Online Colour Naming Experiment Using Munsell Samples

Dimitris Mylonas , Lindsay MacDonald ; London College of Communication , University of the Arts ; London , UK

Abstract

An online colour-naming experiment was designed and developed to determine a broad set of colour words in wide cultural use with their corresponding colour ranges in sRGB and Munsell specifications. The quality of the English dataset was analysed in terms of colour centroids, frequency of names, consistency and response time. The importance of basic colour terms was confirmed while it was also revealed that in free colour naming tasks, the majority of the observers preferred to use non-basic colour terms. The validation of the web-based experimental methodology with previous studies conducted in controlled viewing conditions produced satisfactory results, while a comparison of the 27 most frequent chromatic colour words with a previous web-based experiment showed a remarkable agreement.

Introduction

Extensive research in colour naming and colour categorization has raised philosophical, linguistic, anthropological and psychological implications. The main opposing views call into question whether colour categories are formed under the influence of shared panhuman perceptual and cognitive mechanisms on language, or whether language influences thought and perceptual processes [9, 28]. However, recent linguistic and psychological studies [11, 30] have allowed some degree of compromise, in which both sides recognised the influence of language in the cognitive organisation of colours.

Until recently, the widely accepted nativist view proposed by Berlin and Kay [4] influenced a large number of studies [3, 5, 20, 21, 23, 29] to constrain the focus of their research to a small number of eleven consensual colour categories rather than towards the development of more delicate and refined colour discriminations [12]. This paper investigates colour naming and categorisation through an online free colour-naming experiment [25], accordingly is not looking for direct answers to the question of the origin of the colour categories [17]. Instead it is concerned with the mapping between colour stimuli and a broad set of colour words in wide cultural use. In addition this study improves previous web-based research on colour naming by taking advantage of current web technologies and collecting numerous additional parameters to accommodate some of the disadvantages of the online experimental methodology.

The significance of online colour-naming databases is becoming more vital as more people have access to the global network. This trend implies that colour specification is not only the domain of extensively trained colour users, but it is a facility required by a large audience. Finally, as the Internet forces rapid language and cultural exchange - which often leads to language loss - there is a need for systematic documentation and archiving of data on colour categorisation for future scientific research [16].

Online Experimental Methodology

The World Wide Web offers advanced methods of conducting research on colour naming by providing relatively new tools to access easily a large number of observers from culturally and demographically diverse populations in a short period of time. However, the absence of observer monitoring and interactive clarifications often involves higher observer metamerism and higher rates of dropouts. Studies conducted over the Internet offer greater external validity through greater technical and viewing conditions variance. The general applicability of the results is achieved by replacing systematic control error with random distributed error [27]. In addition web-based research allows the investigation of a wide range of software and hardware components that affect the reproduction of the colour stimuli, while it is possible to explore the variability of the appearance of the colour stimuli affected from viewing conditions related parameters. However, given the lack of calibrated display monitors and accurate measurements of the various components of the viewing field, the results of this paper are restricted to assume the sRGB colour space and categorical viewing conditions.

Infrastructure of Online Colour Naming Experiment

The web-based colour naming experiment was designed and developed in Adobe Flash CS4 S.V. and ActionScript 3 following similar methodology with the high entrance barrier technique as described by Reips [27]. The Flash applet was embedded in HTML and it is connected via PHP bridges to a mySQL database that sends the test images in a random order and in return, stores the information for each participant. The communication form is independent from the database and sends directly an email to the account of the author. Various forms of statistical analysis can be performed either directly with mySQL scripts on the server, or in Matlab (Figure 1).

Stimuli Selection

The 600 in total test samples of the colour naming experiment specified in the sRGB colour space [18] were selected from the Munsell Renotation Data [24, 26]. The original dataset consists of 2729 colour samples specified in xyY and viewed against a neutral grey background under illuminant C. Since the achromatic colours are not included in the original dataset, nine neutral samples, one for each Munsell Value and a White and a Black sample at the extremes of the sRGB cube were added. The Bradford chromatic adaptation transform was used to convert the CIEXYZ values defined under illuminant C to CIEXYZ viewed under D65 [13]. The colour samples lying outside the sRGB gamut were discarded in order to avoid over-representation of the RGB extremes in the sample set. The clipping method involved a simple round trip set of transformations, where samples with more than 4 ΔE were assumed to lie outside the sRGB gamut and were discarded [19].

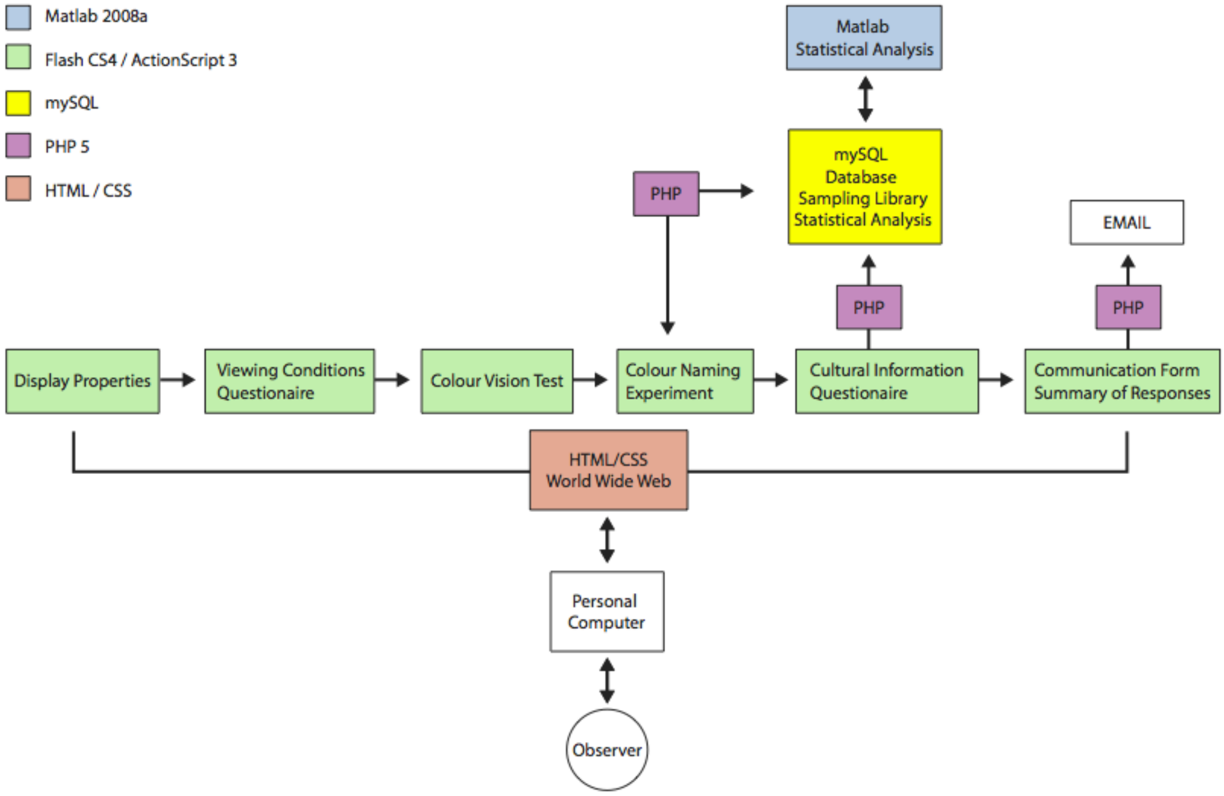

Figure 1. *Flowchart of the online colour naming experiment*

Given the cylindrical structure and the irregular shaped colour solid of the Munsell system, the sub-sampling of the remaining in gamut colour samples followed a similar approach to the advice of Billmeyer to Sturges & Whitfield [29] in order to equalize the perceptual distances between the samples. Specifically, a variable number of hues were sampled at different levels of value and chroma. At chroma 2 - ten hues were sampled, whilst at each successive chroma step the sampled hues were increased by ten. That means from chroma 8 to the boundaries of the sRGB gamut, all 40 hues were sampled. For chroma and value equality in visual spacing a 2:1 ratio was implemented.

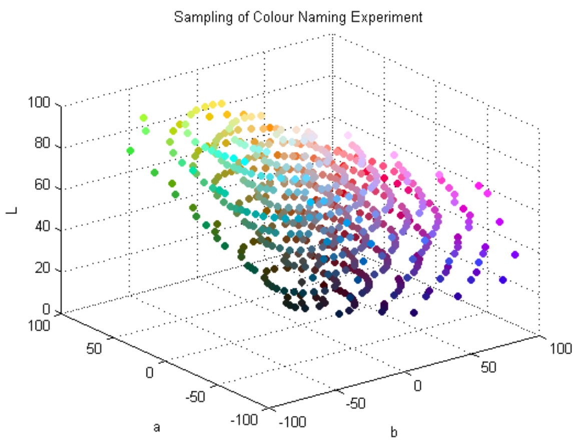

Figure 2. *Sampling of colour naming experiment in sRGB gamut*

Experimental Procedure

The experimental procedure consists of six stages. The first stage presents a greyscale ramp of 21 equally spaced steps where the observers were asked to adjust the levels of contrast and brightness in order to barely see the differences of the first and last steps of the ramp. The second stage consists of a questionnaire on viewing conditions and display properties. The observers are asked to provide information about the display technology, the manufacturer and model (if known), the calibration state, the size of the monitor and their distance from the monitor. Additionally in order to investigate colour naming across different viewing conditions the questionnaire collected verbal descriptions of the lighting conditions, the surroundings and the white point of the display. The third stage hosts a web-based simplified version of the CU Dynamic Colour Vision Test developed by Professor John Barbur [2] at the City University of London [7] to screen the observers for possible colour deficiencies.

The forth stage consists of an unconstrained colour naming experiment that makes use of a similar methodology to Moroney [22] and Dolores Lab [10] of distributed psychophysics. The participants are presented with 20 colour patches from a total of 600 samples in a random order. The size of the stimuli is 147 x 94 pixels, and they are presented one at a time against a neutral grey of L*=51 with a black outline of one pixel thickness. The main task is always visible on the top of the page and asks the participant to name the colour patches with the best representative colour term that they remember. To test the consistency of the free colour naming responses, a sample is repeated twice and both responses are recorded for further analysis. Additionally the observers are informed that their response time will be recorded, and an invisible counter saves the required time for naming each sample in milliseconds in the database.

The fifth stage consists of a cultural information questionnaire in order to investigate any associations between culture, psychophysical channels and linguistic colour spaces. The observers are asked to provide information about the country where they grew up, the country where they are currently resident, the language skills, the level of experience on working with colour, the educational level, the age and finally the gender. The sixth and last stage thanks the participants for their responses and presents a summary of the 20 responses next to the tested colour patches at the same page. Finally an optional communication form is provided for valuable comments and feedback, promoted with a lottery of a limited number of fine art prints from the artist Valero Doval.

All of the data described above is saved in the mySQL database, in a secure Unix server with the time and day of the observation, the IP Address, the operating system (OS), the colour depth and the screen resolution of the display. It is noted that the communication form is separated from the database in order to keep the anonymity of the observations.

Languages

Initially the colour naming experiment was translated into English, Greek and Spanish. According to Berlin & Kay [4], English and Spanish belongs to the top VII stage of colour vocabulary development while the Greek language was studied only through the literature of Homeric Greek and is categorized in the IIIb stage. A Recent study [1] of Modern Greek colour terminology resulted in twelve basic colour terms including two blues, "*γαλάζιο*" (galazio) and "*μπλέ*" (ble), characteristics of at least VII stage.

Subjects

The majority of previous studies involved a small number of observers naming a large number of samples. In contrast the colour naming experiment of this study makes use of distributed psychophysics and presents a small number of colour samples to a large number of observers over the Internet. Taking into consideration recent studies that verified the influence of language in categorical perception for adult humans [11], this paper is interested only in observers over 16 years old, providing colour names in English and possessing normal trichromat vision.

Results

Control Issues of the Web-based Experiment

In the first 3 months the HTML page viewed 3128 times. 1269 visitors decided to start the experiment, and the database hosted a total of multilingual 13.000 colour words from 650 observers resulting a dropout rate of 49%. Multiple submissions were checked through the IP address of each participant with 10% participating twice in English. Given the high rate of participation, and since the observers were not restricted to participate more than once, all the multiple submissions were considered. The raw responses were checked for spelling mistakes, hyphenated, comma separated, and words in parenthesis were treated as multiword colour expressions while the typographic conventions were deleted. The percentage of the disruptive observers totaled 5% in English. The rejected responses included mainly incomplete or numerical responses with a few cases in different languages than those of the experiment. Finally, one set of twenty responses from an observer younger than 16 years old was discarded.

Colour Vision Test

The web-based colour vision test resulted a 17.4% failing to pass the test. It is noted that all the participants were informed to visit their ophthalmologist to diagnose the type and severity of the possible colour deficiency and were allowed to continue the experiment. Since this study is interested only in normal trichromats, all the responses of subjects with positive indication of colour deficiency were rejected, however the colour naming data is available for further investigation.

Validation of the Experimental Methodology

As with any new technological innovation the results of web-based studies need to be validated in order to ensure that the new methods perform what they claim they do. Accordingly the centroids of the basic colour terms were compared with the psychophysically rigorous results of Boynton & Olson [5] and Sturges & Whitfield [29]. Table 1 provides a quantitative comparison in terms of Euclidean colour difference (ΔE) in CIELAB [15].

	B&O vs S&W	**B&O vs M&M**	**S&W vs M&M**
		ΔE (CIELAB)	
Green	17.20	19.79	7.13
Blue	10.94	29.23	24.68
Yellow	16.03	18.36	13.06
Red	15.75	18.77	3.52
Purple	12.78	31.78	19.45
Orange	14.05	10.10	6.15
Pink	12.45	30.52	26.70
Brown	12.17	13.70	12.12
mean	**13.92**	**21.53**	**14.10**

Table 1. *Colour difference of basic colour terms in CIELAB*

The comparison with the results of Sturges & Whitfield resulted in mean ΔE(ab) of 14.1, almost equal to their in-between comparison and smaller than a recent published study conducted in controlled viewing conditions [3]. The different sampling as well as the different illuminant under which the experiment was conducted can explain the relatively larger colour difference of mean ΔE(ab)= 21.23 with Boynton & Olson.

Colour Words in Wide Cultural Use

The refined dataset in the English language resulted in 5428 observations of 1226 unique colour words. The 52% involved mono-lexical responses, 42% two word responses and 6% three or more words. The eleven basic colour terms proposed by Berlin & Kay [4] occurred in 29% while non-basic terms were involved in 23%. Figure 4 shows the colour words relative to their frequency of occurrence.

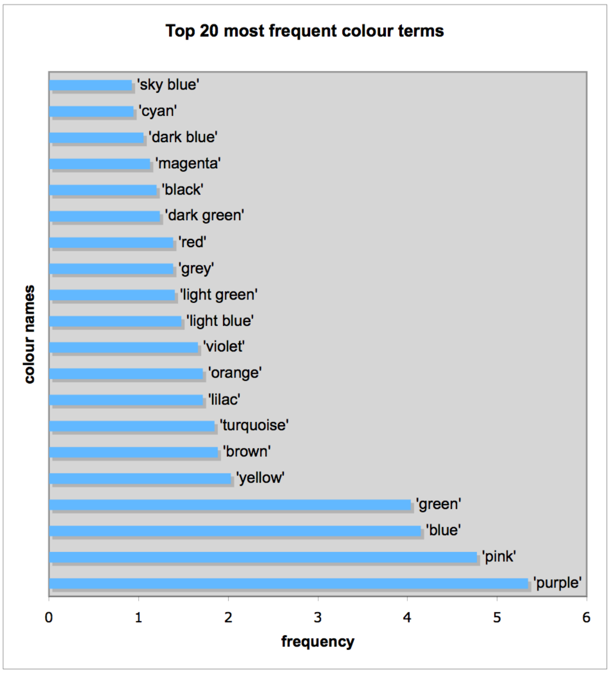

Figure 4. *Top 20 most frequent colour words in English language*

The most frequent colour terms were purple, pink, blue and green while in the top ten also occurred non-basic terms like turquoise, lilac, violet and magenta. From the polylexemic colour terms the most common terms were light blue, light green, dark green and dark blue revealing a preference to use modifiers rather than secondary terms in the blue and green regions.

The 27 top most frequent chromatic colour words were validated against the web-based experiment of Moroney [23] in terms of the location of their centroids in Figure 5.

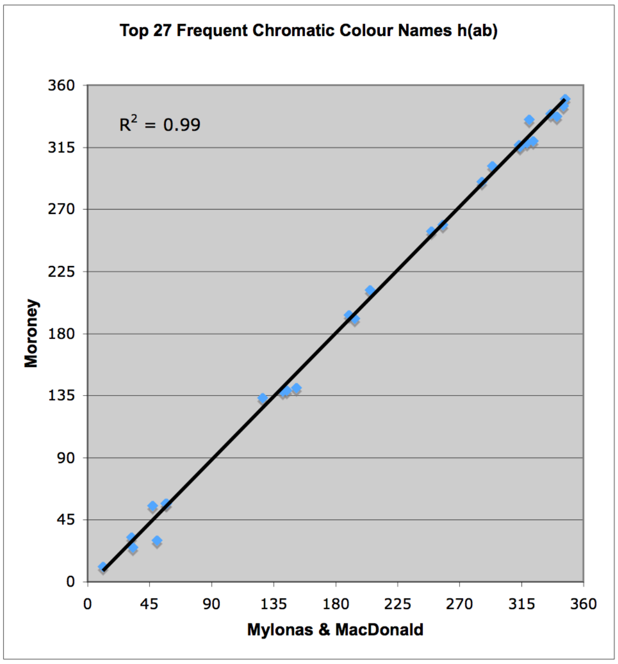

Figure 5. *Comparison of 27 top listed chromatic colour words with the corresponding results of Moroney in terms of hue angles in CIELAB.*

The comparison between the two web-based experiments in terms of hue angles ($h_{(ab)}$) in CIELAB resulted a linear fit of R^2=0.99 For lightness ($L^*_{(ab)}$) the web-based results agree with an R^2 of 0.94 while the agreement for Chroma ($C^*_{(ab)}$) was relatively lower with linear fit of R^2=0.74. Overall the comparison between the two web-based experiments resulted a notable agreement in all dimensions of CIELAB considering that the comparison involved not only the eight chromatic basic colour terms of B&K [23] but also 27 chromatic colour words in wide cultural use.

Consistency

Consistency is a measurement of agreement between two identifications that each observer has made for the same colour samples. In this study the experimental procedure involved the repetition of one colour sample to each observer in order to investigate the consistency of unconstrained colour naming responses. To perform a valid indication two measurements took place following similar methodology to Guest & Laar [15]. The first measured the degree of consistency between the entire colour words used to name the same sample, for example between bright pink and pale pink while the second involved only the main hue components, for instance only pink and pink. The results are shown in Table 2.

Consistency	TRUE %	FALSE %
Entire Colour Name	35.87	64.12
Only Hue Component	66.66	33.33
Total Tested Samples	315	315

Table 2. *Consistency measurements*

The first measurement revealed a low overall consistency value of 35.87% confirming the problem in repeatability of the free colour naming experiments. However the responses were not random, the consistency value of 66.66% obtained by the second measurement upset the first results relatively close to the reported overall consistency of 78.3% and 81.5% in the monolexical constrained experiments of Boynton & Olson and Sturges & Whitfield respectively [29]. It is noted that the observers of this experiment were not informed for the repeated samples.

Response time

Although the distributed psychophysics limited the outcomes of this paper to report consensus colour samples, the response time was recorded in order to provide an indication of ease in recognising and identifying regions of the colour space. It was also aimed at encouraging the observers to provide a natural response without the advice of other reference colour naming systems. Given the absence of observer monitoring, the response time presented in Figure 6 was calculated by averaging the 98% percentile of the response times recorded for each colour name.

The basic colour terms were obviously responded faster than non-basic colour words. Red, blue, white and green were the easiest to name regions of the colour space while the only non-basic colour term in the first 11 top list was teal. In the top 20 list, three colour terms, peach, flesh and tan were occurred to describe the area between pink, yellow, orange and red.

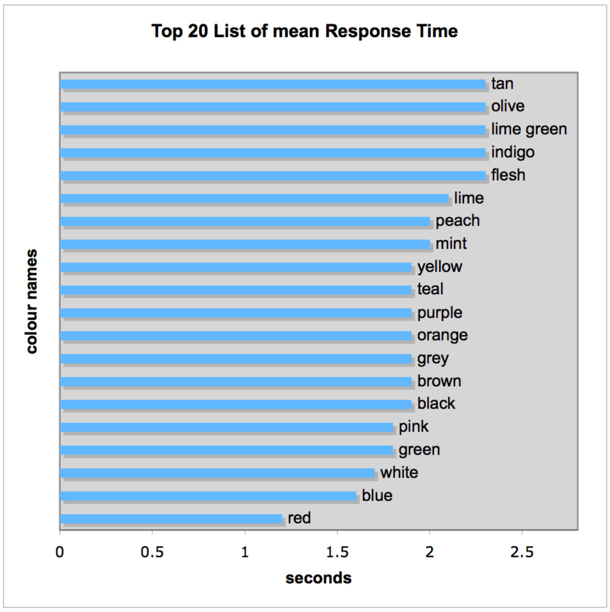

Figure 6. *Top list of response time*

Conclusions

The online experimental methodology provided high quality colour naming data when validated against previous experiments conducted in controlled environments of laboratories. The data analysis confirmed that basic colour terms are used more consistently and were identified more quickly than non-basic colour words. However, it was also revealed that 71% of the responses involved non-basic terms. The comparison of the top 27 most frequent colour words resulted a remarkable agreement and is coherent with a body of existing research, which has suggested the partition of the colour space into 30 or more colour categories [6, 8, 23].

An interesting finding was that a non-opponent hue, like purple is associated with very popular colour words, when the opponent red is following in the 14th position. The overpopulated purple region may imply a possible bias introduced by the clipped Munsell sub sampling in the relative small sRGB gamut, especially at the cyan area.

Given that each colour name is associated with specific categorical viewing conditions in the colour naming dataset, future directions involve the investigation of colour appearance issues. The data analysis of the cultural information questionnaire will provide an insight of how culture is associated with psychophysical channels and linguistic colour spaces and more languages are to be added in the near future. Finally, further research is required to investigate the microstructures of the lexical colour spaces in order to develop a colour-naming model to automate the colour-naming task in web-based environments.

References

[1] Androulaki, †. et al. "Basic colour terms in Modern Greek: Twelve terms including two blues." J. of Greek Linguistics, 7, 3-47 (2006).

[2] J. Barbur, CU Dynamic Colour Vision Test. Proc. Roy. Soc.B 258. pg. 327-334. (1994).

[3] R. Benavente et.al. "A Data Set for Fuzzy Colour Naming." Color research and application, 31(1): 48-56 (2006).

[4] B. Berlin & P. Kay, Basic color terms: their universality and evolution (Stanford, Calif.: Center for the Study of Language and Information 1969).

[5] R.M. Boynton & C.X. Olson, "Locating basic colors in the OSA space." Color Research & Application, 12(2): 94-105. (1987).

[6] A. Chapanis, "Color names for color space." J. American Scientist, 53, 327-346. (1965).

[7] City University London. A new web-based colour vision test, City University London Department of Optometry and Visual Science (2008). Available at: http://www.city.ac.uk/avrc/colourtest.html [Accessed October 1, 2009].

[8] G. Derefeldt & T. Swartling. "Colour concept retrieval by free colour naming. Identification of up to 30 colours without training." Displays, 16(2), 69-77. (1995).

[9] G. Derefeldt et al. "Cognitive color." Color Research & Application, 29(1): 7-19. (2004).

[10] Dolores Lab. Where does "Blue" end and "Red" begin? Available at: http://blog.doloreslabs.com/2008/03/where-does-blue-end-and-red-begin/.
[Accessed November 1, 2008].

[11] A. Franklin et al., Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. Proc. of the National Academy of Sciences, 105(9), pg. 3221-3225. (2008).

[12] J. Gage, Color and culture: practice and meaning from antiquity to abstraction (Boston: Little, Brown and Company 1993) pf 79.

[13] P. Green, Colour Engineering Toolbox, a collection of Matlab functions, UK: London College of Communication.
Available at: http://www.digitalcolour.org/toolbox.htm
[Accessed October 5, 2007].

[14] P. Green "Colorimetry and colour difference" in Green, P. & Macdonald L. Colour Engineering (UK: Wiley 2002)

[15] S. Guest & D.V. Laar, "The structure of colour naming space." Vision Research, 40(7) 723-734. (2000).

[16] C. Hardin, "Closing Thoughts" in Color categories in thought and language (Cambridge; New York: Cambridge University Press 1997) pg 370.

[17] S.R. Harnad, Categorical Perception: The Groundwork of Cognition Reprint (Cambridge University Press 1990).

[18] IEC, Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space – sRGB. IEC 61966-2-1 (1999)

[19] H.R. Kang, Computational Color Technology (SPIE Society of Photo-Optical Instrumentation Engineering 2006).

[20] J. Lammens, A computational model of color perception and color naming. (PHD. State University of NewYork at Buffalo 1994).

[21] H. Lin et al. "A cross-cultural colour-naming study. Part III-A colour-naming model." Color Research & Application, 26(4), 270-277. (2001).

[22] N. Moroney, Unconstrained web-based color naming experiment. Color Imaging VIII: Proc. Hardcopy and Applications. Edited by Eschbach, Reiner; Marcu, Gabriel G. Proc. of the SPIE, 5008, 36-46. (2003).

[23] H. Motomura, "Categorical Color Mapping Using Color-Categorical Weighting Method Part I: Theory." J. Imag. Sci. and Technol., 45, 2. (2001).

[24] Munsell Science Lab, Munsell Renotation data
Available at: http://www.cis.rit.edu/mcsl/online/munsell.php
[Accessed: 10.10.2008].

[25] D.Mylonas, L.MacDonald. An Online Colour Naming Experiment (2008). Available at: http://colournaming.com

[26] S.M. Newhall, D. Nickerson & D.B. Judd, "Final Report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors." J. of the Optical Society of America, 33(7), 385-411. (1943)

[27] U. Reips, "The Web Experiment Method: Advantages, Disadvantages, and Solutions." in Psychological Experiments on the Internet (Academic Press 2000) pg. 89.

[28] J.A. Schirillo, "Tutorial on the importance of color in language and culture." Color Research & Application, 26(3), 179-192. (2001)

[29] J. Sturges & A. Whitfield, "Locating basic colours in the Munsell Space." Color research and application, 20(6), 364-376. (1995).
[30] L.H. Tan et al., Language affects patterns of brain activation associated with perceptual decision. Proc. of the National Academy of Sciences, 105(10), 4004-4009. (2008).

Author Biography

Dimitris Mylonas received his MSc in Digital Colour Imaging from the London College of Communication, University of the Arts (2009). He is currently working as research fellow in applied colour science at the University of Liverpool and TruColour Ltd. His work focus on visual communication design and colour management solutions in web based environments. He is member of IS&T, SDC and Colour Group GB.