UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY FYZIKY A INFORMATIKY
KATEDRA APLIKOVANEJ INFORMATIKY

BASIC COLOR CATEGORIES

(diplomová práca)

Kristína Rebrová

Študijný odbor: 9.2.8 Umelá inteligencia a 3.1.9 Psychológia
Študijný program: Kognitívna veda
Školiteľ: **RNDr. Martin Takáč, PhD.**
BRATISLAVA 2009

Prehlásenie

Čestne prehlasujem, že som túto prácu vytvorila samostatne, vedená radami a usmernením môjho školiteľa, čerpajúc z uvedenej literatúry a zdrojov dostupných na internete.

Hereby I declare that I wrote this thesis myself with guidance and advice of my supervisor and with use of referenced sources.

V Bratislave, dňa 7.5. 2009

....
Kristína Rebrová

Poďakovanie

Chcela by som sa v prvom rade poďakovať môjmu školiteľovi Martinovi Takáčovi, za cennú inšpiráciu, jeho cenné rady, ochotu a trpezlivosť pri tvorbe tejto práce. Tiež chcem vysloviť vďaku mojim rodičom, ktorí ma trpezlivo podporovali počas celého štúdia, mojim kamarátom a v neposlednej rade Ľudovi Malinovskému.

At first I would like to thank my supervisor Martin Takáč for inspiring me to do this thesis, for his valuable advice, willingness and patience. My thanks go also to my parents, who supported me during the whole study, to my friends and especially to Ľudo Malinovský.

Abstract

The main theme of this master thesis is the color categorization. In its theoretical part it provides information about phenomena underlying the categorization of colors, the color perception and color naming and about main hypothesis from this field. The basic color categories are named with basic color terms, which can be defined as a set of simple words with which all speakers of a language can consistently name any color. The main hypotheses proposed about the basic color categories assume that there is only a restricted amount of basic color terms and that their best examples are universal for any language. The World Color Survey, consisting of color naming experiments with speakers of 110 primitive languages with no written form and of non-industrialized cultures, has been carried out to asses these hypotheses. This master thesis works with data collected in WCS and in its imitation in Slovak language. It proposes innovative methods of visualization of color naming data and provides also software for generating such visualization. Another view on usage of the WCS data is encompassed in the color categorization simulation employing the semantics of distinguishing criteria aimed to probe the learning capability of the distinguishing criteria on real data.

Keywords: categorization, color perception, basic color terms, World Color Survey, distinguishing criteria

Abstrakt

Ústrednou témou tejto diplomovej práce je kategorizácia farieb. Jej teoretické časti ponúkajú potrebný vedomostný základ o dvoch hlavných témach zapojených do skúmaného fenoménu a tými sú farebné vnímanie a kategorizácia, ako aj teóriu týkajúcu sa priamo skúmanej problematiky. Základné farebné termíny, ktoré pomenúvajú základné farebné kategórie, môžeme definovať ako univerzálne akceptovanú množinu jednoduchých slov pomocou ktorých je možné opísať akýkoľvek farebný vnem. Primárne hypotézy týkajúce sa základných farebných termínov tvrdia, že ich v akomkoľvek jazyku existuje len obmedzené množstvo a že farby, ktoré predstavujú tieto kategórie ako ich najlepší reprezentanti sú približne rovnaké vo všetkých svetových jazykoch. S cieľom overiť tieto tvrdenia na plauzibilnej experimentálnej báze vznikla svetová farebná štúdia (World Color Survey), vykonaná v 110 jazykoch bez písomnej formy a z neindustrializovaných spoločností. Dáta z tejto štúdie, ale aj z jej napodobeniny v slovenskom jazyku, sú hlavným predmetom praktickej časti tejto diplomovej práce, ktorej náplňou je jednak dizajn inovatívnej vizualizácie týchto dát a druhak ich použitie v simulácií kategorizácie farieb na báze sémantiky rozlišovacích kritérií. Cieľom tejto simulácie nie je preverovať spomínané hypotézy, ale kategorizačnú schopnosť rozlišovacích kritérií s pomocou reálnych dát. K výstupom tejto práce patrí aj softvér na generovanie vizualizačných obrázkov z dát z WCS.

Kľúčové slová: kategorizácia, farebné vnímanie, základné farebné termíny, World Color Survey, rozlišovacie kritéria

Table of contents

List of figures

1 Introduction

The human thought and cognition is in close relationship with our ability to categorize. The categorization is employed in perception, reasoning, planning and action, speech, and any other cognitive tasks. Similar importance is attributed to the language ability, which can be considered a tangible feature discriminating the quality of human cognition from cognitions of any other species. Therefore the study of basic categorization mechanisms within language could be a way to understand the cognition and thought as such. Within cognitive science, cognitive linguistic school [1] is concerned with such notions.

The main theme of this master thesis is color categorization. We have chosen this topic as a good example of a study of categorization based on the language. The importance of color categorization rests in its universality. Our visual perception can be considered the most important of our senses, the process of color discrimination must be present in our thought, and consequently in language. Since the color perception is a general property of human sight, the ability to discriminate colors is not a culturally specific phenomenon. Therefore we assume that the color categorization is an omnipresent phenomenon, which can be studied and compared across cultures and languages. The studies of color categorization are closely connected to the emergence of new paradigms of principles of categorization, namely the prototype theory.

In the following chapters we will provide an essential background for understanding the phenomena involved in color categorization, the central theory about color categorization and related research, and finally we will describe our contribution to this topic. At first we will present the basics of human color perception, i.e. the main theories of color vision, its physiology and psychology and a short overview on how the color can represented by a color space. The second theoretical chapter is dedicated to categorization, mainly to the prototype theory of categorization. It also provides a definition of the basic color categories and relevant background on color categorization as such.

When aware of the underlying theory on color vision and categorization the reader can continue to the heart of this master thesis. In the fourth chapter we will describe the main hypotheses regarding the basic color terms, their universality and evolution and the World Color Survey aimed to validate, invalidate or modify these hypotheses. The data from WCS consist of the results from color naming experiment from 110 primitive languages with no written form and of non-industrialized cultures. The next part of this chapter consists of a short overview on an imitation of WCS experiment in Slovak language we carried out in spring 2008. At last, we provide an overview of selected studies regarding the data from the WCS, aimed to verify the original hypotheses. We also describe the first visualization of these data, which inspired us to create a new one.

In fifth chapter we propose an innovative visualization method. It consists of colored maps that resemble the stimulus material from the WCS experiment aimed to facilitate the imagination of its task and results. We developed three various complementary ways of displaying the data, employing various dimensions to provide the maximum information. In this chapter we also describe a stand-alone application we programmed, which creates these visualizations in form of .jpg images according the preferences the user sets in graphical user interface. This application, as a software output of this master thesis is attached on a CD (see Appendix C).

The last core chapter describes the color categorization simulation we developed employing the semantics of distinguishing criteria and the results of this simulation carried out using the data from WCS and from Slovak color experiment. The aim of this simulation is not to prove or disprove the original hypothesis on basic color categories, but rather to probe the learning capability of the distinguishing criteria on real data.

2 Color perception and color coding

This chapter provides an introduction to basic principles of human color perception, basic theories of color vision, its physiology and psychology. In short, the visual apparatus perceives color by means of absorption of light by three different types of photoreceptor. The essential mechanism involved in many aspects of color processing is comparative principle. It can be found either on primitive level, where the response of each type of color-sensitive receptor is formed in comparison with reactions of other types of receptors, or on cognitive level, where the color of an object is distinguished based on comparison with its surroundings. Another principle present on every level of color perception is the principle of color opponency, generally known as opponent process theory, based on antagonistic relationship between pairs of primary colors, red-green, blue-yellow and black-white.

2.1 Basic principles of color perception

Color vision is the ability to detect and analyze changes in composition of the wavelength of light [2]. The definition of color in general is an undecided problem. On one hand, it is a property or a byproduct of special range of electromagnetic radiation. Color percepts then can be described as psychophysical properties of perceived object or material. On the other hand, color can be defined as a property of mind. It is possible to imagine certain color with our eyes closed. Therefore to describe color in general without binding on certain light and material conditions we describe it as a perceptual entity with three components – hue, value and saturation, where hue encompasses the quality of the percept. The other parameter, value (intensity, lightness or brightness) stands for the intensity of (white) light apparently coming from the colored object. When the light is at its fullest intensity, colors will become bright, at its least intensity, colors become dim. At last saturation, chroma or colorfulness tells us about the purity of color or more precisely about the amount of white light (or gray paint) mixed with the hue.

The visible spectrum (or simply light) is the portion of the electromagnetic spectrum that is visible to human eye. A typical human eye will respond to wavelengths from about 380 to 750 nm [3][1]. A very common mistake is an idea, that only the wavelength of light falling on perceived object determines the color perceived. This principle works for so-called aperture colors studied by scientists examining the most basic aspects of color vision in laboratory conditions, but most probably nowhere else [4]. Visual perception is in principle based on the comparison of focused objects with their background. Our perceptual system is not able to "count" the specific wavelengths, only compare them among objects perceived.

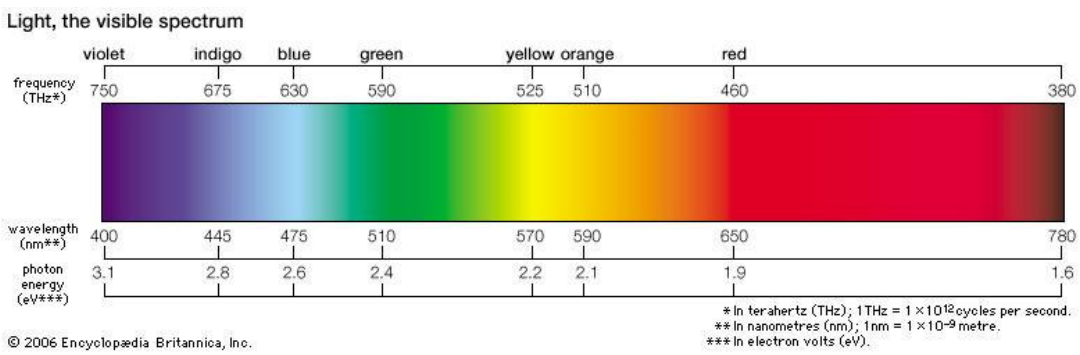


Figure 1: The visible spectrum

This figure depicts the colors of the visible spectrum and corresponding vawelengths and is taken from [5].

[1] The spectrum does not, however, contain all the colors that the human eyes and brain can distinguish. Unsaturated colors, for example pink and purple are absent, because they can only be "made" as a mixture of multiple wavelengths.

2.2 Theories of color vision

There are two generally accepted theories of color vision, both formulated in 19th century and both right in some aspects, while considered to be rival to each other. This section is based on [6] and [7].

The first, trichromatic theory (or Young–Helmholtz theory) suggests that there are three types of color receptors (red, green, blue) sending the values of their excitation the brain according to the color of the perceived light. Note that it was formed before physiological evidence for this phenomenon was found and described. The theory is based primarily on color mixing experiment and suggests that a combination of three channels is sufficient for creating any color. Although it truly corresponds with the three types of color sensitive receptors, this concept fails to explain the uniqueness of four color primaries (explained by the opponent process theory described below), and also why dichromats (people missing one type of color cones) are able to see white and yellow even though it is impossible to mix these colors with one channel missing.

The other theory, which is also important for the further content of this thesis, is the opponent process theory. This one as well distinguishes three perceptual channels, but unlike trichromatic theory it defines six primary colors (also called Hering primaries after the author of this theory) into three antagonistic channels: red-green, blue-yellow and black-white (rather dark-light). It means that at any time either red or green is perceived, but never greenish-red (the same with blue and yellow). This principle easily explains the color-afterimage and other psychological phenomena (described in chapter 2.4), but in its original form it never challenged the trichromacy of initial stage of color processing expressed by the first theory.

The experimental background for this theory was provided by Hurvich and Jameson in 1957. Using hue cancellation method, they isolated psychophysical color opponent channels responsible for the antagonistic mechanisms in the perception of these unique colors. The figure below describes the chromatic response cancellation curves for the three Herring's color channels. Note that there is no single wavelength at which a pure red can be perceived. Hence the pure red is extra-spectral and can be perceived only when the yellow component is canceled by the blue component of specific wavelength.

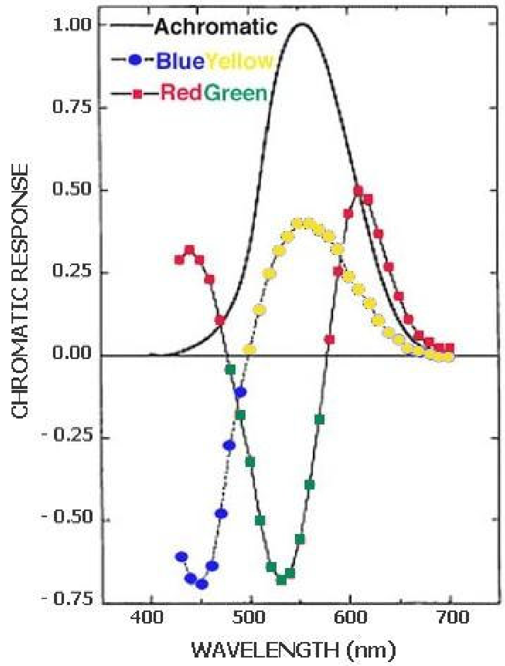


Figure 2: Chromatic response cancellation curves for the three opponent channels
Hurvich and Jameson experiment using blue or yellow and red or green to match all wavelengths of the visible spectrum. The figure is taken from [7] and slightly modified for better comprehension (we changed the colors of points in the graph to resemble the red-green and blue-white distinction).

Some later studies considering color opponent processes summarized in [8] suggest that at physiological level, the pairs of mutually inhibiting color primaries are not exactly red-green and blue-yellow, but rather cherry-teal and chartreuse-violet or some simple rotation of these axes (similar hues with equal spacing in adequate color space). The uniqueness of original colors proposed is explained as resulting from the irregularity of shape of perceptual color space (described in section 2.5).

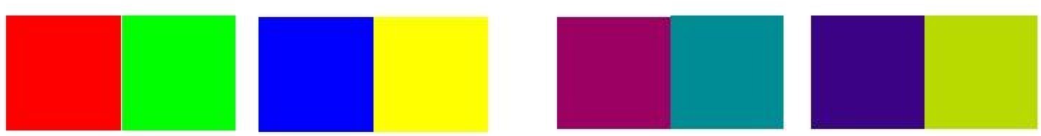


Figure 3: Illustration of opponent process hues
This figure displays color hues which might be considered the classic opponent hues: red, green, blue and yellow and the new hues proposed above: cherry, teal, violet, and chartreuse. Note that this is only an illustration.

Concluding the proposals of these two basic theories of color vision we can say that they are not opponent, but complementary. Firstly, there is the trichromacy of receptors in retina; secondly there are opponent processes in the neural pathways visual cortex. On the other hand, the opponent process theory is valid also in three-type cones part, since yellow is sensed both by red and green cones, but not by blue cones (which's response is mutually inhibited with the "yellow signal"). More details are provided in the following section. The figure below (taken from [5]) displays a scheme encompassing both theories merged together in so-called stage theory, which states that both main theories work, but on a different level of processing.

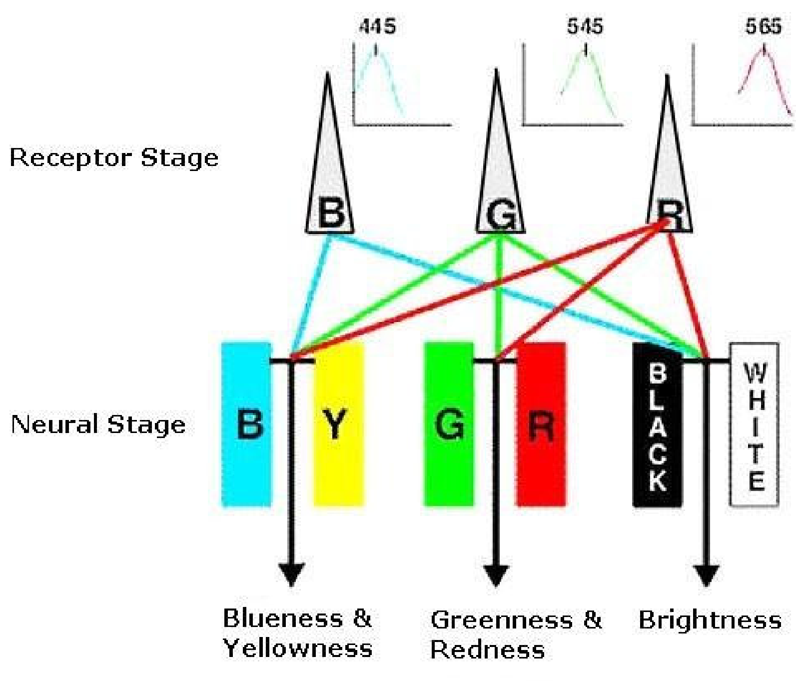


Figure 4: Stage theory of color processing

2.3 Physiology of color vision

The mechanism underlying the visual perception is the absorption of photons, the particles of light, carried out by special, light-sensitive receptive cells on the retina of human eye. There are two types of photoreceptors, rods and cones, named after their characteristic shape. Cones are those responsible for vision at daylight, acuity of vision and for color discrimination. This section is mainly based on [9] and [10].

Unlike many other mammals', the human eye owns three types of cones sensitive to different ranges of wavelength we call red (L), green (M) and blue (S). The red or L-cones (long-wavelength sensitive cones) react to light of wavelength between 500 nm and 700 nm with peak around 564–580 nm, green cones are tuned a little "shorter", to 450–630 nm with peak around 534–545 nm and blue cones to 400–500 nm culminating around 420–440 nm. The response curves of cones are illustrated on the figure below, taken from [11]. Important is that all cones detect wavelength of incoming light only in comparison with at least one other type of cones.

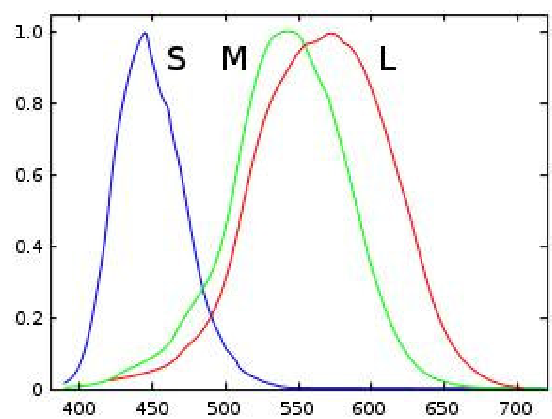


Figure 5: Simplified human cone response curves

Red and green cones are together with the color discrimination responsible for the fine discrimination between bright and dark. On the other hand the short-wavelength sensitive blue cones provide the perception of color contrast and are 10-times shorter in amount than the others. Due to a considerably big interception between the ranges of longer wavelength sensitive cones, the yellowish lights will stimulate them nearly equally, but blue cones will be hardly influenced. In addition, these two types of cones can be also considered red-green and green-red cones. The yellow color, perceived by these two types can be intuitively divided into two parts in the visible spectrum, which we can call the greenish and the reddish yellow. This relationship gives rise to an idea, that even if there were three types of cones, there is a possible explanation for four primaries, because yellow is perceived both by red and green cones opposite to blue, which is only in range of blue cones. The trivariance of color channels allows us to perceive also extra-spectral colors like cyan or magenta.

Color opponent mechanisms are present on the neural stage of color processing as well. After being perceived by photoreceptor in retina the visual information is sent via optic nerve to the thalamus to synapse at the lateral geniculate nucleus (LGN). In the LGN the red-green signal is processed by the parvocellular chromatic channel and the blue-yellow by the koniocellular channel both functioning on opponent principle [12]. The LGN transmits its signals to the first visual area (V1) to process visual signals in cerebral cortex. Opponent processes continue also in V1 where double opponent cells are clustered within localized regions called blobs. Red-green cells compare the relative amounts of red-green in one part of a scene with the amount of red-green in an adjacent part of the scene, responding best to the local color contrast (red next to green). From the V1 blobs, color information is sent to cells in the second visual area, V2. Neurons in V2 then synapse it onto the cells in area V4, which provides input to the inferior temporal lobe ("IT" cortex), which integrates the color information with the information about the shape and form.

2.4 Psychological view

The process of perceiving and distinguishing different colors is, similarly to visual perception mechanisms, based on the comparison of the object with its background. This process is generally based on comparing wavelength and intensity of perceived lights, but there are certain influential factors worth mentioning. The following part will briefly summarize most important of the mentioned properties and effects of cognitive psychological nature, which they cause. All of these effects are apparent changes (or inconsistencies) in the colors perceived. This section is based mostly on [6].

We will use basic color terms of English language, as this study claims in the next chapters, these terms can be universally translated to most of the western civilization languages. Please note that the following pages are not about linguistic categories, but rather about certain color percepts named with exact color terms for the simplicity of the explanation.

First of all, let us describe the physical properties influencing color perception. An important factor influencing the perception in general is size of the object perceived. In case of color, the smaller the object is, the stronger is the convergence of its apparent color – from dark hues to black (e.g. small blue object appearing as black) and from light ones (e.g. yellow) to white. The lightness of a color depends on various physical characteristics, mostly on the luminance of the material, the background and the properties of ambient light. Increase in the brightness of light causes an apparent shift of all colors of the visible spectrum (see section 2.1), those characterized with wavelength below 500nm towards blue and over 500nm towards yellow (e.g. red seems yellowier or more orange-like). This is called the Bezold-Brücke shift. On the other hand in conditions with low illumination the retina becomes more sensitive to shorter wavelengths and less to longer wavelengths. This phenomenon is named Purkinje effect or shift after Czech anatomist Jan Evangelista Purkyně. For example a blue and a red flower, which appear of the same brightness in daylight, will begin to appear unequal in this respect as twilight deepens. The red color will appear darker more quickly than blue, which might appear even brighter than before.

Probably the most influential element of color discrimination principle is the background. Each and every color (similarly to shapes and orientation) is perceived by process of comparison of the focused object with its surroundings. In accordance with the Opponent process theory, the background influences the perception mostly when it is in an opposite-color relationship with the focused object's color. The background can both induce its complementary hue into an object (e.g. if the background is green, the object will appear redder) or reduce apparent saturation of similar hue (e.g. a very red background will induce green into an object), a highly saturated background will de-saturate objects of the same hue and enhance saturation of objects with complementary hue. This principle functions with brightness and is generally called simultaneous contrast.

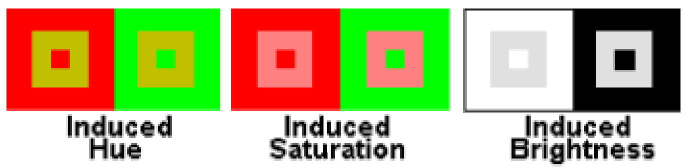


Figure 6: Illustration of simultaneous contrast effects
The figure has been taken from [4]

Assimilation effects, opposite to the first type, cause an apparent change of the color of the background. These are the cases where the foreground instead of producing contrast causes the background to seemingly spread into it. This effect is complete at the point of spatial fusion when the stimuli are no longer viewed as discrete, but fuse into a single stimulus. Spreading, however, occurs at spatial frequencies below those at which fusion occurs. Thus, the stimuli are still observed as distinct from the background, but their colors begin to blend. In the example below white bars spread to make the blue look lighter and the black bars spread to make the same blue appear darker.

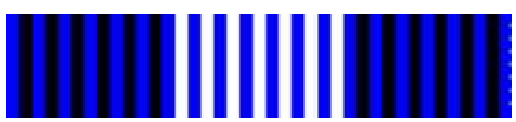


Figure 7: Illustration of assimilation effects
The figure has been taken from [4]

Another principle based on and accounting for the Opponent process theory is the successive brightness contrast. It occurs when the viewer has been under prolonged exposure to light of a particular color. This can produce something like a color afterimage effect resulting in inducement of complementary color. For example, viewing a red field would make a subsequently viewed yellow or white object appear greenish. On the other hand there can be a reduction of apparent saturation. For example, adaptation to a red field would then make a pink object appear whiter. This property can also induce or reduce brightness; for example, viewing a bright field would make a subsequently viewed object appear dimmer[2].

Next group of factors, which influence the human color vision are the properties of perceptual system itself. Together with other perceptual constancies as the constancy of shape, size or distance, color constancy is the perception of an object or its quality as constant under changing conditions. In this case it is the color of perceived object, which remains relatively constant under varying illumination conditions that is caused by the independent changes in responsivity of the three types of cone photoreceptors. This effect is also called chromatic adaptation. Color constancy experiments show that very large spectral changes in illumination cause only small changes in the appearance of objects. There is common consensus that the magnitude of color constancy corrections is very large, but also that the constancy is never perfect [13].

[2] E.g. “rapid light adaptation", the sensation of coming from strong sunlight to a dark room or other way round causes the colors perceived during first few seconds to look lighter and less saturated as there was some kind of white or grey aperture in front of one’s eyes.

2.5 Color spaces

As mentioned before, color can be cognitively defined as perceptual entity with three qualities. A color space can consist of hue, value and saturation or any other reasonable set of quantitative parameters, which will stand for the axis in vector space. An important characteristic is the gamut – the certain complete subset of colors with a color space. The most common examples are color spaces based on popular working color models RGB or CMYK[3], used in CRT displays in first case or printing devices in the latter. Following lines will provide brief introduction to several color spaces designed to depict human color vision. This section is based on [8] and [14].

2.5.1 CIE color spaces

The CIE 1931 XYZ color space was one of the first mathematically defined color spaces, created by the International Commission on Illumination (CIE) in 1931. It is a linear transformation of older CIE RGB space, derived from a series of color perception experiments. It locates color within a tristimulus coordinate system (XYZ are the coordinates) derived from cone responses of retina. A great success is that it resembles perceptual appearance of colored light and the effects of the color mixture of lights by simple addition of vectors. On the other hand it does not correspond well to the perceptual color differences, i.e. the distance (the difference-similarity) measure, between different colors does not correspond directly to human perceptual judgment.

The next generation of color spaces created by CIE are those color-opponent based, derived from CIE 1931 XYZ, from which the most popular is the L*a*b* (or CIELAB) system. Strongly influenced by the Munsell color system (see next section), the intention of CIELAB is to create a space which can be computed via simple formulas from the XYZ space, but is more perceptually uniform[4] than XYZ. Inclining to the opponent process theory and more cognitively defined color spaces (Munsell, HSV or

[3] A color model unlike a color space has no associated mapping function to an absolute color space, in which perceptual difference between colors is directly related to distances between colors and where the interpretations of colors in the space are colorimetrically defined without reference to external factors. In other words it misses a globally-understood system of color interpretation. For example there are several RGB-based color spaces like Adobe RGB and sRGB, created from the model using real colorimetric parameters.

[4] A change of the same amount in a color value should produce a change of about the same visual importance.

HSL[5]) the three coordinates of CIELAB represent lightness of color (L^* = 0 yields black and L^* = 100 indicates diffuse white), its position between red/magenta and green (a^*, negative values indicate green while positive values indicate magenta) and its position between yellow and blue (b^*, negative values indicate blue and positive values indicate yellow).

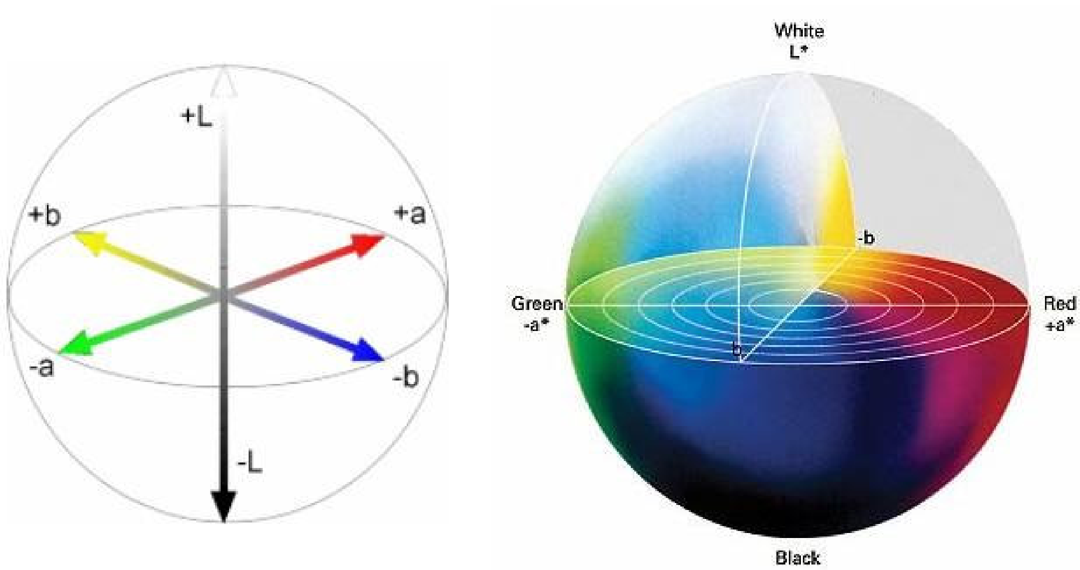


Figure 8: Illustration of the L*a*b* color space

[5] HSV/L coordinates are hue, saturation, value/brightness or lightness/luminance, similar color models based on RGB. In both cases the color hues are arranged on a circle in a spectrum-like manner, while other parameters are in percents. Saturation decreases towards the centre of circle, which is essentially grey and lightness/brightness decreases (or in the second case also increasing) to form a cylinder or a double-cone.

2.5.2 Munsell color system

The goal of perceptual color spaces is to reflect the perception of similarities between colors in the distance between each point of a color space. Munsell color system is one of the best known of these perceptual color spaces. Its coordinates are based on perceptual color properties – hue, value and chroma (approximately corresponding with saturation). The basic idea is that adjacent color samples in each dimension should have a constant perceptual difference, so that the color chips are located at equal perceptual intervals along each dimension. However, no method is given for comparing distances along these dimensions measured in different units.

The theoretical summary from [8] suggests that scaling studies (scaling of color spaces to reach the hypothetic ideal color space) found out an overall pattern which clearly conforms to the Munsell type of organization. However these scaling studies do not describe an axis on which green is opposite to red, but rather red opposite to blue-green and green opposite to red-purple. The results also show the color perimeter divided into five not four equal sections, corresponding to red, yellow, green, blue, and purple. So pure red and pure green, assumed to be opposing colors, are not found to lie at opposite sides of an achromatic point in empirical scaling of perceptual space. This is another factor suggesting that a variant of opponent process theory, involving small changes of hues of the basic primaries is needed. Illustration image below is taken from [15].

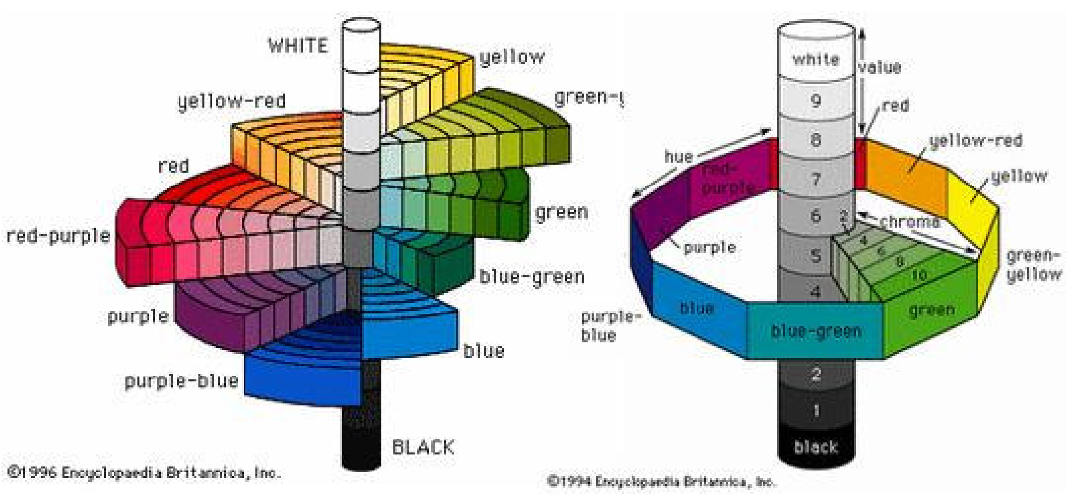


Figure 9: Illustration of Munsell color system

3 Categorization

This chapter is dedicated to categorization, which we consider essential for the human cognition. The categorization as a mechanism is employed in perception, reasoning, planning and action, speech, and any other cognitive tasks. We will describe the classical view on categorization which dominated in philosophy and related sciences for more than two thousand years and is still very influential. However, with recent emergence of cognitive science, new theories arose to overpower the rigid point of view, not only on categorization but on cognition as such. We will briefly describe the classical theory of categorization and thoroughly describe the first view that challenged it, the prototype theory of Eleanor Rosch. Since this view (proposed by Rosch and followed by Lakoff and many others) is the one we consider to be the most coherent and explanatory view, and it is also connected to the latter parts of this thesis, we will not include other influential theories that were constituted later. The two generally accepted examples of such theories are the exemplar and the theory view. The exemplar view claims that there is no single representation of an entire concept; members of a category are formed by specific representations of its instances – exemplars. The theory view (also called the knowledge approach) refers to concepts as to mental theories about the world; membership in a category is decided due to an individual's knowledge [16]. Following text is based on [1] and [17].

3.1 From classic theory to Eleanor Rosch

The classic theory of categorization was established by philosophers of ancient Greece. From that time up to early 18th century, categories were understood as some closed containers filled with things sharing the same properties. Respectively these properties defined the categories. Things were either members of certain category or not (simple Boolean principle). There are other ideas of the "old philosophy" closely related to this notion, for example the idea that reason is a disembodied symbol manipulation, or that mind is separate from body, or that meaning is based on truth and reference (i.e. each word corresponds to one thing in the world).

The first one to notice the flaws in the classical theory was Ludwig Wittgenstein. In his later work he pointed out that there are categories that have no clear-cut boundaries, constituted of members that do not necessarily share the same properties, and that members of the categories might be central and non-central. His most famous example was the word *game*. There are various types of games for example, children games, table games, or sports that share no properties at all, some are based simply on the chance, some have rules, some do not, there are even games that does not have a winner. The only thing they have in common is that they are grouped in the same category. For this principle Wittgenstein found a very nice parallel – the *family resemblances*. The members of a family resemble one another in various ways, but usually there is not a single feature they all share. This principle also accounts against the closed boundaries. There is always a possibility to gain a new family member (marriage, birth, etc.), without it, the world would not function. Another fine example is the category *number*. For a long time, since ancient Greeks, numbers have been only integers, but then rational numbers came along with the need to represent fractions, then real, complex, and transfinite numbers were discovered. This example also shows the centrality and non-centrality of category members. Every precise definition of number must include the integers, but not every definition must include transfinite numbers.

Another important step towards a new definition was the theory of fuzzy sets proposed by Zadeh. A fuzzy set is characteristic with its member function, which allows each member not just to belong or not belong to the category (0 or 1), but to be a part of it to some extend (a number between 0 and 1). This concept provides categories with fuzzy boundaries and accounts for the centrality character of categories, a central member for example may belong to the category at 100%, but non-central only at 20%. This theory is also particularly important for chapter 5 of this thesis.

The first one to propose a new general theory of categorization was Eleanor Rosch (in early 1970's). She saw categorization itself as the one of the most important issues in the cognition. Her main proposal was that thought in general is organized in terms of prototypes and basic level structures. To demonstrate her statements she established new research paradigms in cognitive psychology and proved them on experimental basis.

3.2 The prototype theory of categorization

This theory was also called “the theory of prototypes and basic-level categories”. Note that since Rosch is one of the first cognitive linguists she speaks mostly about categorization in language, with direct consequences on categorization in general. We are not aware of any reasonable objections against this generalization. In her chapter in [17] Rosch describes two principle of the formation of categories[6]:

Cognitive economy: categories are built in a way that maximizes information gain, but also preserves an adequate level of simplicity. The goal is to differentiate which information is important and which is not.

Perceived world structure: since the world is not unstructured total set of equiprobable co-occurring attributes (as the classical theory expected), there are attributes occurring together often and those that are never connected. The perception of such world is always driven towards a high correlation, so people will not think about (consider) features, which are not generally known to be interrelated or that have no logical connection.

The prototype theory, also called “the theory of prototypes and basic-level categories”, proposes that categories and the process of categorization have following properties or characteristics.

1. **Family resemblances** (Wittgenstein): members of a category may be related to one another without all members having any properties in common (i.e. a category does not have to have defining features).

2. **The basic level** (Rosch): the categories are not merely organized in a hierarchy from the most general to the most specific, but are also organized so that categories that are cognitively basic are in the middle of a general-to-specific hierarchy. These basic categories are simple and short words, most commonly used labels and most neutral terms for category members, first named and understood by children, and first to enter the lexicon of a language. Considering

[6] Under the formation of categories she means their formation in the culture, not the development of categories in children born into a culture. She also does not intent to constitute a model of how categories are processed (how categorizations are made) in the minds of adult speakers of a language.

for example a hierarchy *mammal – cat – Siamese*. When talking about an animal sitting on the window it would sound strange saying: "*The mammal sat on the window.*"

Prototypicality (Rosch): there are members of a category that are more "typical" then some others and there are methods for deciding this typicality. The more prototypical of a category a member is, the more attributes it has in common with other members of the category and the fewer with members of contrasting categories.

A very common notion is that according the prototype theory the membership function is a comparison of the category prototype with the considered samples. Rosch herself is against any tendencies trying to present prototypes as a processing model for categories, a theory of representation of categories (prototypes as representations of categories in mind), or the theory of learning categories. The prototypes or more precisely the judgments of degree of prototypicality are in these terms more of a property, than the base of categorization. However, some studies claim that in certain cases categories are built around prototypes – focal colors (see the next section). The comparison with the prototype is also a useful mechanism for artificial intelligence based categorization frameworks. The distinguishing criteria we used for the color categorization simulation described in chapter 6 are working on this basis of such comparison: each distinguishing criterion stores a prototype and reacts to an input with the activity proportional to an exponentially decaying function of the squared distance between the input and the stored prototype. However, the distance function takes into account statistical characteristics of the sample set, such as variances of attributes and their mutual covariances (for details, see [29]). Unfortunately there are shortcomings of this approach, e.g. that categories (or categorization mechanisms) based on prototypes are not able to encompass the composite categories, i.e. those that have multiple different prototypes (best examples). This phenomenon will be described in more detail in the next chapter.

3.3 Basic color categories

According the color categorization definition in [1], there are two basic processes constituting color categorization. The lexical color categorization, characterized as the division of color sensations into classes corresponding to the significata of the color words of a particular language, and the perceptual color categorization, the division of the color sensations into classes by the perceptual processes of an (any) organism. There are various studies aimed to find a correlation between these two phenomena.

In 1969 Berlin and Kay [18] suggested that there exists a final set of basic color terms which describe all percepts in any language. This thesis will be discussed in more detail in the next chapter. Here we would like to concentrate on the definition of the basic color terms. A basic color term should satisfy following criteria:

1. it is **monolexemic,** i.e. its meaning is not predictable from the meanings of its parts (e.g. *blue* vs. *greenish blue*)
2. **generality** criterion
    a. the color referred to is not contained within another color category (e.g. *scarlet* is contained within *red*)
    b. it is not restricted to a narrow class of objects (e.g. *blond* is used only with hair, complexion, and furniture)
3. it is **psychologically salient**, common and generally known (e.g. *yellow* vs. *saffron*)

In other words, the basic color terms can be considered the smallest set of simple words with which the speaker can name any color. Basic color terms name basic color categories. This assumption is in accordance with the basic-level property from the prototype theory of categorization. Berlin and Kay claimed that each and every language has at least two maximum eleven basic color terms. The base for their proposal was a survey in which they studied not only simple color naming based categorization, but also the best examples of categories. They realized that the best examples for color terms used in each examined language cluster together, including the best examples from English language, so they assumed that color categories are built around this prototypes, and that additionally these prototypes are universal for all languages.

4 The World Color Survey

In this chapter we introduce the main hypotheses about basic color terms proposed originally in 1969, and the World Color Survey aimed to asses these hypotheses against a broader empirical basis. Later on we describe an experiment reproducing the WCS experiment in Slovak language with a short overview on its conclusions with emphasis on shortcomings of technical details of the original method that we have encountered while performing the experiment. At last, we provide an overview of selected studies regarding WCS data aimed to verify the original hypotheses.

4.1 The original study of Berlin and Kay

In *Basic Color Terms: Their Universality and Evolution* [18] from 1969 Berlin and Kay proposed two general hypotheses regarding basic color terms and how they emerge in a language:

1. There is a restricted universal inventory of basic color categories
2. A language adds these terms in a constrained order, interpreted as an evolutionary sequence.

The close set of basic color categories consist approximately of terms corresponding to English black, white, red, green, yellow, blue, purple, pink, brown, orange and grey. The evolution of basic color terms starts with the distinction between black and white or more precisely between dark and light and it is in correspondence with the Opponent process theory of color vision (described in section 2.2). Later research of Eleanor Rosch showed that these first two terms represent rather categories covering warm (e.g. yellow, orange and red plus white) and cool colors (black, green, blue) called composite categories, characteristic with multiple best examples (e.g. warm colors are both red and yellow with different prototypes) [19].

Berlin and Kay proposed seven evolutionary stages characteristic by the color categories present in a language of that stage. First stage was the black-white or dark-light distinction, in the second stage red emerged. Stage III contained a term for either green or yellow, stage IV both green and yellow; stage V added blue and stage VI brown. The last stage added the remaining of the eleven basic categories (purple, pink, orange, and grey). To this stage belonged also English and similar western languages.

Although the hypotheses of Berlin and Kay have been substantially confirmed by their research, it had certain drawbacks weakening their findings [20]. First of the important methodological objections against their empirical generalizations was, that the amount of twenty languages studied experimentally is not sufficiently numerous to justify universal conclusions. Moreover the number of speakers per language did not exceed three, which leaves us with roughly 40-60 speakers interviewed. Secondly, the data were obtained in Berkeley rather than in native communities, not all of the subjects were skilled speakers of the languages studied, and all of them spoke English as well as their native (inherited) language.

4.2 The course and methodology of the WCS

The WCS has begun in 1976. Its major purpose was to validate, invalidate or modify the main findings of Berlin and Kay on a broader empirical basis. Data on the basic color term systems of 110 unwritten languages were gathered with mean of 24 speakers per language. Once data gathering was completed (circa in 1980), data processing, quality control, and analysis were undertaken. The online data archive was published in 2003 [21]. The following text is based on [20] and [22]].

4.2.1 Stimulus materials

Stimulus materials used in both the WCS and Berlin and Kay's experiment consisted of 330 color samples from Munsell color system. There were individual chips in glass 35 mm slide and full stimulus board with all samples on it. The stimulus board, in form of a grid (from now Munsell color grid) encompasses 320 samples, which represent forty equally spaced Munsell hues from R2.5 in column 1 to RP10 in column 40, at eight levels of lightness (Munsell value) in rows arranged from the lightest on top to the darkest on bottom. The color in each cell corresponds to the maximum available Munsell Chroma (saturation) for that hue-value combination. In addition a left-most column displays ten levels of lightness of neutral shade (black, grey, white). The next image illustrates the stimulus board and is taken from [22].

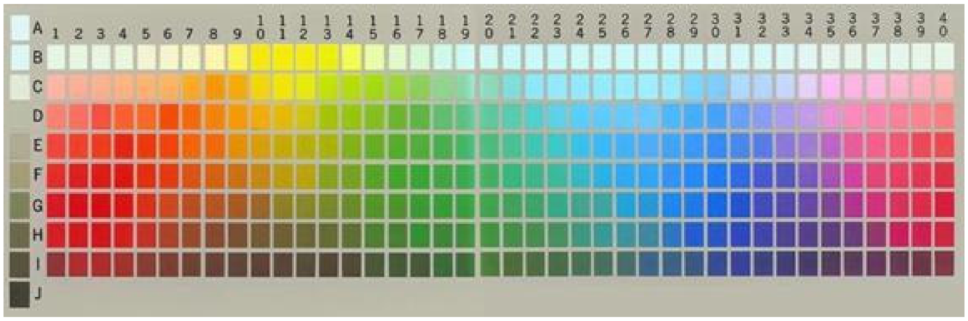


Figure 10: Illustration of the stimulus board from the WCS experiment

4.2.2 The Method

Unlike the original experiment of Berlin and Kay, the WCS method consisted of only two tasks. In the original version of the experiment by Berlin and Kay speakers were initially asked to elicit all the basic color names of their language. In the WCS experiment the task of collecting the set of used terms was carried out by the experimenter. The concept of basic color terms was explained to cooperating speakers as "*the smallest set of simple words with which the speaker can name any color*". Instructed to respond with short, simplest names, observers were shown the 330 samples, one by another in a fixed random order. This was called the naming task.

In the latter task, called foci task, subjects were to choose so called focal color chips – the samples which they considered to be the best examples of color terms they used, separately for each term, on the full stimulus board (described above). Considerable factor influencing results of this task was that the speakers were allowed to choose as many focal samples as they wanted. Considering that they had a possibility to select all samples they previously labeled as forming one category (e.g. ten or twenty samples), the misunderstanding of the task might cause a full failure in detecting the best examples. Additionally there was no backward checking whether the samples selected as best examples of a category were previously classified as members of that particular category during the naming task.

4.3 Slovak color experiment

In spring 2008 we carried out an imitation of the World Color Survey experiment. We gathered data from 25 speakers, 12 men and 13 women, whose mother tongue was Slovak and who had no color vision deficiencies. With slight modification of the method we gained data of two types, first consisting only of 11 basic categories proposed by Berlin and Kay and second including some other color terms considered as significant because mentioned very often. Results of this experiment on one hand confirm the original thesis, on the other hand show that the evolution of basic color terms may continue.

4.3.1 The method

We used nearly the same methodology as in the WCS experiment with several changes. At first, the stimulus material, unlike the original study, consisted of color samples[7] printed on ordinary office paper with an all-purpose office laser printer. Although we used creation of the samples the original L*a*b* color values and a professional software, the quality must have been remarkably different from the original material.

At the beginning of the color naming task speakers were instructed to use the shortest and most basic names. However the discussion about generality of almost every newly added color name was necessary unlike we predicted. People frequently considered the names of their favorite colors, or those color terms they personally used often, as generally known and used. However, after further explanation of the task, most of them agreed upon using more general terms.

The ideological modification of the original method consisted of a two-answer questioning mechanism. Speakers were urged not only to use general terms, but also, when they used a word outside Berlin and Kay's 11 basic color terms (e.g. turquoise), they had to name also a counterpart for it from this closed set (e.g. turquoise-green or turquoise-blue). In this case we recorded both the first answer and the second specifying answer, so the data from this experiment finally consisted from two answer sets for each speaker.

[7] Rectangles app. 4x3cm glued to app. 5x4cm white thick paper frame.

Since the English language is one of the most popular and nowadays obligatory in compulsory education it would have been hard to get to monolingual speakers. Therefore we did not count this aspect as significant. Probably all of our subjects speak English, some of them German or French as well.

The last important factor influencing the results of this experiment is that speakers were examined not one by one, but in small groups. They were instructed that there are no bad answers and that it is important, that they would speak for themselves, but still there reamains a great chance that they influenced one another.

4.3.2 Results and important findings

As a result of this experiment, we found out that besides the 11 basic color categories there were several color terms that appeared often. The most frequent of these were: *tyrkysová, kaki, béžová, bordová*, and *okrová*, which roughly correspond to English: turquoise, khaki, beige, burgundy (or maroon) and ocherous[8]. We will call them *outer* color categories.

What our results certainly confirm is that Slovak language contains firmly established 11 basic color categories as proposed by Berlin and Kay. To this also adds a fact that in foci task speakers had some difficulties with labeling the outer color categories, but not with selecting best examples for the basic ones. Only one outer category, turquoise, was strong (i.e. frequent) enough to win in a few color samples, a more accurate description of the state of being is provided in the visualization.

The figures below are outputs of the visualization method described in chapter 5. In short the map on top of the first figure displays the results of Slovak color naming and foci tasks in *classic* way, the small number in the left corner is the identification number of the category, which was elicited by most of the speakers. The larger numbers indicate if the color chip was selected as the best example of the category with the displayed serial number. The second map is so called *fuzzy* visualization, based on proportional mixtures of colors elicited categories derived from the best example colors, for more

[8] Note that as there exists no one-to-one translation between any languages, there is a certain problem translating color names from one to another. Be aware that these translations are only informative. On the other hand, it is alarming that the 11 basic color terms are easily translatable between various languages, try to think about it and imagine it for example between any two languages you know.

detail see chapter 5. The latter figure displays the same visualization for the results without outer categories, but only in the fuzzy stile.

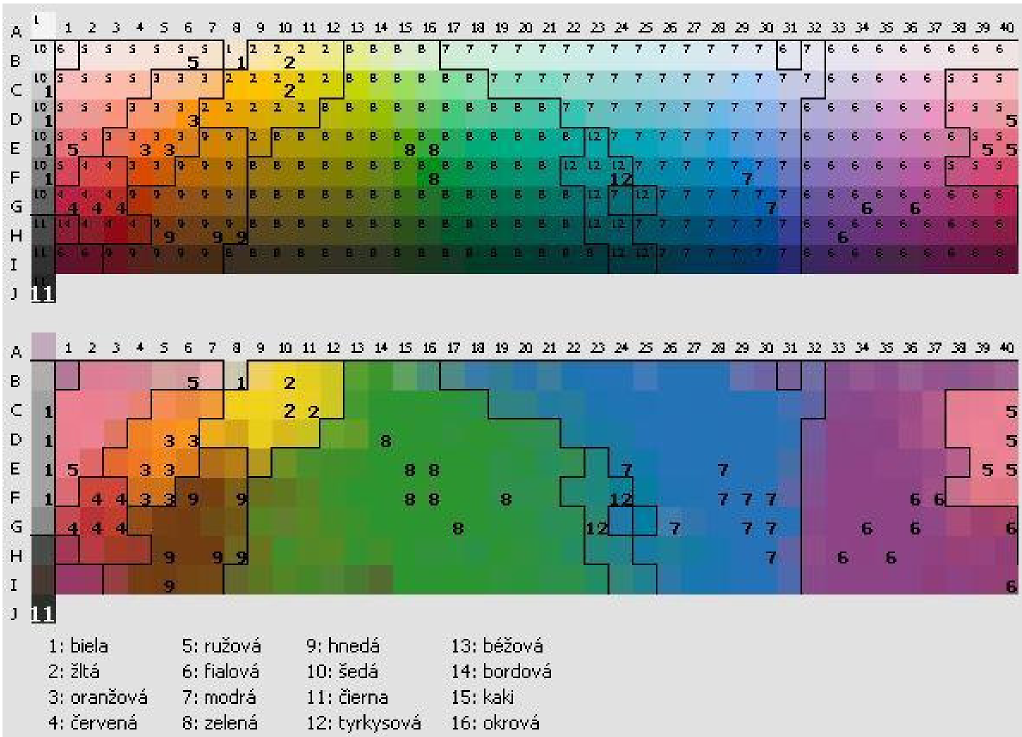


Figure 11: Visualization of the Slovak color experiment with outer categories
This is the visualization of the results including the so called outer categories (the last 5 from the list). Note that only category number 12 was able to gain some space on the grid among the 11 basic color categories. For better understanding of the visualization and evaluation of the data from color naming experiment see chapter 5.

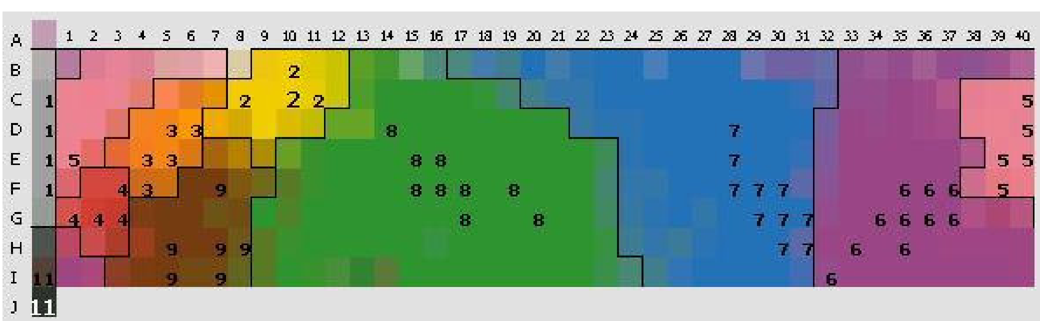


Figure 12: Fuzzy visualization of the Slovak color experiment without outer categories

The greatest drawback of this experiment is that it was not executed on exactly the same stimulus material as the original WCS experiment. During the foci task we encountered that the subjects complained that the color samples missed particularly representative hues of some colors, for example of red[9]. This might by caused by the print quality, or more precisely the compression from Lab color space to printer color space, but it could also be a property of the original color values.

We also noted some disadvantages of the method in general. These findings can be useful for a general overview on the WCS experiment as such. First of all, the 330 sample set is considerably large amount of presenting material. Most of the subjects got bored and tired after only one third of the samples. A significant portion of the samples presented are very similar color hues. Considering the ten degrees of lightness for each hue, it might cause some samples to look nearly the same. This is on the other hand not a general rule for all hues and all degrees of lightness.

As mention before, as a great disadvantage we see the option of choosing unlimited amount of focal colors. We examined that some speakers, when not forced, pick a large number of samples just to "say something". The other drawback of this part of the method of the experiment was that when not corrected by the experimenter subjects picked up the samples previously named as one category, to be the best examples of another category, mostly in case of so called outer categories. This fact accounts for the original thesis about 11 basic colors, but could do better when a comprehension of the examined speaker was present.

The most important fact we learned about this and the WCS experiment is that the speakers must be aware of the task that they are to produce (name) the smallest set of simple words with which they can name any color and agree upon this color naming with all speakers of their language. Here we would like to emphasize that subject should be made aware of the task during the whole course of the experiment and to be urged to keep focused only on the task.[10]

[9] A large number of the subject considered an ideal red a color hue, which can be in RGB color space coded as (255,0,0). To this hue there is no sample in the WCS sample set.
[10] Unfortunately, people generally expect this experiment to be about presenting their opinions and theories about colors or color naming in general, e.g. "The basic color terms are red, green and blue because the computer encodes color this way".

4.4 Further studies based on the WCS data

There are several studies concerning the WCS data aimed to prove to some extend the original hypothesis of Berlin and Kay. They are based either on analysis of clustering of color terms, their distribution and saliency, but also on artificial intelligence simulation or on the human color vision abstraction into the perceptual color spaces. The oldest way proposed to analyze the WCS languages was through a special conceptual framework, and on the basis of preliminary data summary and visualization.

4.4.1 Preliminary processing: the first visualization

Shortly after the WCS data were gathered, data processing, quality control, and analysis were undertaken using computer programs for both data entry and data analysis. The preliminary data summary, presented in [22]] and [20] included among the others the first visualization of the data, created using simple ASCII symbols. First, the "naming arrays" displayed the responses from the experiment from both naming and foci task, for each single speaker and for the whole language. These were aggregated results of the naming task across all speakers at various levels of inter-speaker agreement. The modal agreement array (100%) displays for each stimulus chip a symbol corresponding to the term most often applied to that chip, regardless of how often that was. The 30% Agreement array displays for each stimulus chip the symbol corresponding to the term most often applied to that chip only if that term was used for that chip by at least 30%, similarly for other levels of agreement (i.e. 70%, 40%, etc.). The "term maps" are created for each term separately. In the map for a given term, each chip receives a typographical symbol (including blank) of visual 'density' showing the frequency with which speakers named a particular chip with displayed term. High-agreement symbols tend to occur in the interior of categories and lower agreement symbols at the edges.

A positive feature of this type of processing is that the terms are imaginable on the WCS color grid, but only for those who already can imagine the stimuli from the experiment. Additionally, there is no possibility to display more term maps in one image, or to display the other categories elicited for a single sample than the winning one (the most frequently used). The visualization method we propose in the next chapter can handle all of these issues, because it uses colored maps instead of characters, and those are able to display more information.

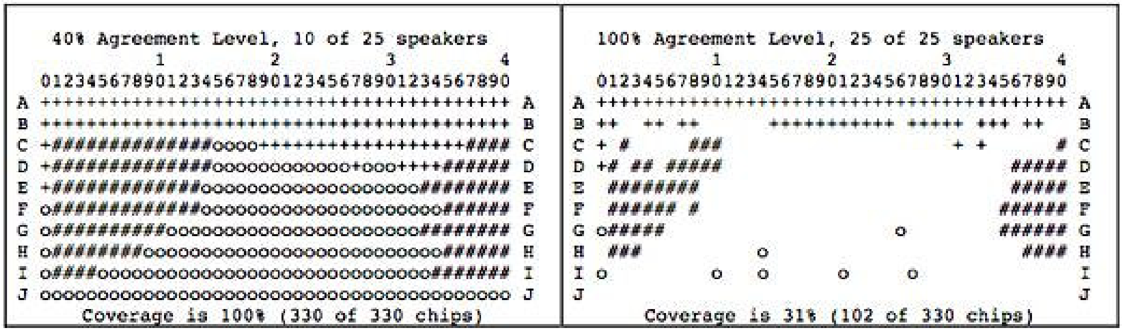


Figure 13: Aggregate naming arrays for 25 Wobé speakers
Symbols **+, #** and **o** represent color categories. Note that at the 40% level of agreement all 330 chips were named, that is, at least ten speakers gave the modal response for each of the 330 chips, we can consider Wobé a high consensus language [22].

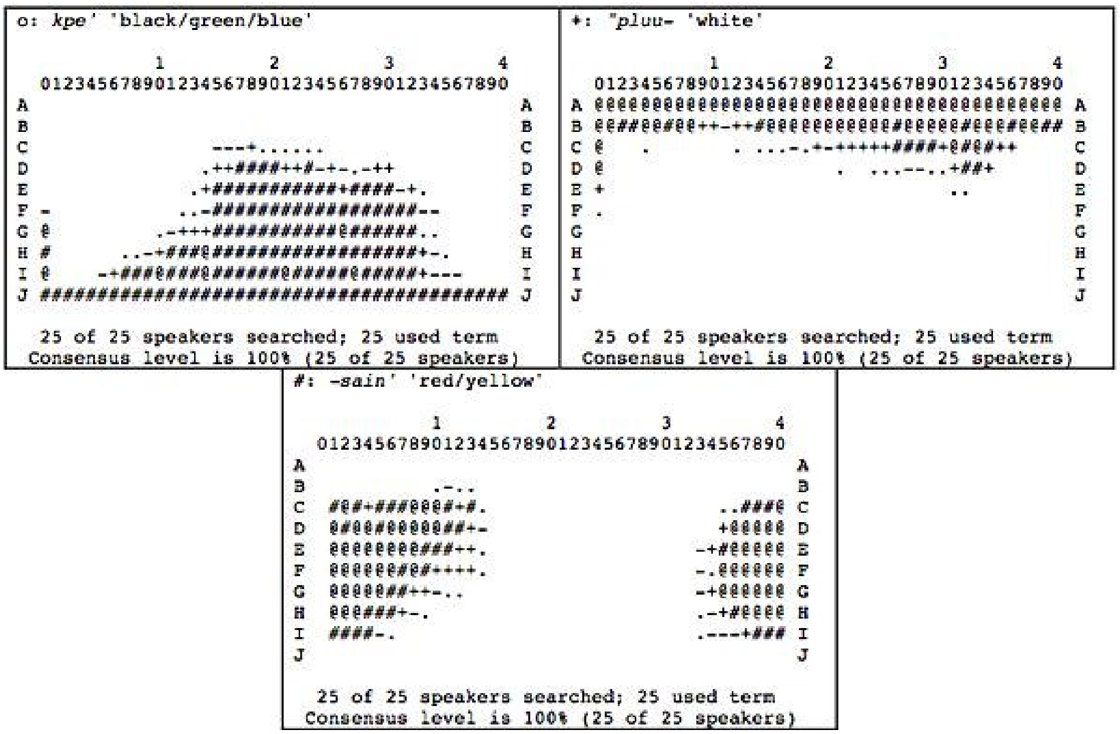


Figure 14: Term maps for the three terms of Wobé
Clue: (@ = perfect agreement/100%, # = 81-100% agreement, + = 61-80 % agreement, - = 41-60% agreement, . = 21-40% agreement). Figure taken from [22].

4.4.2 Conceptual framework of evolutionary stages

The conceptual framework proposed in [21] divides the WCS languages into groups based on the number of color terms they use. Note, that this significant number of terms is always derived from color naming task winners, i.e. most frequent names collected from all fields[11].

This concept suggests that an evolutionary development of basic color terms should not be seen as a single process, but as two partially independent processes: the division of composite categories into the six fundamentals (black, white, red, green, yellow, and blue) and later the combination of fundamental categories into derived categories (the other 5 basic color terms). The first progress finally yielding six primaries is categorized into five stages, corresponding with systems containing two to six composite or fundamental categories. Beginning with two basic composite categories, there follow two partially independent processes: dissolution of the white/warm channel and dissolution of the black/cool channel.

Through the two-process mechanism the authors explain the fact that a large set of languages developed separate terms for white, red and yellow (or similar distribution of colors covered by warm category), but did not developed separated terms for green and blue, or even the whole cool category. However, this phenomenon still accounts against the originally proposed opponent process theory basis, stating that the next step of the progression will be the opponent color (e.g., when a language already has black, white, red and green, the next stage will be yellow and blue so no primary will be left without its opposite).

In summary, the stage I. consists of two categories – warm and cool, in stage II there happens the separation of warm channel to white and red-yellow in all cases. Differentiation starts in stage III where either decomposition of cool channel or of the red-yellow category occurs. The latter decomposition is influenced by the previous distribution. There are three possible pairs of colors remaining together in this stage,

[11] We mean a set of terms, from which each term appeared at least once as a most frequent answer in color naming task. Please note that this winner-take-all-like method, despite its simplicity and power to generalize, has its real drawbacks, mostly in case of colors, that are not typical members of a category. It is possible that from 25 speakers in 13 cases the winner is category *a* and in 12 category *b*, so *a* is a winner, but a weak one.

green-blue, black-blue and yellow-green. In the last stage all six primaries are separate. The table (figure) below describes this process in more detail.

		[W R/Y G/Bu Bk III.G/Bu] w2→	[W R Y G/Bu Bk IV.G/Bu] c2↓	
[W/R/Y Bk/G/Bu] w1→	[W R/Y Bk/G/Bu] c1↑ c1→ w2↓	[W R/Y G Bk/Bu III.Bk/Bu] w2↓		[W R Y G Bu Bk]
		[W R Y Bk/G/Bu III.Bk/G/Bu] c1↑ c1→	[W R Y G Bk/Bu IV.Bk/Bu] c2↑	
		[W R Y/G/Bu Bk III.Y/G/Bu] →	[W R Y/G Bu Bk IV.Y/G] ↑	
		[W R Y/G Bk/Bu III.Y/G] ↑		
I	II	III	IV	V

Figure 15: Table of stages of color naming systems

4.4.3 Color naming universality

This part will briefly describe selected studies of WCS data regarding the existence of universal tendencies in color naming with various means. In [23] the authors examined whether color terms from different languages in the WCS do cluster together in color space to a degree greater than chance and whether they all (unwritten languages of non-industrialized societies) fall near color terms of written languages from industrialized societies as examined in the Berlin and Kay's original survey (e.g. English). In short, they analyzed whether categories of WCS languages represented in L*a*b* color space[12] were more clustered across languages than would be expected by chance.

[12] They used L*a*b* for its good resemblance of perceptual differences between two colors (see section 3.5)

Similarly they compared the measure of separation between WCS languages and B&K[13] languages with WCS languages and hypothetical data. In both cases they discovered relationships closer than chance.

Another study [24] provides analysis of WCS data by using k-means cluster and concordance analyses. The results show that when divided into 2 to 10 clusters, the average color-naming patterns of the clusters all gloss easily to single or composite English patterns, and also that the structures of these k-means clusters unfold in a hierarchical way similarly to evolutionary scheme described above. Processing of the WCS data also showed that 8 was the optimal number of WCS chromatic categories: red, green, yellow-or-orange, blue, purple, brown, pink and grue[14] (green-blue), what is also roughly consistent with the framework proposed above. In the second part the analysis of concordance in color naming within WCS languages revealed statistically significantly high concordance across languages in small regions in color space that agreed well with five of six primary focal colors of English.

A different approach to finding universalities in color naming can be found in [25]. In this work authors introduced a computational model of the acquisition of color categories with an aim to prove that the universal character can be explained as the result of learning on the basis of linguistic communication. The model was in form of a multi-agent simulation, in which agents perceived colors in L*a*b* color space and interacted in two ways, in discrimination (training color categorization) and guessing game (communicating categories to others). The preliminary results of this study were compared with results from the WCS and showed that agents were able to acquire color categories, and not only discriminate them but also communicate them well. Interestingly, the categories resulting from the simulations took up regions in color space that corresponded well to the WCS data.

This surprising phenomenon is understandable, when explained using the properties of perceptual color spaces as suggested in [8]. This study proposes that a possible explanation for color naming universalities is that the developmental order of color names is due to the irregular shape of the color space. They illustrated the irregularity of

[13] Languages from the original Berlin and Kay's study

[14] Unfortunatelly this term refers not only to a composite green and blue category, but also stands for a color which is a mixture of green and blue. The latter case is severe, but still present in some WCS languages.

the perceptual color space (described in section 5.3) pointing at the shape of Munsell color system (see Figure 15 below), where hue interacts with saturation and lightness producing large bumps. These are located at focal yellow and focal red. The entire blue-green area is depressed (of low chroma), as is the area below focal yellow. The authors also assume that a property of names assigned to the color space at any stage is that they have to be most informative about color. In case of two color terms, the most informative categories will be dark/cool versus light/warm. In correspondence with framework described in 4.4.2, the region of color space that is most distant from the regions specified by these two terms is red. Further, after three terms specified it becomes more difficult to determine which is the next most distant region, because the differences in distances are smaller and depend in part on how the focal areas are determined. Expected is either yellow or blue to be the next, followed by green, purple, pink, orange, brown, and grey.

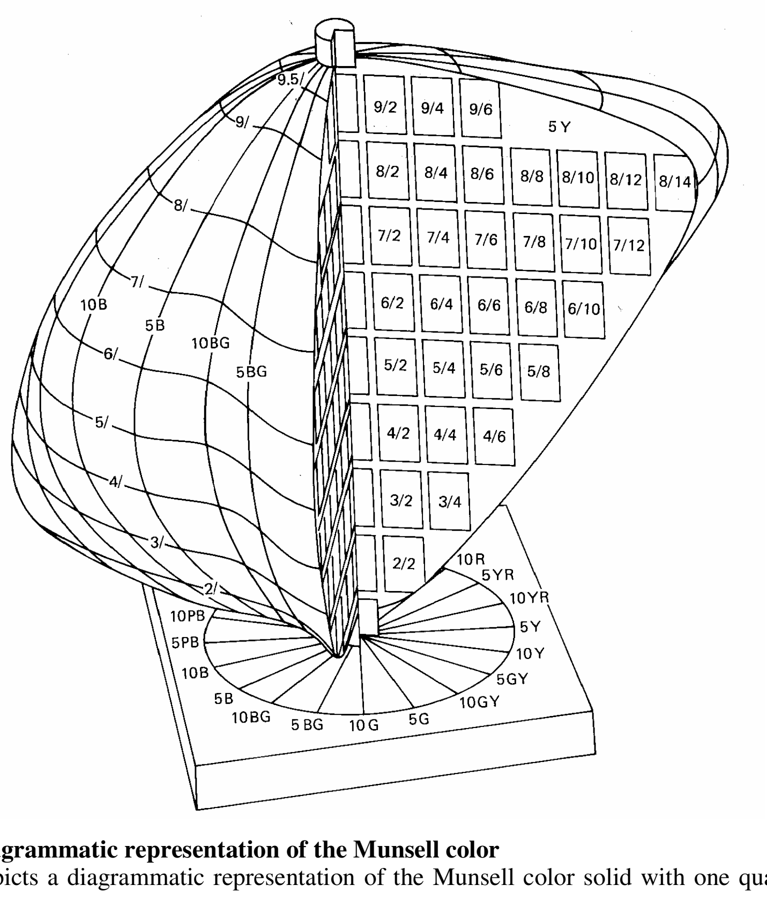


Figure 16: Diagrammatic representation of the Munsell color
This figure depicts a diagrammatic representation of the Munsell color solid with one quarter removed. The numbers displayed in boxes represent individual color samples – with various hue, value and chroma. The circle on the bottom displays various hues, the value (lightness) increases from bottom (black) to top white and the chroma (saturation) increases from center to the side. Figure taken from [26].

5 World Color Survey Visualization

In this chapter we would like to present our visualization of WCS data. Inspired by the old visualization based on typographical symbols described in section 4.4.1 the ambition of our conception is to overcome the limitations of the old one by employing more dimensions to display more information and to facilitate the imagination of the original data stimulus material. We project each from the WCS languages onto the Munsell color grid in three various complementary ways called classic, reliability and fuzzy visualization. In this chapter we will at first define some specific terms that will be used in the whole chapter, describe the three types of visualization, and the program for generating these visualizations in form of .jpg images. We will present an illustration of a confrontation study of selected outputs with the conceptual framework from section 4.4.2 and finally discuss the possible usage of this visualization, its advantages and drawbacks.

5.1 Visualization types and basic terminology

5.1.1 Terminology

Firstly it is important to note that we will interchange freely the terms *color term* and *color category*, despite the fact that they do not mean the same thing. Generally, the color term is the name of the color category. We are aware of the distinction between them, but it makes in no semantic difference in this case.

Secondly, a category will be most of the time described only by its serial (identification) number, because the lexical name in the case of primitive languages plays no role for us. However the program is able to create visualization with a clue of the mapping between serial numbers and concrete terms of a language. For simplicity we will shorten the *serial number of the category* into just *category number* or sometimes just *category*. Additionally the color *sample* is the same as the color *chip* and refers to one field of the Munsell color grid.

A *winning category* for a certain color sample stands for the color term of particular language that was elicited for certain color sample at most. Similarly to winner-take-all algorithm in this case the winner may be one of several frequently used terms with advance of small percentage, so it might not be exactly a dominating color term of that color sample. Still, this method has already been used in the first visualization of the WCS data described in 4.4.1 and has no equivalently simple and strong counterpart. A *focal chip* is the color sample selected as the best example of a certain category. Respectively a *focal category* for a chip stands for the serial number of the particular category of which the chip was selected as focal.

By *reliability* we mean the percentage of speakers that contributed to the winning category, i.e. how many speakers of all elicited the winning color term for a certain chip. Average reliability is the average of reliability values from all 330 color samples. The *Munsell color grid*, as mention before, stands for the full stimulus board described in section 4.2.1.

5.1.2 Classic visualization

This type of visualization is the simplest one projecting the winning category numbers on the Munsell color grid with original colors used in the color naming experiment. The black lines around color chips represent the borderlines between categories. A small number in the left-top corner of each field represents the winning category number. The larger number in the right-bottom corner of some chips stands for the focal category, i.e. labels the best example chip with the number of category for which it was considered the best example. The size of the font of this number represents the agreement of speakers. We include this focal category numbers only in fields with at least 10% agreement of speakers.

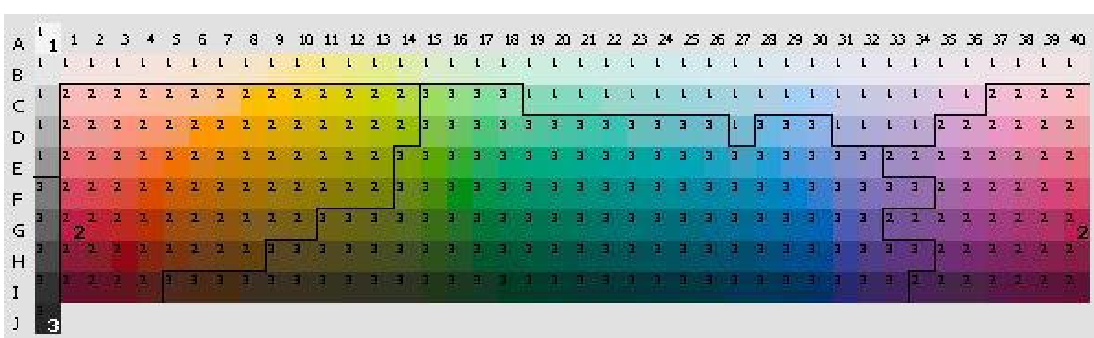


Figure 17: The visualization map for Wobé
Note that this language was also displayed on the examples in section 4.4.1.

5.1.3 Reliability visualization

In this type of visualization the color of the chips of color grid depends on how many speakers contributed to the winning category of that chip. The weaker agreement, the lower is the saturation of a chip. The new color of each sample is counted as the proportional mixture of its original color and grey color from the first column in the same row, to preserve the lightness of the sample. For example when only 50% of speakers agreed upon calling a chip *c* with the winning term *t*, the color of this chip will be half the original color and half grey. One drawback of this visualization principle is that there is no possibility to desaturate the shades of grey, so the rightmost column does not show the reliability of its category.

Two examples below display Wobé with the highest average reliability from all WCS languages (89%) and the language with the lowest average reliability Tifal (42%). Note that in this case we omitted the numbers of winning categories, just for simplicity; we kept the focal choices and borderlines. In the case without category numbers, but focal numbers present, we still include category numbers in separated chips, i.e. those which are not surrounded by any samples of the same category.

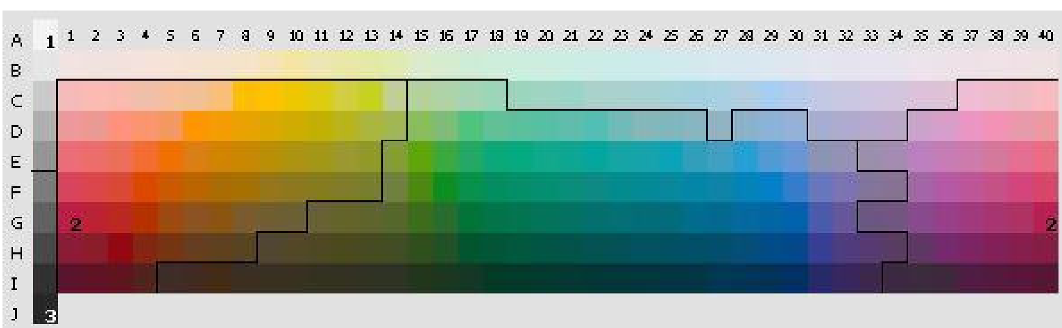


Figure 18: Reliability visualization for Wobé

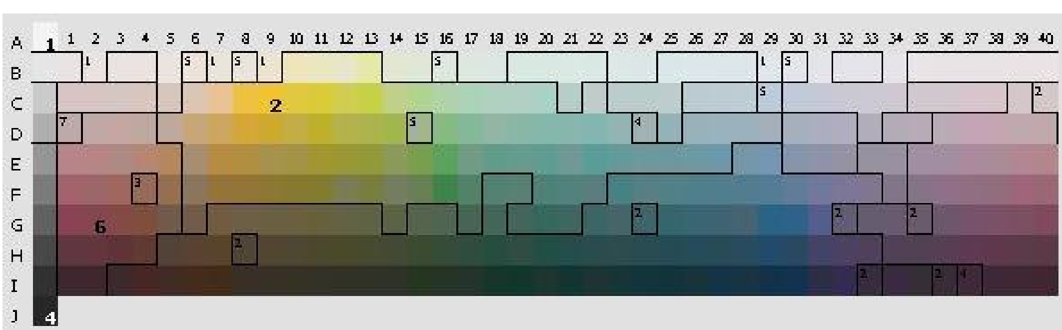


Figure 19: Reliability visualization for Tifal

5.1.4 Fuzzy visualization

The most interesting and innovative is the fuzzy visualization. In this case the color chips of the grid no longer represent the original samples, but visually encode the results of color naming task and in certain sense also the foci task. Color for each chip is computed as a weighted mean of elicited categories represented by the average focal color, i.e. the mean of colors of all focal chips. In symbolic representation: $c = \sum_{i=1}^{N} r_i f_i$, where c is the resulting color[15], N is the number of color terms in language, r_i represents the portion of speakers responding with the category i and f_i stands for the averaged focal color counted for the category i globally from all samples labeled as best examples of it, in a similar proportional way.

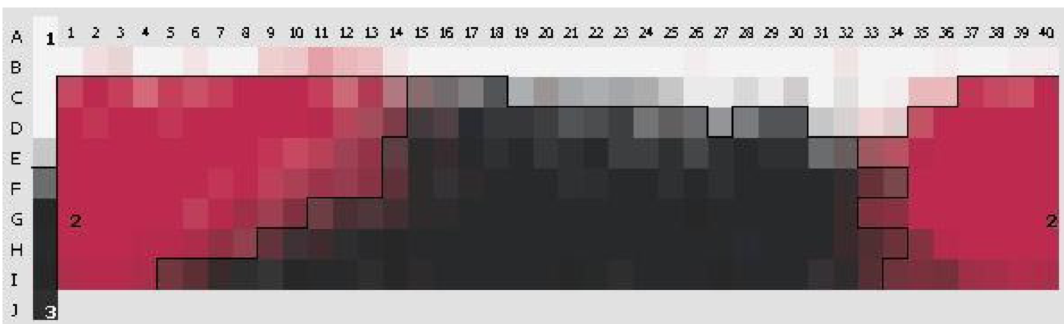


Figure 20: Fuzzy visualization for Wobé
For continuity we present also the fuzzy visualization for Wobé. Note an interesting phenomenon – the focal colors are consistent, but consist only of 3 primaries, even if the winning term area for these categories, especially number 3 (dark/cool color) covers also blue, green and other hues. This accounts for the stage theory of color term systems.

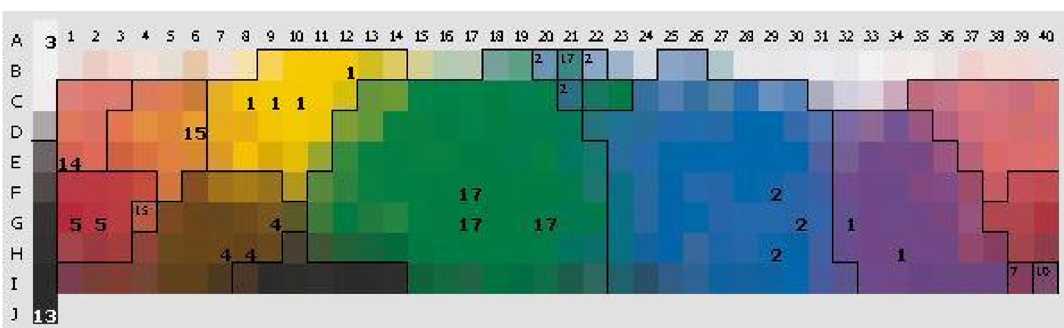


Figure 21: Fuzzy visualization for Chavacano
This primitive language from Philippines with relatively high average reliability (73%) demonstrates a fully developed language with 11 winning categories similar, if not identical to 11 basic color terms proposed by B&K. Note that distribution of these categories is very similar to those from Slovak language (chap. 4.3) and probably with other western languages as well.

---

[15] This is the abstraction independent of color space in which the color is encoded. The parameter c can either represent a triple of RGB or L*a*b values, or whichever else coordinates.

The interesting property of this visualization is that it displays distribution of all color categories on the grid and with fuzzy borders, since chips, which are on the edges of categories or with low reliability gain a more neutral color[16]. On the other hand, the colors that are placed over the grid are highly related to the foci task responses, so they are prone to the errors and imperfections of the experimental data. If data contained no focal responses, the simplified colors for representation of categories were counted from the most salient chips (with high reliability).

It is important to note that the colors on the grid are not equivalent neither to the real best examples of the categories from the foci task (since we used an average from all foci chips for a category) nor to the real percepts of the speakers interviewed. The strength of this visualization method is the simplicity and intuitiveness of display. Even if the colors on the maps do not resemble the real percepts, they perfectly show the real distribution and saliency of color terms and their best examples[17].

A notable drawback of this visualization is that it does not respect the nature of composite categories. For example the category *grue* will not appear green or blue as the speakers would perceive or, more precisely, categorize in finer distinction, since they have named multiple foci for any composite categories. In a typical case the focal responses for grue are located in the middle of blue and in the middle of green samples (approximately around F16, G16 and F28, G28 on the grid). However, the visualization will display grue as a proportional mixture of foci, so it will produce something like teal or turquoise. On the other hand if the prototypical grue was really green-blue color it will account for the hypothesis about the different hues of Hering primaries mentioned in chapter 3.2. At least the fuzzy visualization can show, which of these composite sub-categories of basic terms (not yet emerged into two separate terms) is stronger. Here again, to emphasize the useful properties of this visualization, we will bring to front the intuitiveness. The colored areas of the map are intuitively comprehensible and comparable with each other and also with the real percepts of the observer, so it indirectly suggests the number and shape of basic color categories in a language.

[16] More precisely a mixture of colors of categories on which borderline the chip is. For example a chip between red and white will be light pink.

[17] Note that in hypothetical case, if a certain color category had only one focal chip and the reliability of this chip was 100%, in the fuzzy visualization it would be displayed with the same color value (the same color hue) as in the Munsell color grid.

5.2 WCS Visualizer

In the following part we will describe the software output of this master thesis, a standalone application programmed in Java. At first we will describe the structure of data from WCS archives and how they had to be processed to form reasonable data structures to be used for visualization task. In the next part we will portray how the application functions, then its graphical user interface, features and usage.

5.2.1 WCS data archives

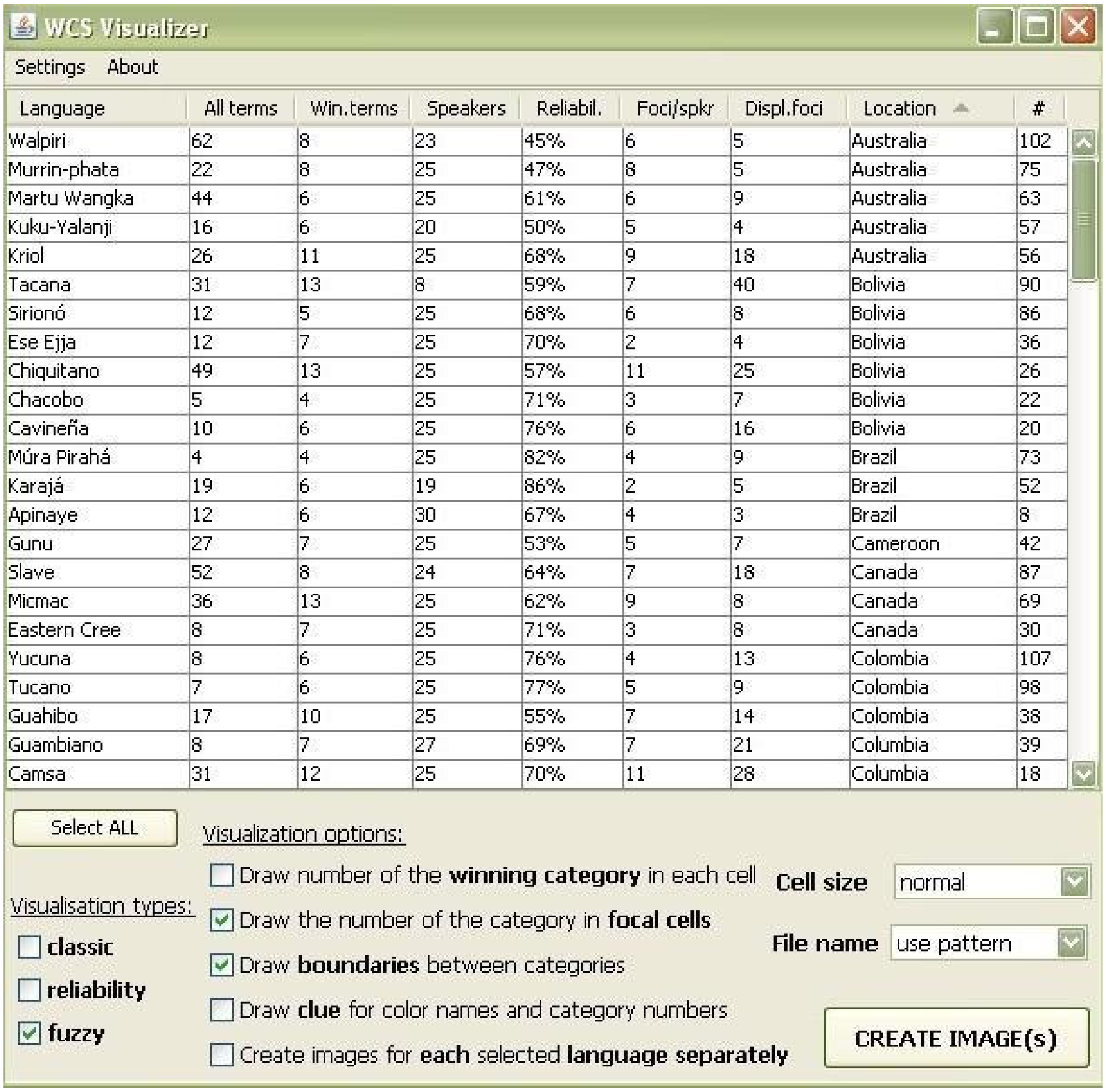
The World Color Survey online data archives [21] consist of several downloadable text files with quite peculiar structure, which is probably related to file size economy (even though the main data file has nearly 10MB). At first there is a file with mapping of WCS coordinates – chip numbers to their position in grid (rows in letters, columns in numbers). Then there is a file with language names and information about languages (which are incomplete, for the purposes of this thesis we completed them manually from [23]). The first file important for data processing is the dictionary, assigning to each term in each language an abbreviation, which is used in essential files with color naming and foci results. In the process of building reasonable structures the visualization program must at first read the dictionary file to create mapping between serial numbers of categories (terms), and words and abbreviations. Only then it is possible to process the color naming and foci task data. We aggregate the responses into arrays representing all color names for each color chip. In case of focal responses we do not only simply aggregate color samples proposed, but when a speaker has chosen multiple chips, only the right portion for each chip is added to the general sums. This mechanism ensures that each speaker's answers have a same weight.

5.2.2 Visualization process

Each time application is launched, it reads the data from the WCS data text files and builds the data structures necessary for generating the visualization images. These images are created according to properties selected by the user described in the next section. Since the color values of the color chips are encoded in CIE L*a*b* color space and the output image in RGB we use a special library for transfers between color spaces we developed in [14].

5.2.3 Graphical user interface and usage

The figure below displays the graphical user interface, the following text its parts, their meaning and usage.



Language	All terms	Win.terms	Speakers	Reliabil.	Foci/spkr	Displ.foci	Location	#
Walpiri	62	8	23	45%	6	5	Australia	102
Murrin-phata	22	8	25	47%	8	5	Australia	75
Martu Wangka	44	6	25	61%	6	9	Australia	63
Kuku-Yalanji	16	6	20	50%	5	4	Australia	57
Kriol	26	11	25	68%	9	18	Australia	56
Tacana	31	13	8	59%	7	40	Bolivia	90
Sirionó	12	5	25	68%	6	8	Bolivia	86
Ese Ejja	12	7	25	70%	2	4	Bolivia	36
Chiquitano	49	13	25	57%	11	25	Bolivia	26
Chacobo	5	4	25	71%	3	7	Bolivia	22
Cavineña	10	6	25	76%	6	16	Bolivia	20
Múra Pirahá	4	4	25	82%	4	9	Brazil	73
Karajá	19	6	19	86%	2	5	Brazil	52
Apinaye	12	6	30	67%	4	3	Brazil	8
Gunu	27	7	25	53%	5	7	Cameroon	42
Slave	52	8	24	64%	7	18	Canada	87
Micmac	36	13	25	62%	9	8	Canada	69
Eastern Cree	8	7	25	71%	3	8	Canada	30
Yucuna	8	6	25	76%	4	13	Colombia	107
Tucano	7	6	25	77%	5	9	Colombia	98
Guahibo	17	10	25	55%	7	14	Colombia	38
Guambiano	8	7	27	69%	7	21	Columbia	39
Camsa	31	12	25	70%	11	28	Columbia	18

Figure 22: Graphical user interface of the visualization application

1. **Table of languages** displays all languages and their characteristics. The first column contains their names (Languages), then follows the number of all terms of the languages vocabulary (*All terms*), winning terms (*Win.terms*), and speakers (Speakers), then average reliability per field (*Reliab.*), the average number of focal responses per speaker (*Foci/spkr*), the number of focal responses to be displayed in the visualization (*Displ.foci*), i.e. the number of chips which were selected as focal

with more than 10% agreement, the location (*Location*), where the language is spoken, and the serial number of the language (#). All columns can be used as keys for sorting, in this example the data are sorted by the location. This table enables the user to select languages to be visualized.

2. **The control panel** placed below the language table enables user to set basic properties of the visualization. Starting from left it contains:

    a. A button for selection of all languages in the table.

    b. Checkboxes for selection of the visualization type.

    c. The *visualization options* group for selecting whether numbers of categories and the focal numbers should be written in each field, then whether to draw the borders between categories and the clue, which represents the mapping for the serial numbers of the terms and their written transcript[18]. The last option is important, because it sets whether the selected languages should be placed together in one output image or rather in separate files per language.

    d. The *cell size* drop-down selection box containing values *normal*, *medium* and *small* denoting the size of the visualization images, more precisely the size of one chip from the grid. Since the program has its limitations it preprogrammed to set a smaller size when the amount of displayed items (i.e. types of visualization multiplied by the number of languages) rises beyond certain level. These levels are set to 40 items for normal size and 150 for medium. Another constrain given is that when small size is selected, both the clue and the numbers of categories are disabled, because a font in appropriate size will not be legible. Another modification is that labels for focal chips are not written numbers, but only small black rectangles. The clue is disabled also for the medium size.

    e. The *file name* drop-down selection box which allows the user to select how the output image should be named in case of the multi-language image. In case of generating separate images per languages the images will be always

[18] Note that the written form of the terms is the transcription of how their sounded to experimenters who recorded particular language data, since none of the languages surveyed has a written form.

named after the language. In the other case, there are two options. The first is to use the naming pattern, option *use pattern*, which can be changed in the upper menu under *File naming pattern* option (described later). The function of the file naming pattern mode is that the program always finds the largest serial number of the file present in the output directory and generates the image named using the pattern and a number higher then found. The other way is to use one fixed name for all images generated, so the previously made image will be overwritten by a new one. This function is active when the file name drop-down selection box is set to *use the same* option and the name in the *Universal filename* option in the upper menu.

f. The button *Create image(s)* which executes the creation of visualization images.

3. **The upper menu** contains a drop-down item: *Settings*, which contains:

a. *Output location*: an option for changing the location of the output images.

b. *File naming pattern*: an option for setting the pattern string for file names of multi-language images, displayed on the figure below. The symbol # stands for the serial number of the image. The right-bottom side dynamically displays how the names for output images will look like.

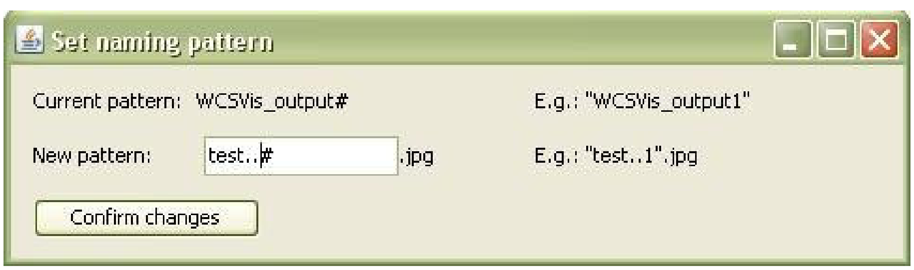


Figure 23: A window for changing the file naming pattern

c. *Universal filename*: an option for setting a fixed name for the *one name* option described before.

At the end of this section let us describe a model case of usage. We want, for example, one output image with six languages of largest numbers of elicited categories. We sort the table of languages by field *All terms* (click twice on the column header) and select the first six items. We want to display the visualization maps in reliability and the fuzzy mode, with foci, but without borderlines, and in small size, so we select the visualization types on the right, and second and third option from the middle. We pick the option *small* from the first drop-down selection box on the right and click the *create images(s)* button. The information about the language above its visualization informs the user about the number of color terms in the vocabulary, winning terms, and speakers, and average reliability.

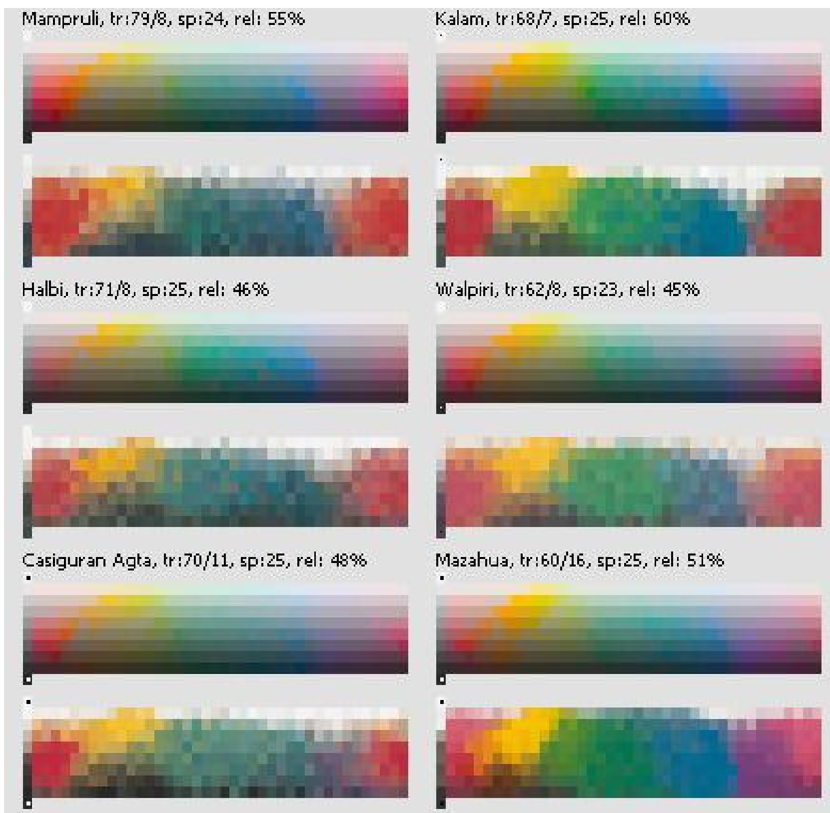


Figure 24: Fuzzy visualization for six largest vocabulary languages

This figure displays six languages with the highest number of elicited term. Note that in all six languages the number of elicited terms highly exceeds the winning terms. This inconsistency might have been caused by the misinterpretation or weak following of the instructions. Here you can also see that the fuzzy visualization is powerful even in case of such a great amount of categories, reducing the original 79 to approximately 6, consistently with Berlin and Kay's theory, in the topmost language (taking a closer look you can notice that this language already started the process of separation of green and blue but it is not yet so obvious) or in case of left-bottom language where the distribution obviously resembles the 11 basic categories also proposed by B&K.

5.3 Confrontation with existing studies, an example

In section 4.4.2 we described proposed conceptual framework for analysis of structure of the WCS languages. It emphasized the process of division of two composite categories into six Hering primaries in five stages with various possibilities of the distribution of these colors. Here we would like to show that intuitive – visual conclusions, which could be derived from the outputs of the fuzzy visualization, can be both in complementary and contradictory relation with this framework.

We would like to illustrate an interesting observation regarding the stages of development. The fuzzy visualization shows that composite category grue is present not only within languages with four winning categories, but also in larger schemes (up to 10 categories). The following figures display firstly one ideal case of language with four winning categories and their distribution according to proposed framework. Secondly a language with ten winning categories including *grue*, which's other categories are not in accordance with the proposed framework, since there are areas certainly depicting violet (or violet-pink), brown and even grey, which are by this framework and also by Berlin and Kay considered to be the second level categories which are "allowed" to emerge only when a language already has the Hering six primaries (black, white, red, yellow, green, and blue).

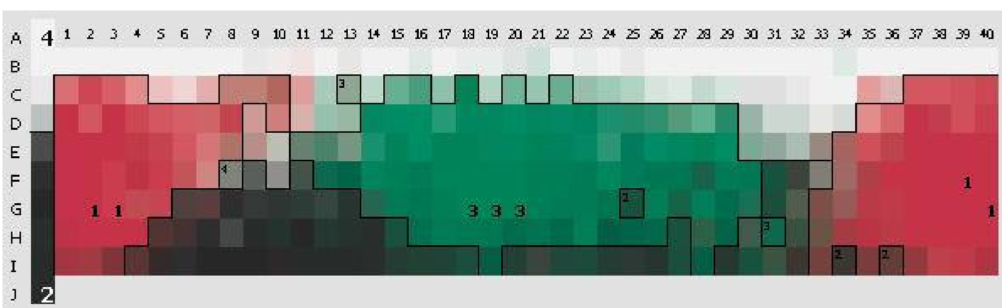


Figure 25: Fuzzy visualization for Múra Pirahá

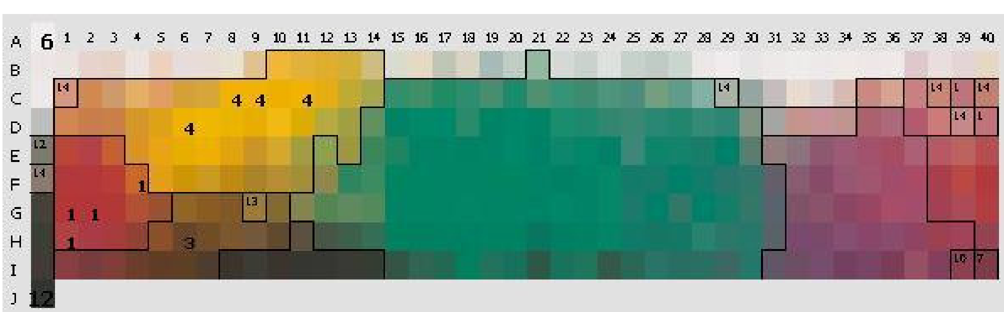


Figure 26: Fuzzy visualization for Mixteco

Since there can be an objection that the inconsistency example we used as a language with relatively low average reliability (circa 65%), we provide another example, a language with the second highest reliability from the whole set accounting for the same inconsistency. Chayahuita is a language from Peru with average reliability of 88%, which can be similarly to Wobé considered a high consensus language. Dissimilarly to Wobé, this language does not validate the universality of proposed framework. There is no additional computation needed to induce that this language has its seven winning basic terms equivalent to black, white, red, yellow (or yellow-orange), pink, grey-brown, and grue. Interesting is that grue which covers the whole green and blue area has its prototype only in blue, like there was no green at all. For better understanding we provide also classic visualization.

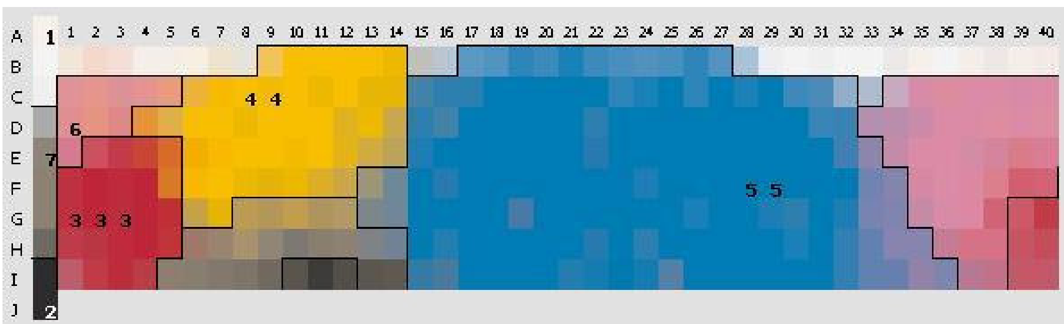


Figure 27: Fuzzy visualization for Chayahuita

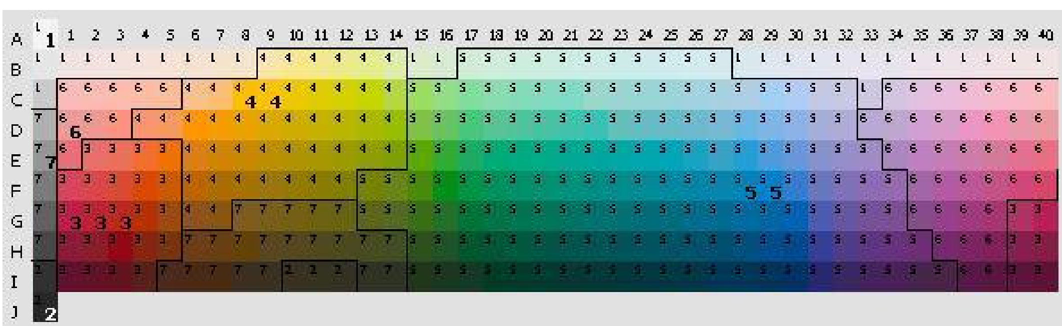


Figure 28: Classic visualization for Chayahuita

Hereby ends our introduction into various opportunities which our visualization tool provides. For further inquiries of the WCS data by means of this visualization see the application on the attached CD (see Appendix C). For more information about the WCS languages like location, number of speakers and terms, etc. see Appendix A.

5.4 Conclusion

The visualization we developed provides means for a wide range of comparisons, studies and discussions involving the WCS data. We briefly outlined a model case study regarding the fuzzy visualization, but without further explorations and without drawing any consequences, because that is not the aim or main theme of this thesis.

Considering the main contribution of this concept, we would like to emphasize the multi-dimensional character of the information displayed. Using the shape and colors of original stimulus grid we remind the observer of the experiment from which the data are drawn. Adding the numeric notation we display both the color naming and foci task responses, and using another visual component, the thin black lines, we denote the fixed borderlines between categories determined by the winning categories, which can be useful when displaying visualization maps without numbers. On the other hand, the fuzzy visualization encodes all information into color hues and relays on the intuition of the observer. Combining the colors of focal samples into proportional mixtures it provides not only the information about the most frequently named categories, but about all given responses, so the categories are displayed with fuzzy borders similarly to real perception and categorization.

The most important drawback of the fuzzy visualization is, that it does not reflect the composite character of some categories, e.g. grue, that usually have two or more focal hues located in different basic colors terms, that are expected to emerge later (e.g. grue decomposed to green and blue). Other inconvenience is that without knowledge of basic theory behind it, the direct interpretation of the fuzzy maps might cause confusions and misunderstandings. Crucial is to distinguish, that the fuzzy visualization does not directly reflect the perception of speakers of particular language, but rather the distribution of its color terms and its relationship to other distributions and intuitive relation to any color system the observer (user) can imagine. Unfortunately, there is no artificial intelligence based simulation program with pre-learned 11 basic color categories from English or from the ideal partitions of perceptual color space [8][27], that could be used to compare the fuzzy visualization with its output or otherwise produce a list of categories that form salient regions on the fuzzy maps. Therefore this task remains for the judgment of the observer.

6 Distinguishing criteria based simulation of color categorization

In this chapter we would like to present the color categorization simulation based on the semantics of distinguishing criteria and the results of this simulation carried out using the data from the World Color Survey and from Slovak color experiment (for more details on these studies see chapter 4). The aim of this simulation is not to prove or disprove the original hypothesis of Berlin and Kay (section 4.1), but rather to probe the learning capability of the distinguishing criteria on the real data represented in a perceptually uniform space. We will at first provide a brief notion of what the distinguishing criteria are, then the description of implementation of this semantics, the methods used for evaluation of the results of this simulation, and finally the results and discussion.

6.1 Distinguishing criteria

As originally proposed by Šefránek [28] a distinguishing criterion is a computational abstraction of meaning modeled by a locally tuned detector reacting to some part of its input space [29]. Under the term meaning we understand one category; in our case is it one color category. Distinguishing criteria (DC) have these important properties:

1. **Learnability**: they can be (and ought to be) incrementally and continually constructed from an incoming sequence of examples[19].

2. **Identification**: each single criterion can express for each given input its *degree of membership* in a category represented by given criterion. In other words, the distinguishing criterion can express if the given input is an instance of the concept, which it represents. The value of activity of the DC for an input is expressed by a real number from [0, 1]. The closest it is to 1, the more the input represents the category, if it is 1, the input represents the prototype of the category, its best example.

3. **Auto-associativity**: even for a possibly incomplete or noisy input, it returns the best example (prototype) of the represented concept.

[19] For details of learning mechanism see [28].

The second property of the distinguishing criteria is due to their character of locally tuned detectors, which can be intuitively represented by conceptual spaces. A conceptual space is a geometric space with dimensions corresponding to the attributes of represented objects organized in domains, while a particular object is represented as a point (vector of coordinates) in a subspace of one or several domains. In our case of color categories we will have no difficulties with the accordance of the domains of compared objects (categorical inputs represented as points in L*a*b* color space), so there is no need to describe how the distinguishing criteria deal with entities with attributes from varying domains. Important is, that the natural categories (which include also color categories represented in perceptual color space, see section 2.5) are represented by convex regions in the space[20]. The best examples are then the geometrical centroids[21] of these regions.

The implementation we use computes the prototype as an average of all inputs. Each color category c is represented by a locally tuned detector $r_{\vec{p}}$ ($\vec{p}$ stands for the prototype of the category c). The degree of membership of a certain color input $\vec{x}$ (one point in L*a*b* color space, i.e. three-dimensional vector) will be computed as an exponentially decaying function of its distance from the prototype: $r_{\vec{p}}(x) = \exp(-k.d(\vec{p},\vec{x}))$, where k is some positive constant, and d is the metric used. We used the covariance based distinguishing criteria, which track down the distribution of inputs in the covariance matrix sigma and compute the activation for an input using the Mahalanobis distance $d^2_{\Sigma^{-1}}(\vec{p},\vec{x}) = (\vec{x}-\vec{p})^T \Sigma^{-1}(\vec{x}-\vec{p})$, where $\vec{p}$ and $\vec{x}$ are column vectors and Σ^{-1} is the inverse of the covariance matrix of the sample set used for training. For more mathematical details see [29].

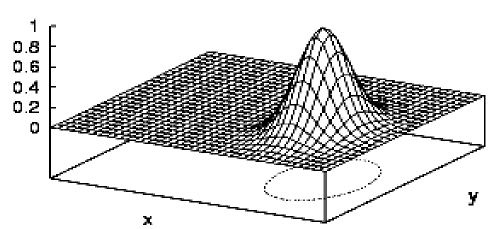


Figure 29: Illustration of a distinguishing criterion
The receptive field (with the threshold of 0.1) of a 2-dimensional locally tuned detector, figure from [29].

[20] If two points represent objects that are good examples of a category, then any point in between them must also be a good example of that category.
[21] A centroid is a geometric center of the object's shape. In the simples case of a triangle it will be the intersection of three lines connecting its vertices.

6.2 Simulation of color categorization

6.2.1 Implementation

We have implemented the simulation in Java using the distinguishing criteria framework developed by Takáč [29]. From his project we used the distinguishing criteria libraries for the covariance-based criteria and the related math libraries. Similarly to what had to be done in WCS Visualizator application, this case required the processing of WCS data archives into the form of "reasonable" data structures (for more details see section 5.2.1). For this purpose we developed a class *WCSDataStore*, which in its initialization puts the WCS data together. Then it produces train sets with all naming task results from a language to be presented to the distinguishing criteria and data sets with summarized data from color naming task and foci task to be used for the final comparison of the simulation outputs with the original WCS data. The categorization simulation program also includes a module for visualization of the results, a library with necessary mathematical equations, and a package for conversion between color spaces from [14].

Constructing the simulation for each language we at first created a set of distinguishing criteria (instances of DC classes), one for each color category from the language's color vocabulary (see WCS data structure in 5.2.1). For representation of inputs we have chosen the CIE L*a*b* color space mainly because it is a perceptual color space, so the distance between any two point in this space is equivalent to the perceptual difference between the two colors (see section 2.5). Secondly, we used this particular perceptual color space, because the color samples from the experiment were encoded in it within WCS data archives.

In the process of teaching we took each response of each speaker for each sample[22] and "fed" them to the distinguishing criteria. Each criterion received only the values of the samples which were called with the corresponding color term (a name of the color category). We used the same mechanism for the Slovak color experiment data. Note that also in this chapter we have allowed us to interchange the terms *color term* and *color*

[22] That is approximately 330 x 25 = 8250 categorical responses (depending on the number of speakers), which can be imagined in a form: *#: {L, a, b}*, where # represents the serial number of the category assigned to a color sample by one speaker and *L*, *a*, *b* are values of the color coordinates of that sample.

category freely. However we are aware of the distinction that the color term is the name of the color category.

6.2.2 Methods of evaluation

This section will describe how we evaluated the results of the categorization simulation. Here we will use the specific terminology related to WCS data evaluation and visualization described in section 5.1.1. We compared the results for each of WCS languages in two different ways.

First the **winning category** for each of 330 samples was computed from the data as the term which was used to name the particular sample at most and compared with the winning category from the simulation represented by the distinguishing criterion with maximum activity for the given sample from all criteria. If the winning criterion matched the winning category computed from the data, we added 1 to the sum of matches. Then we divided this sum by the number of all samples and recorded the resulting measure of agreement in percents.

For more complex comparison of the categorization outputs we designed the **vectors of activities**. These are normed vectors of the categorical responses for each sample for both the summarized speaker's data and distinguishing criteria outputs. In case of the speaker responses it stands for the frequency of how often a term was elicited for a given sample. For example if there were 3 terms (categories) and 10 speakers for each sample a distribution of the elicited terms was created in form of a vector, lets say: [2,3,5], where the serial number of the category is the serial number of the component and its value is equal to the number of speakers, that have used the given term. In the latter case the values of components of the vector (corresponding to all terms of the language, in the same order as in the speaker-responses case) correspond to the value of the activity function of the particular criterion for the given sample.

Both vectors are normed using the Euclidean norm (each component is divided by the size of the vector, which is the square root of the sum of squares of each component), so they can be compared as two unit vectors in n-dimensional hyper sphere (where n represents the number of elicited categories). In two-dimensional case it would be two unit vectors on the unit circle. For the comparison of the two vectors we have chosen the scalar product (dot product), which directly corresponds to the cosine of the angle between these two vectors (since we have unit vectors, their size will not count). If the scalar product is 1, they are identical; if it is -1, they are opposite to each other. For the summary comparison with other languages we recorded the average value of scalar products of normed activity vectors for all samples.

6.2.3 Visualization of results

Since we developed a useful method of visualization of the WCS data, we decided to apply it on the results of our categorization simulation to provide also a visual comparison. For the types of the visualization and a detailed description of it see chapter 5 (namely section 5.2). For the best understanding of the results we decided to create both the classic and the fuzzy visualization. In case of displaying the results of DC we used the above methods to decide the winning categories and according to them the boundaries between categories were drawn. In case of fuzzy visualization of DC outputs we used a similar algorithm as for the WCS data visualization. The color of each field is a sum of the color values of the prototypes of categories, each multiplied by the normed activity of that category for given sample. The full visualization output (an image in .jpg format) for one language contains the name and some characteristic information of a language (e.g. number of elicited terms, winning terms, speakers, etc.; see section 5.2.3) plus information about the average results of the simulation (one line above the images), then the classic visualization and the fuzzy visualization of the original WCS data on the left side and the same types of visualization for distinguishing criteria on the right side.

Since the distinguishing criteria compute their prototypes as an average of the WCS color samples, the fuzzy visualization maps created without any normalization are always rather brownish due to lowered saturation. To compensate this side effect we decided to increase the saturation of the prototypes of the DC to match the saturation of the original samples. For each category we counted the average distance of all samples presented to the corresponding discrimination criterion from the center of L*a*b* color space (point with coordinates [0,0,0]) and used it to norm the distance of the criterion's prototype. We did so by multiplying all the prototype's coordinates by the computed distance and dividing them by the original distance of the prototype from the center. We used this technique only for the visualization.

6.3 Results and discussion

In this section we would like to present a general overview of the results of the categorization simulation together, some interesting examples of the visualization outputs and the results of the Slovak data simulation. In general the distinguishing criteria have proofed themselves quite good in resembling the distribution of categorical responses of the speakers, and in smaller extend they also gained success in the winning terms comparison, mostly within languages that have a small difference between elicited and winning color terms. The full result table is included in the Appendix B and the whole set of visualization images is included on the attached CD (see Appendix C).

6.3.1 Winning categories comparison

For each language we compared the winning categories computed from the data (as the each sample's most frequently used term[23]) with the winning categories from the language's distinguishing criteria (the criterion with the strongest activity). We compared them for each sample and noted the *agreement* between them in percents. The winning categories from the simulation matched with those from data at 49,253% in average. The best case was Chayahuita with 88,788%; the worst was Casiguran Agta with only 6,061% simulated winning categories matching the real winning categories.

We found out that there is a correlation between the number of elicited[24] terms and the agreement of the winning terms from data and from simulation. The more categories were present in a language, the more "confused" the distinguishing criteria got, i.e. the smaller was the percent of the matching winning categories. This happened especially within languages with a high amount of elicited terms and significantly smaller amount of winning terms (2, 3, or more times smaller).

The phenomenon is displayed on the following grapg. It displays the average winning category agreement for each language. Languages are sorted primarily by the difference between the amounts of elicited and winning categories (elicited minus winning), secondly by the amount of elicited categories and thirdly by the amount of winning

[23] For more details about winning terms see section 5.1.1. The details of the comparison mechanism are in section 6.2.1.

[24] By elicited terms (or categories) we mean the whole color vocabulary of a language, i.e. all terms elicited in the color naming task. In comparison the winning terms/categories include only terms, from which each was at least once (for at least one sample) the most frequently used term.

categories. For completeness it also displays the similarity measures of the activity vectors, which will be discussed in the next section. Note that the labels of the x-axis do not display all languages because there is not enough space, but each language is represented by one line of the grid.

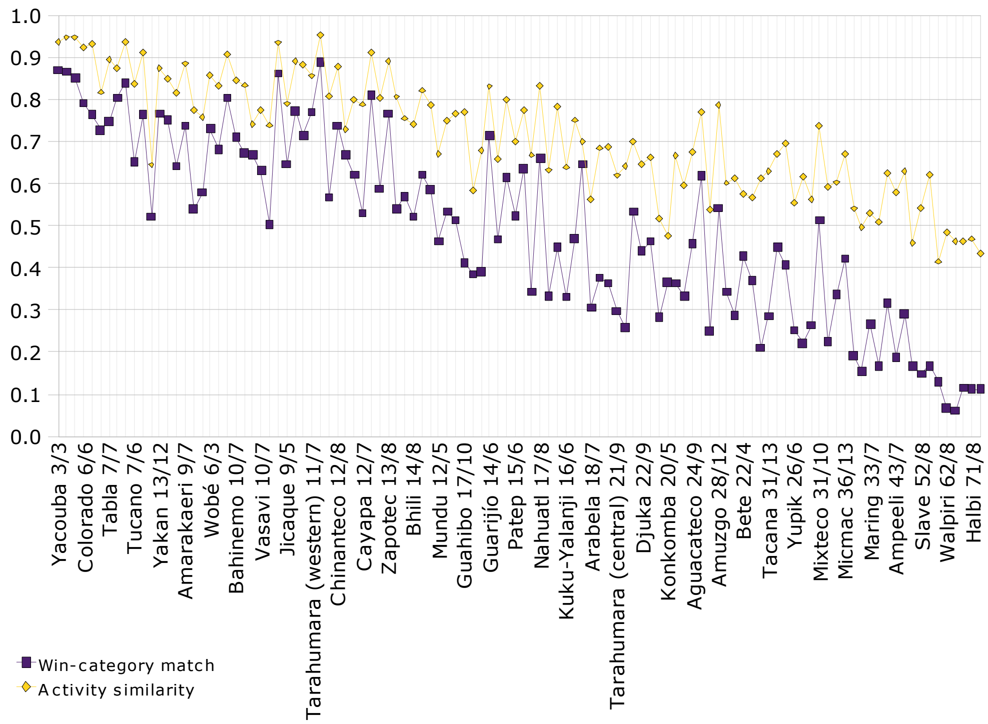


Figure 30: Graph of average winning term agreement and average activity similarity values of WCS languages ordered by the difference between the amounts of elicited and winning categories

Despite the deviations in some points the graph shows a decreasing tendency in both winning category matches and activity vector similarities. However the decreasing tendency of the activities starts not with the smallest differences, but within languages that have circa two times more elicited terms than winning terms. The possible explanation is that the spurious terms used only by few speakers are mixed up with the real, generally used terms. Since when computing the activities of criteria and choosing the winning category we do not make a difference between the criteria presented with small and presented with large amount of samples by (we only compare the activities of the criteria for certain input), these small categories might "steal" the place of the more important categories just because their prototype somehow fits the sample better.

6.3.2 Activity vectors comparison

Activity vectors represent the distribution of categorical responses for a given sample from both speakers' data and activation of distinguishing criteria, each of them with the same dimension equal to the number of categories in a given language. Each component of such vector represents the "strength" of one category (with serial number equal to the serial number of the component) for the given color sample. We computed the relationship of these two vectors (normed to unit vectors, see 6.2.2) on the basis of scalar product, which is in this case equal to the cosine of the angle between the two vectors. The closer this value is to 1, the more similar the vectors are; if their scalar product is -1, they are opposite (the angle is 180˚).

The activation similarity measures ranged from 0,415 (± 0,166) in language Iduna up to 0,954 (± 0,087) in Chayahuita. The average value from all languages was 0.715 (± 0.139). The correlation between the measure of similarity of activity vectors and the winning terms agreement is already visible on the graph on the previous page. The graph below displays their direct relationship, which has roughly linear character. With increasing similarities in the activities of all categories rises the chance to decide the winning terms properly. However there are cases with good performance in winning categories matching independently of the below standard performance in activity vectors similarity (the local maxima). These are, as we assume, caused by the high inter-speaker agreement (reliability), which on the other hand does not have a general influence on all languages.

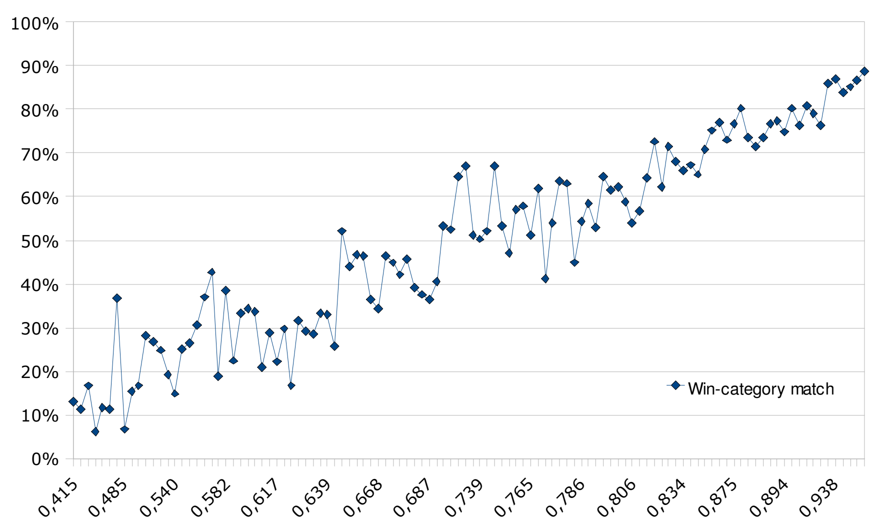


Figure 31: Graph of the relation of activity vector similarity measure and winning term agreement

6.3.3 Visual output

The visual outputs from the color categorization simulations are quite interesting. As we emphasized before, the advantage of the fuzzy visualization we proposed in chapter 5 is the more acute resemblance of the distribution of categorical responses (activities), which were also used as one measure of success of the simulation. The winning terms are displayed directly on the classic visualization maps and are also responsible for the borders between categories (for more information about visualization see chapter 5). The fuzzy visualization, on the other hand, encompasses the distribution of categorical responses for both speakers (left side) and criteria (right side). Since the visualization of the distinguishing criteria (section 6.2.3) was created using the portions of prototypes (color values in L*a*b*) of the categories (criteria), that were computed from all samples presented to a criterion, they can better reflect the real colors of the whole set of color samples from the experiment. The focal samples' colors used in the fuzzy visualization for the original data are, on the other hand, computed only from the best examples chosen by the speakers, which might not take into account all sub-categories of a composite category, or categories that are in the process of emergence.

The following example displays language Vagla, an example of successful categorization simulation as well as an example of a good dataset, since the number of elicited terms is equal to the number of winning terms (all terms were used by all speakers). Also the average reliability (inter-speaker agreement) is relatively high (70%). The average winning category agreement was 76% and an average categorical activity accordance was 0.932 (± 0.063).

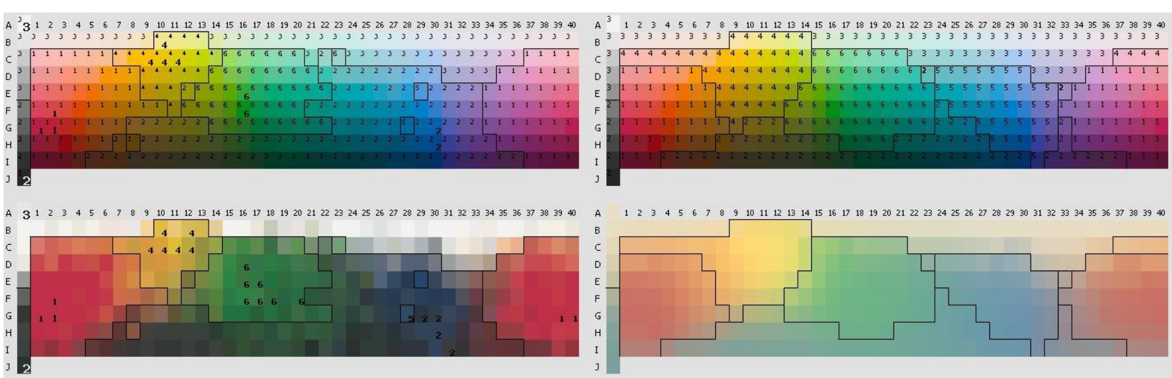


Figure 32: Classic and fuzzy visualization of languageVagla and its output from the simulation
Note that in the left case there are only too fields where the new emerged category for blue had won. On the other hand according to the distinguishing criteria it takes up space which we Europeans would also label as blue. Might this mean that the criteria are able to predict the emergence of new color terms?

The second example will display another one, but not so welcome epiphenomenon of the distinguishing criteria based categorization. As we mentioned before, in case of a high amount of elicited categories, the distinguishing ability decreases. A simple explanation is that since there are too many criteria groups of criteria with nearly the same prototype emerge and cause the confusion. In this case a criterion, which has a high prototype and spatial overlap with another criterion, will consider a good example also a sample that was never presented to it and "steal" it from the other criterion which was trained to it. The example below depicts the worst case of the activity vector comparison and the third worst from the winning category comparison. Note that in case of simple processing, where the spurious responses are removed by the winner-take-all mechanism, it becomes obvious that this is a nice 5-term language example, with black, white, red, yellow and grue. However there are also some spots within the fuzzy visualization which show inconsistencies in category activities, characteristic by slightly desaturated colors.

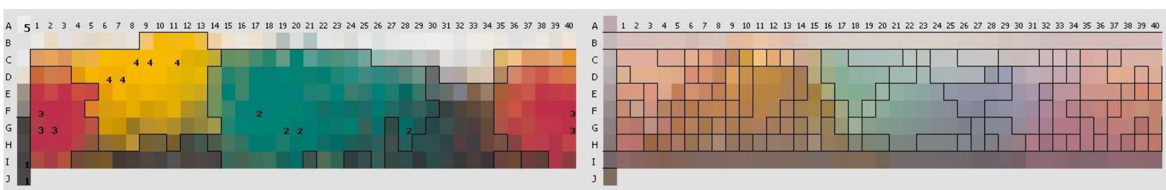


Figure 33: Fuzzy visualization of Iduna and of its simulation
The first map is the fuzzy visualization from the data, the second from the simulation. The dictionary part of the data for this language consisted of 51 terms, from which only six became the winning terms.

The last example accounts for the same problem as previous one. It shows as well that in case of diverse inputs the prototypes get more saturated and hence less productive. This language was one of the worst with only 15.45% average agreement of winning terms and nearly the worst average similarity of the activity vectors.

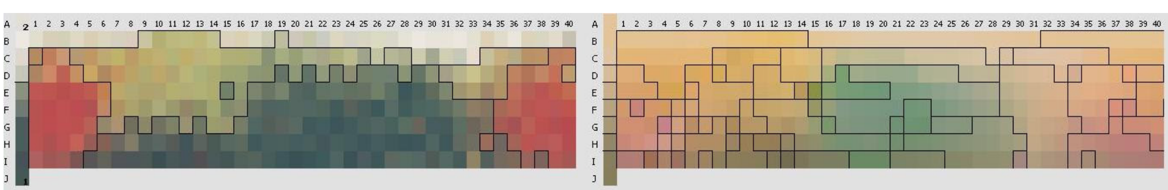


Figure 34: Fuzzy visualization of Culina and of its simulation
The left map is the fuzzy visualization from the data, the right from the simulation. This language has total 29 color terms from which only 4 were winning terms.

6.3.4 Slovak language simulation

This section will discuss the results from the simulation carried out on the data from the Slovak color experiment (see section 4.3). Since we collected two sets of data, with and without the six *outer* color categories[25], we have executed the simulation twice. The data set with the outer categories resulted as less consistent and similar to the middle cases from the WCS languages. The similarity with problematic cases with spurious categories can be confirmed also by the fact that from 16 terms of the used vocabulary, only 13 were winning (and one of them had won only for one sample). In section 4.3 we concluded that only one new category (turquoise) has a chance to emerge in the Slovak language, but the simulation ended differently.

The average winning category agreement for the enhanced set (11 basic plus 5 outer categories) was only 52.72% and the average activity vector similarity was 0.769 (± 0.202). The results for 11-term set were better, 79.1% and 0.873 (± 0.176), placing this simulation into the group of 20 best results. However these results were not as good as expected, what might have been caused by inconsistencies in the responses of the speakers and generally bad conditions under which this experiment has been carried out (discussed in 4.3.2). The pictures below display the fuzzy results of both simulations.

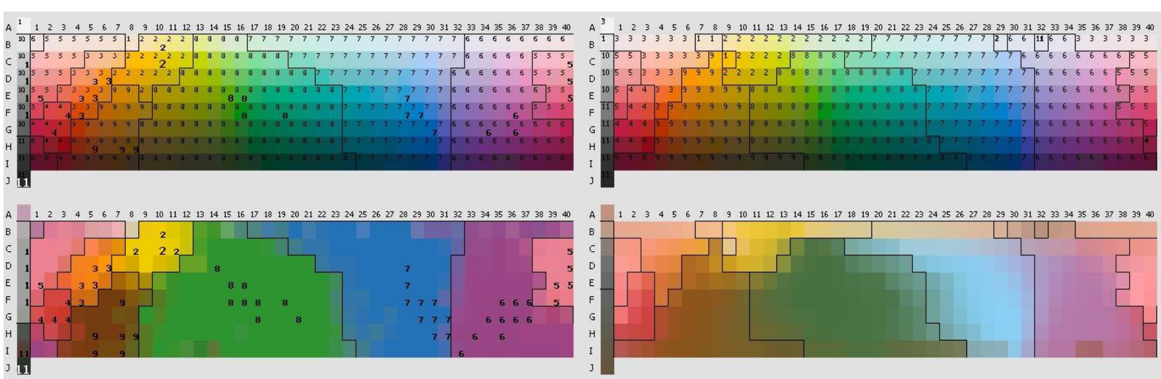


Figure 35: Classic and fuzzy visualization of Slovak data and simulation with 11 categories

[25] The colors which were successfully recognized among significant amount of speakers other then 11 basic categories proposed by Berlin and Kay, more details in section 4.3.

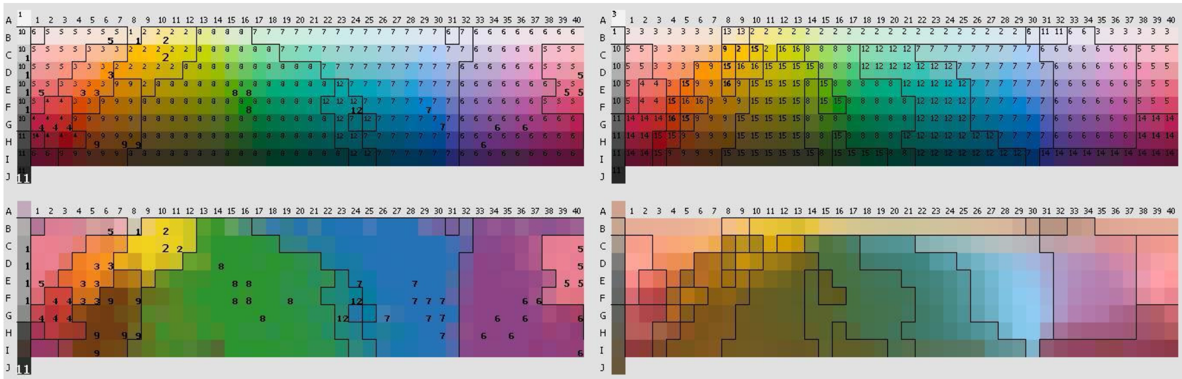


Figure 36: Classic and fuzzy visualization of Slovak data and simulation with outer categories

Concluding the visualizations, the resemblance of color categories was in some cases really good, but in some cases it failed. The visualization maps from the original data (on the left side) of both versions of simulation reserve only small spots for the brown color, on the other hand the distinguishing criteria considered as brown large parts of orange and green. Although there were also speakers who had difficulties with so-called khaki, in hue columns 10 - 13 (remarkably large area in the first image on the right side of brown), they finally decided that it was a subset of green (the map in the left bottom corner). However the criteria display it differently.

The most probable reason for the brown-blending and unadequate khaki-winning effect is what we already mentioned in the visualization section of this chapter (6.2.3) as the prototype desaturation[26]. A middle of two higly saturated points in L*a*b* color space is in general closer to the center of the space than the two original points, i.e. it has lower saturation. Since the samples for brown are darker and dimmer we assume that they overruled larger areas of samples, since the prototypes for orange, yellow and green ended up less salient and unable to compete with brown prototype. On the other hand many speakers complained about the colors of samples as brownish and ugly, so this effectmay be caused also by the quality of samples. In case of khaki, we assume that even small amount of samples could build a prototype strong enough to cover the whole controversial area (column 10-13), because the prototype for green is more general, constituted from relatively larger amount of samples.

[26] The average value of two points in the L*a*b* color space is in their middle. Since this color space has a spherical shape, if the two points have the highest saturation, i.e. the maximum distance from the 0 point, the distance of the middle point must be shorter.

7 Conclusion

In the first part of this master thesis we provided an essential theoretical background for understanding the phenomena involved in color categorization – color perception and categorization. We presented a definition of basic color terms (the labels for basic color categories) as the smallest set of simple words with which all speakers of a language can consistently name any color. Then we described main hypotheses proposed about them, which stated that each language has at least 2 and maximum 11 basic color terms and that the best examples of these color terms are similar (identical) with those from English: black, white, red, green, blue, yellow, purple, pink, brown, orange and grey. The first six of these terms are so-called Hering primaries that are involved in opponent processes underlying the color vision and are expected to emerge in any language at first. These hypotheses were tested in the World color survey, which's output consists of the results from color naming experiment from 110 primitive languages with no written form and of non-industrialized cultures. We took this data from the freely available online archives and processed them.

Concluding the progress and results of our contribution to this topic, we at first gained a personal experience with methods of the World Color Survey and color naming data for Slovak language by a simulation of the WCS experiment. In the second step we developed a new method of visualization for the WCS data (or any other data from the same color naming experiment). At last we used the WCS data to probe the learning capability of the categorization simulation based on the semantics of distinguishing criteria.

The results from the Slovak color experiment confirmed that Slovak language contains firmly established 11 basic color categories as proposed by Berlin and Kay as we expected. What we consider a more important aspect of the results of this experiment is that we found some general shortcomings of the WCS method. First of all, the 330 sample set seems to be too large and tiring amount influencing the performance of subjects. We also noted that a significant portion of samples were too similar. The most important fact we learned is that the speakers must be fully aware that the task is to use basic color terms, not the terms they personally like or consider important. Another problem we found is related to the method of the foci task, since speakers are allowed to produce unlimited amount of focal responses for any category and since the method does not involve any kind of direct checking of the responses with the color naming task responses, speakers could both choose too many samples or the samples that they had previously labeled as some other category and so produce inconsistencies in the results.

The visualization we developed provides means for a wide range of comparisons, studies and discussions. It presents the WCS data in three complementary ways. In the first two types it uses the shape and colors of original stimulus material and thus informs the observer about the experiment in which the data originate. The most innovative type called fuzzy visualization displays not only the winning terms (which were used for a certain sample at most), but encompasses the whole distribution of the subjects' responses, combining the colors of focal samples into proportional mixtures according to the results of the naming task The strength of this visualization method is the simplicity and intuitiveness of display. Although the colors on the resulting maps do not resemble the real percepts, they do show the real distribution and saliency of color terms and their best examples. However there is at least one important drawback, which is that this visualization is unable to reflect the character of composite categories (that have multiple different foci), since it uses proportional mixtures of all prototypes. It is important to distinguish that the fuzzy visualization does not directly reflect the perception of speakers of a particular language, but rather the distribution of its color terms according to the original stimulus material.

The results of the color naming simulation suggest that the distinguishing criteria are able to catagorize colors similarly to humans. More than a half of languages proceeded with 50% success in the winning categories aspect and even more of them were quite good in resembling the distributions of categorical responses. We found out that the success of the simulation significantly decreases with the increasing size of the color vocabulary of a language, especially in the cases of languages with a high amount of elicited terms and significantly smaller amount of winning terms (2, 3, or more times smaller). A possible explanation is that there were too many spurious terms (those used only by few speakers) which got mixed up with the real, generally used terms and since they were trained only on few samples, and their prototypes were closer to these samples than the prototypes of the more general critera.

8 Bibliography

[1] Lakoff, G.: *Women, Fire and Dangerous Things*. Chicago: The University of Chicago Press. 1987.

[2] Wilson R. A. and Keil F. C. (Eds.): *MIT Encyclopedia of Cognitive Sciences*. Cambridge, MA: MIT Press. 1999.

[3] Starr, C.: *Biology: Concepts and Applications*. Pacific Grove, CA: Thomson Brooks/Cole. 2000.

[4] Green, M.: *SBFAQ Part 1: Basic Terms and Definitions*. Visual Experts. 2004. <http://www.visualexpert.com/FAQ/Part1/cfaqPart1.html>

[5] *Light*. Encyclopædia Britannica Online. 2009. <http://www.britannica.com/EBchecked/topic/340440/light>.

[6] Fairchild, M.D.: *Color Appearance Models, 2nd Ed.*. Chichester, UK: Wiley-IS&T. 2005.

[7] Gouras, P.: Color Perception. In *WebVision*, *The organization of the Retina and Visual System,* John Moran Eye Center, University of Utah. Last update: February, 2008. <http://webvision.med.utah.edu/KallColor.html>

[8] Jameson K.A. and D'Andrade R.G.: It's not really Red, Green, Yellow, Blue: An inquiry into perceptual color space. In Hardin, C.L. and Maffi, L. (Eds.): *Color Categories in Thought and Language*. Cambridge, MA: MIT Press. 1997, p. 295–319.

[9] *Ľudské Telo*. Bratislava : GEMINI. 1991. (orig.: The Human Body. London: Marshall Editions Ltd. 1989)

[10] Gouras, P.: Color Vision, In *WebVision*, *The organization of the Retina and Visual System,* John Moran Eye Center, University of Utah. Last update: February, 2008. <http://webvision.med.utah.edu/KallColor.html>

[11] File:Cones SMJ2 E.svg. in *Wikipedia the Free Encyclopedia*. Last update: October 2008. <http://en.wikipedia.org/wiki/File:Cones_SMJ2_E.svg#filelinks>

[12] Conway, B.R. Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). In *Journal of Neuroscience*. 21 (8). 2001. pp. 2768-2783.

[13] McCann, J.J., Local-Global Mechanisms for Color Constancy, in *Die Farbe 34*, 1987.

[14] Rebrová, K.: *Vnímanie a pomenovávanie farieb a farebných kategórií* (Color perception and color naming), Bc. thesis. Comenius University. Bratislava, Slovakia. 2007.

[15] Munsell colour system. In Encyclopædia Britannica Online. Last update: 2009. <http://www.britannica.com/EBchecked/topic/397642/Munsell-colour-system>

[16] Murphy, G.L.: *The Big Book of Concepts (Bradford Books)*. Cambridge, MA: MIT Press. 2004.

[17] Margolis, E. and Laurence, S. (Eds.).: *Concepts: Core Readings*. Cambridge, MA: MIT Press. 1999.

[18] Berlin, B., Kay, P.: *Basic Color Terms: Their Universality and Evolution*, Berkeley, CA: University of California Press. 1969.

[19] Heider (Rosch), E. Universals in color naming and memory. *Journal of Experimental Psychology* 93. 1972. pp. 10-20.

[20] Kay, P., Berlin, B., Maffi L., and Merrifield, W.: Color naming across languages. In Hardin, C.L. and Maffi, L. (Eds.): *Color Categories in Thought and Language*. Cambridge, MA: MIT Press. 1997

[21] Cook, R. S., Kay, P, Regier T.: *WCS data archives*. Page Created: 2003/06/01. Last Update: 2006/11/14. <http://www.icsi.berkeley.edu/wcs/data.html>

[22] Cook, R. S., Kay, P, Regier T.: The World Color Survey database: History and use. In Cohen, H. Lefebvre, C.: *Handbook of Categorization in Cognitive Science*, St. Louis : Elsevier, 2005.

[23] Kay, P. and Regier, T.: Resolving the question of color naming universals, *PNAS* 100, 2003. pp. 9085-9089

[24] Lindsey, D. T. and Brown, M.A.: Universality of color names. In *Proceedings of the National Academy of Sciences* 103 (October 2006). pp. 16608-16613.

[25] Belpaeme, T. and Bleys, J.: Colourful language and colour categories. In Second International Symposium on the Emergence and Evolution of Linguistic Communication.2005

[26] *Systems by George Hernandez*. Last update: 2007/10/21. <http://www.georgehernandez.com/h/xzMisc/Color/Systems.asp>

[27] Regier, T., Kay, P. & Khetarpal, N.: Color naming reflects optimal partitions of color space. In *PNAS* 104. 2007. pp. 1436-1441

[28] Šefránek, J.: Kogníçia bez mentálnych procesov. In Beňušková, Ľ. et al. (Eds.): *Kognitívne vedy.* Bratislava: Kaligram, 2002.

[29] Takáč, M.: Autonomous Construction of Ecologically and Socially Relevant Semantics. *Cognitive Systems Research* 9 (4), October 2008, pp. 293-311.

Appendix A

Table of characteristic information about WCS languages processed from WCS data.

Language	Location	Terms All	Terms Winning	Spkrs.	Reliability (in %)	Avg.foci /speaker	Displ. Foci
Abidji	Ivory Coast	9	7	25	55.01	6	8
Agarabi	Papua New Guinea	27	10	24	53.72	6	7
Casiguran Agta	Philippines	70	11	25	48.26	5	6
Aguacateco	Guatemala	24	9	35	61.10	8	10
Amarakaeri	Peru	9	7	6	73.65	8	26
Ampeeli	Papua New Guinea	43	7	27	52.33	9	8
Amuzgo	Mexico	28	12	25	67.10	10	19
Angaatiha	Papua New Guinea	16	7	25	61.15	6	10
Apinaye	Brazil	12	6	30	67.04	4	3
Arabela	Peru	18	7	25	58.98	6	10
Bahinemo	Papua New Guinea	10	7	25	65.02	7	10
Bauzi	Indonesia	8	5	25	78.61	4	10
Berik	Indonesia (Irian Jaya)	29	6	25	61.66	7	13
Bete	Ivory Coast	22	4	25	75.22	2	7
Bhili	India	14	8	25	66.70	7	11
Buglere	Panama	27	6	25	50.12	7	11
Cakchiquel	Guatemala	28	17	30	61.70	8	6
Ucayali Campa	Peru	16	7	25	45.82	5	3
Camsa	Columbia	31	12	25	69.53	11	28
Candoshi	Peru	18	8	11	72.35	6	14
Cavine?a	Bolivia	10	6	25	76.36	6	16
Cayapa	Ecuador	12	7	25	67.19	4	7
Chacobo	Bolivia	5	4	25	70.99	3	7
Chavacano	Philippines	19	11	25	72.72	8	21
Chayahuita	Peru	11	7	25	87.64	3	11
Chinanteco	Mexico	12	8	25	70.43	7	15
Chiquitano	Bolivia	49	13	25	57.04	11	25
Chumburu	Ghana	24	9	25	63.93	8	15
Cofán	Ecuador	9	6	25	78.72	6	13
Colorado	Ecuador	6	6	25	73.19	5	9
Eastern Cree	Canada	8	7	25	71.14	3	8
Culina	Peru, Brazil	29	4	25	66.73	3	1
Didinga	Sudan	7	7	25	85.58	7	14
Djuka	Surinam	22	9	25	60.57	8	13
Dyimini	Ivory Coast	12	9	25	69.59	7	13
Ejagam	Nigeria, Cameroon	7	3	25	86.92	3	5
Ese Ejja	Bolivia	12	7	25	69.75	2	4
Garifuna (Black Carib)	Guatemala	18	10	28	60.28	8	26
Guahibo	Colombia	17	10	25	54.73	7	14
Guambiano	Columbia	8	7	27	69.07	7	21
Guarijío	Mexico	14	6	25	79.33	6	12
Guaymí (Ngäbere)	Panama	30	7	25	59.45	7	14
Gunu	Cameroon	27	7	25	53.08	5	7
Halbi	India	71	8	25	45.95	4	6
Huastec	Mexico	12	8	25	75.38	5	9
Huave	Mexico	17	11	25	71.05	9	20
Iduna	Papua New Guinea	51	6	25	69.75	4	10

Ifugao (Keley-i)	Philippines	23	8	25	63.09	5	7
Iwam (Sepik)	Papua New Guinea	10	7	25	55.60	6	11
Jicaque	Honduras	9	5	10	79.57	5	15
Kalam	Papua New Guinea	68	7	25	60.49	7	9
Kamano-Kafe	Papua New Guinea	32	8	25	62.07	8	16
Karajá	Brazil	19	6	19	86.20	2	5
Kemtuik	Indonesia	24	7	25	63.92	6	11
Kokni (Kokoni)	India	11	9	25	61.5	8	17
Konkomba	Ghana	20	5	25	69.07	4	5
Kriol	Australia	26	11	25	67.73	9	18
Kuku-Yalanji	Australia	16	6	20	49.81	5	4
Kuna	Panama	45	11	25	59.05	7	12
Kwerba	Indonesia (Irian Jaya)	22	4	25	71.01	4	6
Lele	Chad	19	5	15	68.72	4	9
Mampruli	Ghana	79	8	24	55.11	5	6
Maring	Papua New Guinea	33	7	25	48.67	6	3
Martu Wangka	Australia	44	6	25	61.34	6	9
Mawchi	India	9	7	25	65.13	7	11
Mayoruna	Peru	4	4	25	82.15	2	2
Mazahua	Mexico	60	16	25	51.04	11	15
Mazateco	Mexico	19	10	25	58.79	9	20
Menye	Papua New Guinea	17	8	25	64.77	6	6
Micmac	Canada	36	13	25	61.73	9	8
Mikasuki	United States	13	8	25	72.98	8	20
Mixteco	Mexico	31	10	25	64.71	9	11
Mundu	Sudan	12	5	18	54.28	4	12
Múra Pirahá	Brazil	4	4	25	82.35	4	9
Murle	Sudan	14	7	25	70.68	7	10
Murrin-phata	Australia	22	8	25	46.72	8	5
Nafaanra	Ghana	4	3	29	87.74	3	2
Nahuatl	Mexico	17	8	6	74.36	6	30
Ocaina	Peru	11	6	25	79.26	5	12
Papago	United States, Mexico	25	9	25	78.15	5	10
Patep	Papua New Guinea	15	6	24	74.52	5	10
Paya	Honduras	13	5	20	73.76	4	6
Podopa	Papua New Guinea	18	7	14	56.02	6	9
Saramaccan	Surinam	44	11	25	58.83	10	13
Seri	Mexico	12	7	25	80.59	5	14
Shipibo	Peru	21	14	25	69.50	4	14
Sirionó	Bolivia	12	5	25	67.61	6	8
Slave	Canada	52	8	24	63.58	7	18
Sursurunga	Papua New Guinea	11	6	26	74.75	5	9
Tabla	Indonesia (Irian Jaya)	7	7	25	71.10	6	10
Tacana	Bolivia	31	13	8	59.45	7	40
Tarahumara (central)	Mexico	21	9	9	67.15	5	26
Tarahumara (western)	Mexico	11	7	6	77.45	5	18
Tboli	Philippines	10	7	25	73.09	6	14
Teribe	Panama	27	8	26	64.28	7	13
Ticuna	Peru	10	6	25	74.65	5	11
Tifal	Papua N. Guinea	25	7	25	41.87	2	4
Tlapanec	Mexico	15	9	25	64.18	8	13
Tucano	Colombia	7	6	25	77.42	5	9
Vagla	Ghana	6	6	25	70.71	5	13
Vasavi	India	10	7	25	65.05	7	13
Waorani (Auca)	Ecuador	18	5	25	63.08	4	6

Walpiri	Australia	62	8	23	44.97	6	5
Wobé	Ivory Coast	6	3	25	88.72	1	4
Yacouba	Ivory Coast	3	3	27	80.79	3	5
Yakan	Philippines	13	12	25	70.54	10	25
Yaminahua	Peru	16	5	25	72.34	5	5
Yucuna	Colombia	8	6	25	75.95	4	13
Yupik	United States	26	6	25	73.39	6	10
Zapotec	Mexico	13	8	25	77.68	7	12

Appendix B

Table of results from the color categorization simulation for all WCS languages

Language	Reliability	Terms		Winners	Activity vectors	
		All	Winning	match	Mean	STD
Abidji	55,01	9	7	53,94	0,774	0,113
Agarabi	53,72	27	10	28,79	0,612	0,149
Casiguran Agta	48,26	70	11	6,06	0,462	0,110
Aguacateco	61,10	24	9	45,76	0,675	0,148
Amarakaeri	73,65	9	7	73,64	0,885	0,131
Ampeeli	52,33	43	7	18,79	0,579	0,132
Amuzgo	67,10	28	12	54,24	0,786	0,128
Angaatiha	61,15	16	7	63,64	0,775	0,151
Apinaye	67,04	12	6	56,97	0,756	0,173
Arabela	58,98	18	7	30,61	0,564	0,138
Bahinemo	65,02	10	7	70,91	0,844	0,097
Bauzi	78,61	8	5	68,18	0,832	0,150
Berik	61,66	29	6	22,42	0,592	0,142
Bete	75,22	22	4	42,73	0,576	0,192
Bhili	66,70	14	8	52,12	0,740	0,148
Buglere	50,12	27	6	26,36	0,561	0,119
Cakchiquel	61,70	28	17	36,36	0,687	0,142
Ucayali Campa	45,82	16	7	34,24	0,668	0,119
Camsa	69,53	31	12	40,61	0,695	0,193
Candoshi	72,35	18	8	46,97	0,751	0,179
Cavine?a	76,36	10	6	71,52	0,881	0,131
Cayapa	67,19	12	7	53,03	0,789	0,175
Chacobo	70,99	5	4	83,94	0,938	0,083
Chavacano	72,72	19	11	61,52	0,799	0,187
Chayahuita	87,64	11	7	88,79	0,954	0,087
Chinanteco	70,43	12	8	73,64	0,879	0,121
Chiquitano	57,04	49	13	29,09	0,628	0,174
Chumburu	63,93	24	9	33,33	0,596	0,152
Cofán	78,72	9	6	80,30	0,908	0,117
Colorado	73,19	6	6	79,09	0,923	0,086
Eastern Cree	71,14	8	7	52,12	0,644	0,394
Culina	66,73	29	4	15,45	0,498	0,137
Didinga	85,58	7	7	72,73	0,818	0,189
Djuka	60,57	22	9	43,94	0,645	0,167
Dyimini	69,59	12	9	50,30	0,739	0,161
Ejagam	86,92	7	3	86,06	0,935	0,133
Ese Ejja	69,75	12	7	58,79	0,804	0,180
Garifuna (Black Carib)	60,28	18	10	46,67	0,659	0,138
Guahibo	54,73	17	10	41,21	0,771	0,146
Guambiano	69,07	8	7	76,36	0,910	0,098
Guarijío	79,33	14	6	71,52	0,831	0,158
Guaymí (Ngäbere)	59,45	30	7	33,64	0,602	0,148
Gunu	53,08	27	7	22,12	0,617	0,151
Halbi	45,95	71	8	11,21	0,469	0,094
Huastec	75,38	12	8	56,67	0,808	0,257
Huave	71,05	17	11	58,48	0,787	0,182
Iduna	69,75	51	6	13,03	0,415	0,166

Ifugao (Keley-i)	63,09	23	8	36,36	0,666	0,153
Iwam (Sepik)	55,60	10	7	66,97	0,743	0,115
Jicaque	79,57	9	5	64,55	0,789	0,209
Kalam	60,49	68	7	11,52	0,464	0,124
Kamano-Kafe	62,07	32	8	19,09	0,540	0,120
Karajá	86,20	19	6	53,33	0,699	0,254
Kemtuik	63,92	24	7	34,24	0,602	0,153
Kokni (Kokoni)	61,50	11	9	57,88	0,758	0,118
Konkomba	69,07	20	5	36,67	0,477	0,131
Kriol	67,73	26	11	61,82	0,770	0,169
Kuku-Yalanji	49,81	16	6	33,03	0,639	0,129
Long-haired Kun	59,05	45	11	31,52	0,624	0,138
Kwerba	71,01	22	4	36,97	0,568	0,200
Lele	68,72	19	5	46,36	0,664	0,198
Mampruli	55,11	79	8	11,21	0,433	0,100
Maring	48,67	33	7	26,67	0,529	0,095
Martu Wangka	61,34	44	6	16,67	0,458	0,121
Mawchi	65,13	9	7	64,24	0,816	0,124
Mayoruna	82,15	4	4	85,15	0,947	0,093
Mazahua	51,04	60	16	16,67	0,620	0,147
Mazateco	58,79	19	10	44,85	0,783	0,173
Menye	64,77	17	8	33,33	0,631	0,143
Micmac	61,73	36	13	42,12	0,672	0,249
Mikasuki	72,98	13	8	53,94	0,806	0,144
Mixteco	64,71	31	10	51,21	0,736	0,149
Mundu	54,28	12	5	46,36	0,670	0,167
Múra Pirahá	82,35	4	4	86,67	0,947	0,099
Murle	70,68	14	7	51,21	0,765	0,174
Murrin-phata	46,72	22	8	28,18	0,518	0,114
Nafaanra	87,74	4	3	80,30	0,875	0,144
Nahuatl	74,36	17	8	66,06	0,834	0,154
Ocaina	79,26	11	6	62,12	0,799	0,179
Papago	78,15	25	9	24,85	0,536	0,212
Patep	74,52	15	6	52,42	0,700	0,203
Paya	73,76	13	5	39,09	0,680	0,192
Podopa	56,02	18	7	37,58	0,685	0,145
Saramaccan	58,83	44	11	16,67	0,509	0,121
Seri	80,59	12	7	80,91	0,913	0,128
Shipibo	69,50	21	14	38,48	0,582	0,150
Sirionó	67,61	12	5	53,33	0,748	0,141
Slave	63,58	52	8	14,85	0,540	0,159
Sursurunga	74,75	11	6	66,97	0,730	0,153
Tabla	71,10	7	7	74,85	0,894	0,094
Tacana	59,45	31	13	28,48	0,630	0,185
Tarahumara (central)	67,15	21	9	29,70	0,619	0,161
Tarahumara (western)	77,45	11	7	76,97	0,855	0,143
Tboli	73,09	10	7	67,27	0,834	0,147
Teribe	64,28	27	8	44,85	0,670	0,183
Ticuna	74,65	10	6	77,27	0,892	0,118
Tifal	41,87	25	7	20,91	0,612	0,128
Tlapanec	64,18	15	9	62,12	0,822	0,122
Tucano	77,42	7	6	65,15	0,838	0,168
Vagla	70,71	6	6	76,36	0,932	0,063
Vasavi	65,05	10	7	63,03	0,776	0,143
Waorani (Auca)	63,08	18	5	25,76	0,641	0,212

Walpiri	44,97	62	8	6,67	0,485	0,104
Wobé	88,72	6	3	73,03	0,857	0,153
Yacouba	80,79	3	3	86,97	0,938	0,079
Yakan	70,54	13	12	76,67	0,875	0,126
Yaminahua	72,34	16	5	64,55	0,701	0,138
Yucuna	75,95	8	6	75,15	0,849	0,154
Yupik	73,39	26	6	25,15	0,553	0,188
Zapotec	77,68	13	8	76,67	0,892	0,143

Appendix C: the software CD

Content of the attached CD:

FOLDER **WCSCategorizer_output**:

FILES: the visualization images from categorization simulation for all WCS languages + Slovak

FOLDER **WCSVisualizer**:

FOLDER: **data**

FOLDER **input**: the whole content of WCS online data archives

FOLDER **stats**: statistic information about WCS languages (appendix A)

FILES: **run.bat, run.sh, WCSVisualizer.jar, readMe.txt**

FILE **Rebrova_thesis09.pdf**: electronic version of this thesis

Software requirements: JAVA 6

To **run** the visualization **application**, go to **WCSVisualizer** folder and execute:

1. on Windows: **run.bat**
2. on Linux or Mac OS: **run.sh**