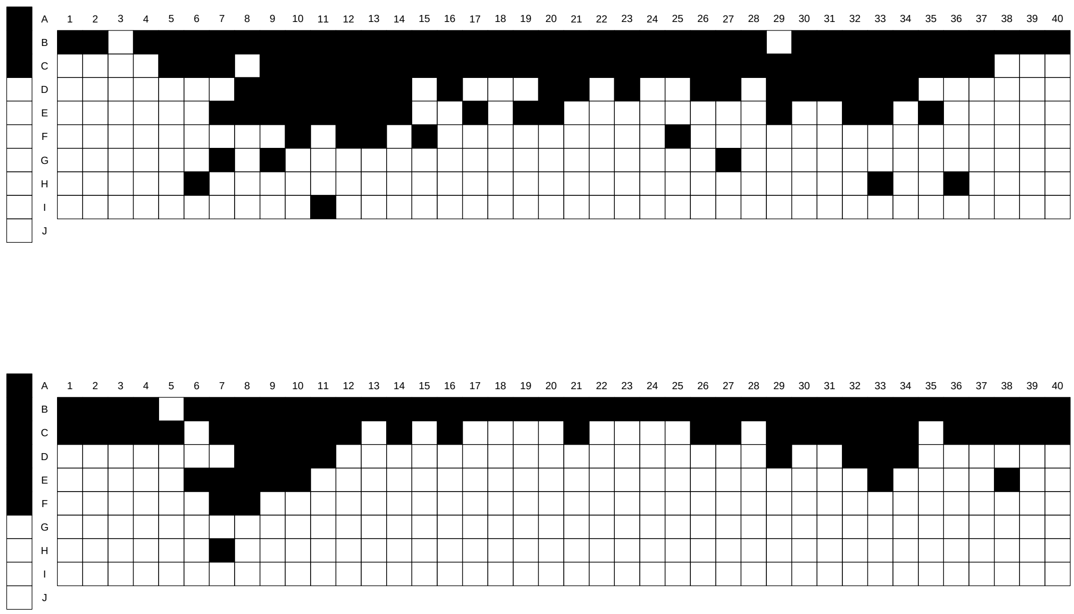

Figure 8: *White/yellow* composite categories excluding *red*

6 Towards a new model

Kay and McDaniel (1978) proposed that the fundamental ordering principle for the typology of color term systems is the distinction between *white/red/yellow* (warm colors) on the one hand and *black/green/blue* (cold colors) on the other hand — all partition types derive from this. Kay et al. (1991) noted, however, that a non-negligible number of languages use color categories that comprise *yellow* and *green* and thus transcend the boundary between warm and cold colors. According to the classification used here, 82 speakers out of 1,771, i.e. 4.6%, use such a category. Therefore the collection of partition types from Table 3 seems somewhat unsystematic.

In Kay et al. (1991) a rule for color category formation is formulated that is substantially more accurate if applied to our data. It is based on the diagram in Figure 9 (Figure 2 on page 15 in Kay et al. 1991). The authors propose the following rule, based on non-statistical evaluation of whole-language data of the WCS:

> "**Composite Category Rule:** A possible composite category is any fuzzy union of a subset of fundamental neural response categories which, in Figure [9], forms an unbroken associational chain not crossing the diagonal line."
>
> Kay et al. (1991), p. 16

The notion of a *fundamental neural response category* refers to the six primary colors (see my discussion in the next section regarding the neural substrate of color cognition).

I	II	III	IV	V
		[white red/yellow green/blue black] 46	[white red yellow green/blue black] 742	
[white/red/yellow black/green/blue] 0	[white red/yellow black/green/blue] 111	[white red/yellow green black/blue] 6		[white red yellow green blue black] 447
		[white red yellow black/green/blue] 57	[white red yellow green black/blue] 74	
		[white red yellow/green/blue black] 29	[white red yellow/green blue black] 9	
		[white red yellow/green black/blue] 22		
	[white/yellow red green/black/blue] 36	[white/yellow red green/blue black] 61		

Table 3: Frequencies of the various partition types

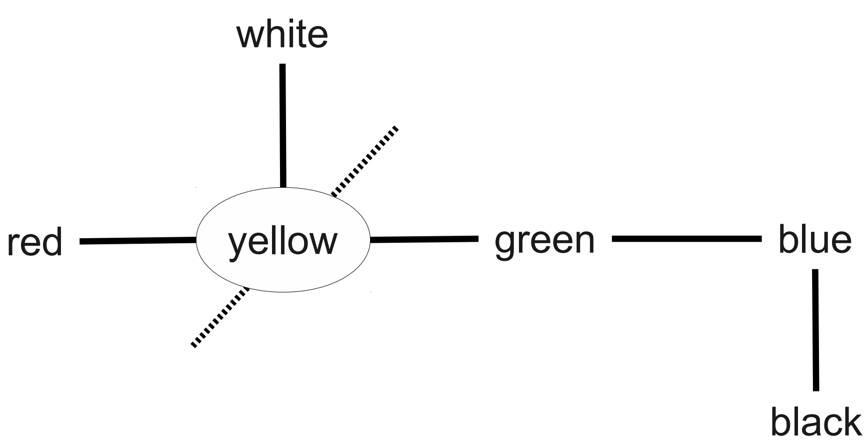

Figure 9: Association graph; from Kay et al. (1991)

Kay et al. (1991) note that of the nine composite categories that are admitted by this rule, eight are actually attested.[4] The only one that is unattested is the four-element category {yellow, green, blue, black}. Our procedure did not produce a single example of such a category either. This can be stated as another constraint:

> **At Most Three:** No composite category comprises more than three primary colors.

The mentioned authors do not further investigate the question which partitions types are admitted by the CCR (Composite Category Rule). All thirteen partitions in Table 3 are in fact addmited. Furthermore, there are three licit partitions obeying AM3 (At Most Three) that are attested and three that are not attested in the data:

- {green}, {white/yellow}, {red}, {black/blue} (14 occurrences)
- {green}, {white/yellow}, {red}, {black}, {blue} (8 occurrences)
- {green}, {white}, {red/yellow}, {black}, {blue} (2 occurrences)
- {green}, {white/yellow/red}, {black/blue} (0 occurrences)
- {green/blue}, {white/yellow/red}, {black} (0 occurrences)
- {green}, {white/yellow/red}, {black}, {blue} (0 occurrences)

The four partition types that are predicted but do not occur (the three mentioned above, and the leftmost one in Table 3) all contain the hypothetical composite category {white/ yellow/red}. The bipartite partition {green/black/blue} vs. {white/yellow/red} has been reported by Eleanor Rosch Heider for the Papua language Dani (Heider 1972), and this system is assumed to be basic for all other systems in the research following Kay and McDaniel (1978). The existence of composite categories comprising *yellow* and *green* show, however, that not all attested partition types are refinements of the Dani system. On the basis of the WCS data, the composite category {white/yellow/red} is unattested. Therefore I propose a third constraint:

> ***CompositeRed:** Red co-occurs with at most one other primary color in any composite category.

With this proviso, all admitted partition types are in fact attested in the data, and 93.6% of all data points fall into one of these categories.

The fit of the model can actually be further improved if a link between *green* and *black* is added to Kay et al.'s association graph. This modification licenses four more partition types, which are all attested:

[4] The composite category {white/red/yellow} does not occur in the WCS, but it has been reported for other languages.

- {green/black}, {white}, {red}, {yellow}, {blue} (14 occurrences)
- {green/yellow/black}, {white}, {red}, {blue} (3 occurrences)
- {green/black}, {white/yellow}, {red}, {blue} (3 occurrences)
- {green/black}, {white}, {red/yellow}, {blue} (2 occurrences)

The existence of a category that includes *black* and *green* but excludes *blue* may be surprising. Such a system is illustrated by the data from speaker 7 of language 1 (Abidji, a Niger-Congo language spoken in Ivory Coast); see Figure 10.

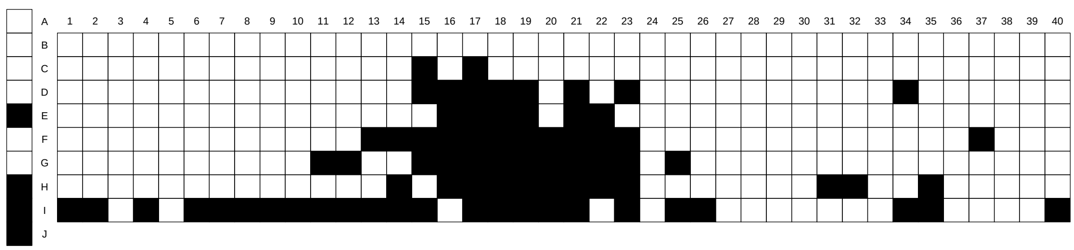

Figure 10: A *black/green* category excluding *blue*

Comparing this to the data from the other Abidji speakers showed that the term that speaker 7 exclusively uses for the *black/green* category is used by other speakers for the less exotic category *black/green/blue*. Additionally, the language has a specialized term for *blue*. The apparent *black/green* composite category for speaker 7 may thus be an artifact of the way the data were elicited. Perhaps the term in question ("lobu") also means *black/green/blue* for speaker 7, but he or she chose the more specific term for *blue* ("gale") where possible and used the more general term only if no more specific one was available.

If this account is on the right track, this would cast doubt on the assumption that the basic color vocabulary of a language necessarily partitions the color space, because Abidji apparently has color terms with overlapping extensions.

Be this as it may, at the level of the raw data the *black/green* category is attested.

The fit of the model is improved even more if a link between *white* and *black* is added. This modification licenses four more partition types, two of which are attested:

- {green/blue}, {white/black}, {red}, {yellow} (14 occurrences)
- {green}, {white/black}, {red}, {yellow}, {blue} (9 occurrences)
- {green/yellow}, {white/black}, {red}, {blue} (0 occurrences)
- {green}, {white/black}, {red/yellow}, {blue} (0 occurrences)

The existence of a *black/white* composite category may seem even more surprising. There are undeniable instances though. Figure 11, displaying the data for speaker 1 of language 82, the Meso-American language Paya, may serve as illustration.

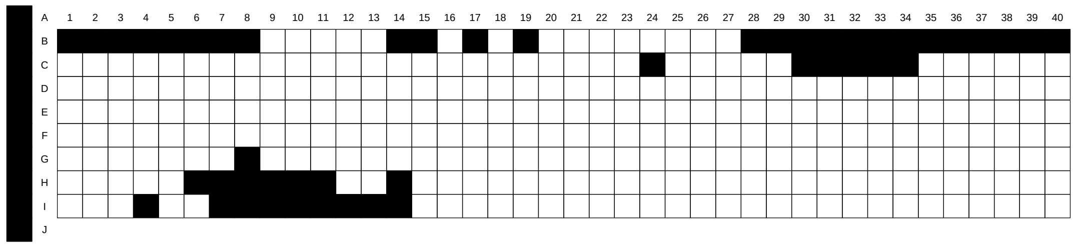

Figure 11: A *black/white* category

The modified association graph is given in Figure 12. The constraints CCR, AM3 and

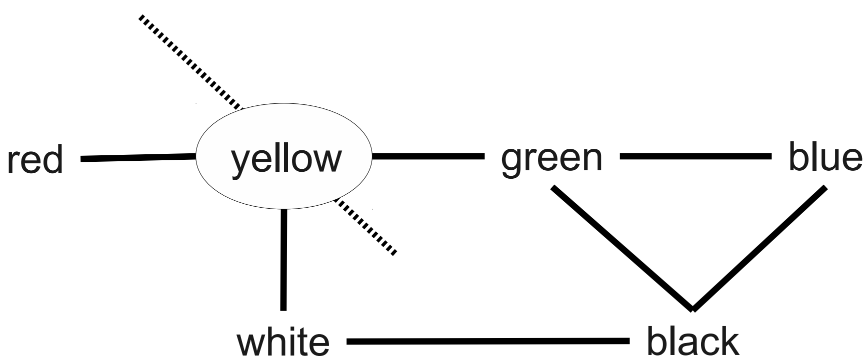

Figure 12: Modified connection graph

*CompositeRed, applied to this connection graph, licenses 23 different partitions, 21 of which are empirically attested. The data from 1,703 out of 1,771 speakers belong to one of the 23 licensed partition types, this are 96.2%.

7 Is there a perceptual/neurophysiological basis for universal constraints?

Since the typological variation in color naming systems is so severely limited, an obvious question is whether the universal constraints that we observe in the data have a basis in the cognitive organization of color perception and categorization, and ultimately in neurophysiology. At least since Kay and McDaniel (1978), this issue has been discussed extensively, but so far no conclusive consensus has been reached. Since this is not the main issue of this paper, I will touch upon it only briefly.

The model proposed here has an obvious base in the properties of the psychological color space, even though the latter underdetermines the former. As pointed out in Section 2, the color solid (i.e. the psychophysical representation of all perceivable colors in a geometric model that represents similarity as spatial proximity) is a three-dimensional object (see Figure 14). The three coordinate axes are usually identified with the oppositions *white/black*, *red/green* and *yellow/blue*. The six primary colors thus represent the vertices of an irregular octahedron. With the exception of the link between *white* and *black*, all associations in Figure 12 thus correspond to edges of this octahedron. The distance between *red* and *blue* is larger than the distance between the other pairs of chromatic primary colors. This may account for the absence of a direct connection between *red* and *blue*. Likewise, *yellow* is very close to *white* in the color solid, which may explain the privileged association between these two colors.

Other aspects of the topology of the graph in Figure 12 do not directly correspond to perceptual facts. The strong affinities between *green* and *blue*, and between *blue* and *black* apparently do not follow from the geometry of the color solid. Likewise, it is unclear why *red* only co-occurs with *yellow* in composite categories. Also, the constraint that is indicated by the dotted line — the generalization that *yellow* may associate with *green* or *red* or *white*, but never with two of those simultaneously — seems mysterious. Finally, the fact that a non-negligible portion of the participants of the WCS use a composite category for *black* and *white* — which motivates the corresponding link in the association graph — seems distinctly odd. Arguably these categories are no regular color categories but refer to a one-dimensional conceptual space related to saturation or chromaticity.

There are suggestive correlations between the structure of the psychological color space and the neurophysiology of color perception, but at the current state, the issue is far from being resolved. The human retina contains two kinds of photopigments, rods and cones. Only cones are sensitive to color differences. There are three kinds of cones, which are, roughly speaking, responsible for the perception of red, green and blue light respectively. Subjective color differences thus correspond to differential excitation levels of these three kinds of receptors, which explains the fact that the psychological color space has three dimensions.

The particular shape of the human color solid is not determined by the retina though. Higher aspects of vision are neuroanatomically located in the visual cortex, which is located near the back of the skull. Between retina and visual cortex, important processing steps take place in the *lateral geniculate nucleus* (LGN) of the thalamus. A special kind of cells there — P-cells — are operative in color perception. It has frequently been assumed that P-cells provide a direct substrate to psychological notions of primary colors. Research like de Valois and Jacobs (1968) indicated that P-cells of macaques fall into two pairs of opposing types. +R-G-cells respond positively to red and negatively to green stimuli, while -R+G-cells have exactly the opposite response pattern. Likewise, +Y-B-cells and -Y+B-cells respond in opposite ways to yellow and blue stimuli respectively. It is suggestive to link these two types of opposing cells to the two chromatic axes of the perceptual color space.

However, more recent research has shown that this picture is too simple. For instance, all P-cells in humans respond to achromatic white light. Also, the 0-point where neither of the four P-cell types are excited or inhibited, is subjectively perceived as a hue sensation between yellow and green. So while the LGN does play a crucial role in color vision, the precise neurological substrate of color psychology is still elusive. (For more details, see for instance the discussion in Abramov 1997 and in Jameson and D'Andrade 1997.)

8 Conclusion

While the model proposed above is concise and seems to give a good fit of the empirical data, I would like to close with some cautionary notes.

- The method that I used to classify the data from individual speakers has several advantages. It can be automatically applied — and thus provides a massive amount of data with comparatively little effort, and it does not suffer from theoretical biases. On the other hand, it is somewhat crude and gives clear results even for cases where the data actually aren't clear at all. My discussion of the *white/yellow* and the *black/green* categories from the previous section point to some problematic cases where the method suggests generalizations that are perhaps questionable. Before an explanatory model can be attempted, such cases should be investigated in more detail. A combination of more fine-grained statistical methods and manual annotation seems to be promising.

- I argued above that the variation between speakers of the same language is often theoretically significant and should not be neglected by simply averaging over all speakers of a language. Therefore I only used the data from individual speakers. I thereby ignored the information about the variation of the extension of the same term between the speakers of the same language. The informal discussion of the *black/green* issue above indicates that such information is also relevant. A more detailed investigation should take both the level of individual speakers and of languages into account.

- It is tempting to extend the method used here beyond the six primary colors, perhaps applying it to all the 15 features that were automatically extracted. In an exploratory study I found that the connection graph for all 15 features, with very few exceptions, connects two colors if and only if they are adjacent in the CIELab space. This might seem encouraging because it allows to predict parameters of the model from an independent information source. However, demanding that each category is a connected subgraph of the thus enlarged connection graph only amounts to the insight that color categories are mostly continuous regions of the CIELab space. Such a model is much less restrictive than the one presented here. For instance, there are *red/pink/orange* composite categories, so the constraint that *red* combines at most with one other color cannot be maintained in the larger model. But since there are also *white/pink* composites, for instance, the absence of a *red/pink/white* composite category cannot readily be accounted for anymore.

These considerations should stress the exploratory nature of this study. Nevertheless I would like to argue that statistical techniques like automatic feature extraction provide valuable information for typological studies. It will certainly not replace more traditional methods, but augment them.

References

Abramov, Israel (1997); 'Physiological mechanisms of color vision'. In Clyde L. Hardin and Luisa Maffi (eds.), *Color Categories in Thought and Language*, Cambridge University Press, Cambridge (UK), 89–117.

Berlin, Brent and Paul Kay (1969); *Basic color terms: their universality and evolution.* University of California Press, Chicago.

Bryant, Fred B. and Paul R. Yarnold (1995); 'Principal-Components Analysis and Exploratory and Confirmatory Factor Analysis'. In Laurence G. Grimm and Paul R. Yarnold (eds.), *Reading and Understanding Multivariate Statistics*, American Psychological Association, 99–136.

Cook, Richard, Paul Kay, and Terry Regier (2005); 'The world color survey database: History and use'. In Henri Cohen and Claire Lefebvre (eds.), *Handbook of Categorisation in the Cognitive Sciences*, Elsevier, chapter Amsterdam, 223–242.

de Valois, Russell and Gerald H. Jacobs (1968); 'Primate Color Vision'. *Science* **162**:533–540.

Heider, Eleanor Rosch (1972); 'Universals in color naming and memory'. *Journal of Experimental Psychology* **93**:10–20.

Jameson, Kimberly and Roy G. D'Andrade (1997); 'It's not really red, green, yellow, blue: an inquiry into perceptual color space'. In Clyde L. Hardin and Luisa Maffi (eds.), *Color Categories in Thought and Language*, Cambridge University Press, Cambridge (UK), 295–319.

Jolliffe, Ian T. (2002); *Principal Component Analysis.* Springer.

Kay, Paul, Brent Berlin, Luisa Maffi, and William Merrifield (1997); 'Color Naming Across Languages'. In C. L. Hardin and Lusia Maffi (eds.), *Color categories in thought and language*, Cambridge University Press, Cambridge, 21–58.

Kay, Paul, Brent Berlin, and William Merrifield (1991); 'Biocultural implications of systems of color naming'. *Journal of Linguistic Anthropology* **1**:12–25.

Kay, Paul and Luisa Maffi (1999); 'Color appearance and the emergence and evolution of basic color lexicons'. *American Anthropologist* 743–760.

Kay, Paul and Chad K. McDaniel (1978); 'The linguistic significance of the meanings of basic color terms'. *Language* **54**:610–646.

Kay, Paul and Terry Regier (2003); 'Resolving the question of color naming universals'. *Proceedings of the National Academy of Sciences* **100**:9085–9089.

Lindsey, Delwin T. and Angela M. Brown (2006); ‘Universality of color names’. *Proceedings of the National Academy of Sciences* **103**:16608–16613.

Regier, Terry, Paul Kay, and Richard S. Cook (2005); ‘Focal colors are universal after all’. *Proceedings of the National Academy of Sciences* **102**:8386–8391.

Roberson, Debie, Ian R. L. Davies, and Jules Davidoff (2000); ‘Color categories are not universal: Replications and new evidence from a Stone-age culture’. *Journal of Experimental Psychology: General* **129**:369–398.

Appendix

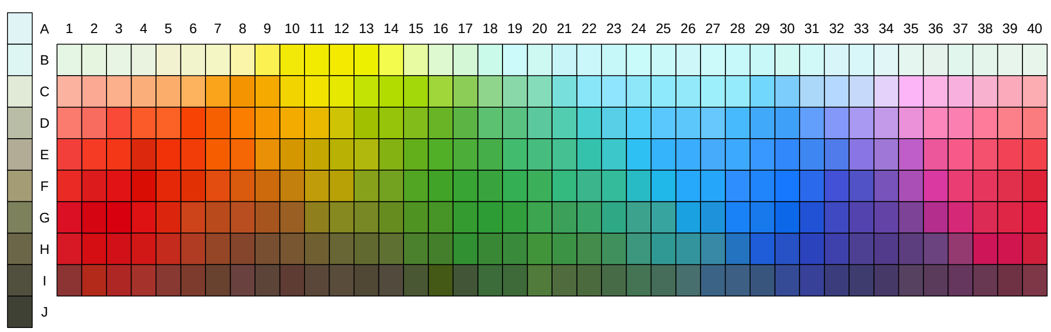

Figure 13: The Munsell chart

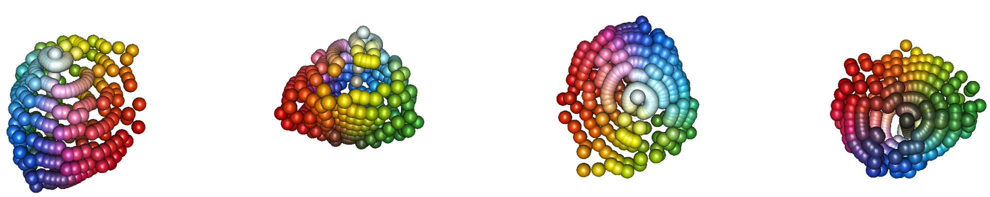

Figure 14: The CIELab color solid

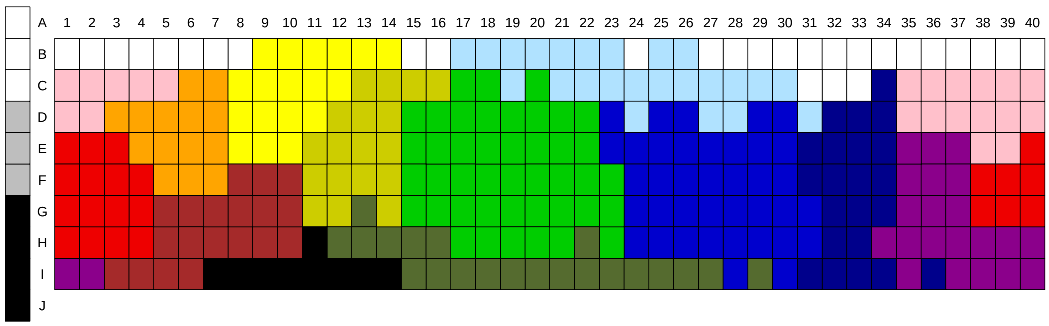

Figure 15: Partition induced by the extracted features

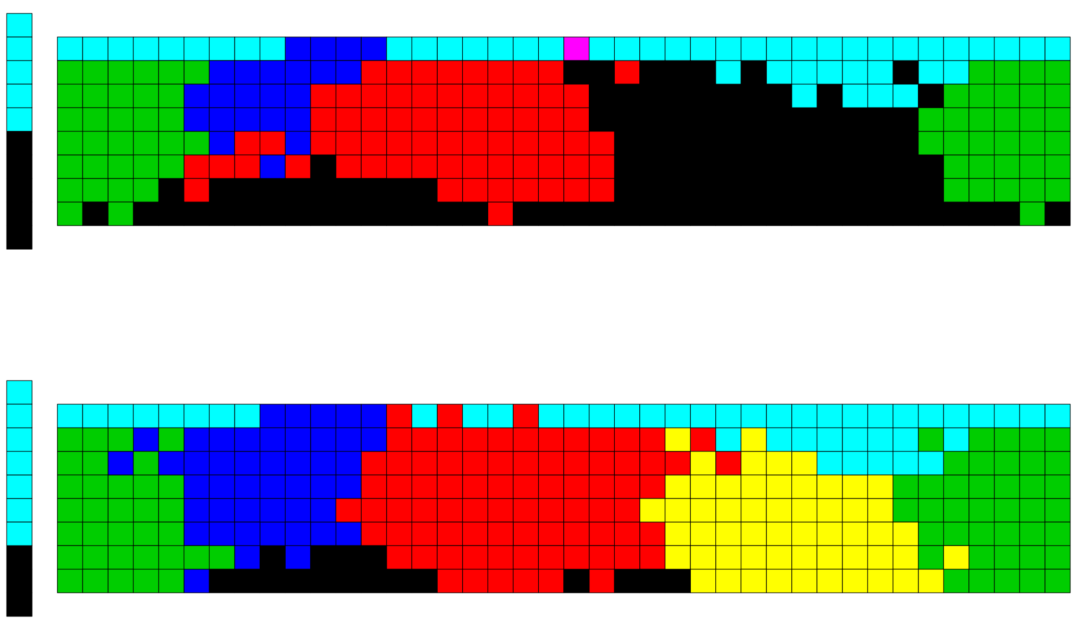

Figure 16: Intra-linguistic variation: two speakers of Patep

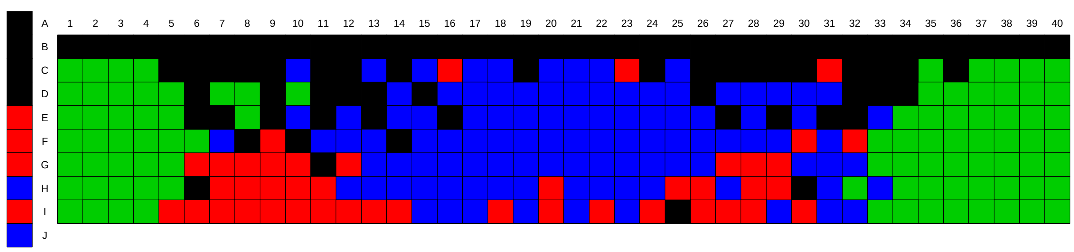

Figure 17: Fuzzy categorization of *yellow*