

ROBERT PIŁAT

# Colour Names And The Concepts Of Colours

*There is growing body of knowledge about how humans and animals perceive colours; we may safely say that both physiology and physics of colour perception are becoming less and less mysterious. Still it doesn't help to solve a philosophical puzzle: What do exactly mean expressions like "perceived red" or "perceived green"? What do perceived colours refer to in the world? There are three problem fields I am touching on in this paper: (i) semantics of colour names, (ii) ontological status of colours, (iii) cognitive relevance of colours. I am trying to formulate ontological and epistemological assumptions for semantics of colour names. I am especially focused on classical problem of objectivity of colours. While pursuing my task I am making some critical remarks about Wittgenstein's views on colours as formulated in "Tractatus" and modified in "Remarks on Colours". I am using Park's (1983) solution of Hempel's paradox of confirmation by recasting colour names as being about properties of shapes and not objects. I am also discussing briefly a recent theory of colours proposed by Gärdenfors (2000), who takes colours to be regions in certain conceptual space. Finally I turn to some basic notions of possible worlds semantics in hope to obtain a good ground for semantics of colour names.*

## 1. Johann Wolfgang von Goethe and his *Farbenlehre*

Goethe's program of systematic study of colours can hardly be taken seriously today. Goethe searched for strict laws governing colour perception in humans, but he wasn't very strict himself. Besides, the knowledge about the mechanisms of colour perception in his times was rather rudimentary despite great progress, that had been made by discovery of colour circle. Goethe's method of inquiry escapes easy labelling. It is neither analytic nor phenomenological nor psychological. It is rather a mixture of them all like in remark 768 of *Farbenlehre*, to the effect that yellow produces an impression of warm and cosiness and therefore in painting this colour plays big role in lighted, active parts. In just one brief passage Goethe contains: (1)

psychological statement (about producing impressions), (2) a physical statement about the role of yellow in lighted scenes; (3) an aesthetic statement about the art of painting. Even more bizarre is his remark, that purple glass, when one looks through it, shows the landscape in horrifying light, which will most probably be one that will cover earth and sky on the day of the Last Judgement (*Farbenlehre*, no. 798).

If I am nevertheless starting from Goethe's *Farbenlehre* it is to define precisely my approach, which is loosely speaking, phenomenological, so as Goethe's was in his best remarks, like for instance *Farbenlehre*, no. 805 and 806 to the effect that human eye, when stimulated by a colour patch, automatically produces an impression of complementary colour (a pair of such colours embrace the whole circle of colours). It is because – explains Goethe—a perception of single colour provokes in the eye a **desire** for certain generality or totality. In order to satisfy that desire the eye starts to **search** for some colourless region in the neighbourhood of the perceived colour spot and **places** there a complementary colour (words printed bold refer to what very much reminds phenomenological constitution of experience husserlian style). Now, this kind of phenomenological approach was rather distasteful for Wittgenstein, who in several important notes in *Remarks on Colour* argues with Goethe's approach. Neither in *Tractatus* nor in *Philosophical Investigations* would Wittgenstein agree that there might be any phenomenological laws of colours (or of anything else). This, I think, shows the fundamental theoretical tension which will be my starting point.

## 2. Wittgenstein's Logical Theory of Colours

Wittgenstein formulated his view on colours in the context of discussing logical properties of sentences like "This is red and this is green". (Following some commentators let's mark this kind of sentences **A**). The question is: Is sentence **A** contradictory? Following an insightful paper by Allaire (1996) we may speak of three stages of Wittgenstein's coping with this problem.

**First stage:** (*Tractatus*) On thesis 6.3751: "For two colours, e.g. to be at one place in the visual field, is impossible, for it is excluded by the logical structure of colour. (...) (It is clear that the logical product of two elementary propositions can neither be a tautology nor a contradiction. The assertion that a point in the visual field has two different colours at the same time is a contradiction.)". So Wittgenstein's claim is to the effect that:

- (1) Colour predicates are not simples, but they can be defined;
- (2) Sentences like "This is green" do not express elementary propositions (see Allaire 1996, 203).

If neither (1) nor (2) were the case, then sentence **A** would not be contradictory but simply false or nonsensical. By endorsing both (1) and (2) Wittgenstein commits himself to explaining the logical structure that could possibly prohibit these two colours to be at one place at the same time. Such explanation, if seriously taken up, would probably presuppose a translation of both constituent sentences into more complex sentences, the would be free from demonstratives. Only for so translated sentences it could be showed that they cannot be both true. But Wittgen-

stein does not do this job in *Tractatus* and it was Ramsey who pointed out the resulting gap. What Ramsey himself suggests as a solution of the problem is to refer to the law of nature rather than to logic. It is laws of nature that may prevent two sentences to be true at the same time without resorting to pure syntax—namely by presenting two descriptions of the same state such that on certain theory the evidence supporting one description would speak against the other.

**Second stage:** (*Remarks on logical form*) Here Wittgenstein seems to do justice (not quite but close) to Ramsey's criticism of the *Tractatus* position. He says that due to limitations of our logical notation we can not make visible the contradictory character of sentence A. And he goes on to say: "It is, of course, a deficiency of our notation that it does not prevent the formation of such nonsensical constructions and a perfect notation will have to exclude such structures by definite rules of syntax. These will have to tell us that in the case of certain kinds of atomic propositions described in terms of definite symbolic features certain combinations of T's and F's must be left out. Such rules however cannot be laid down until we have actually reached the ultimate analysis of the phenomena in question. This, as we all know, has not yet been achieved" (Wittgenstein 1929, p. 170-71, after Allaire 1996, 205)

Allaire points out that Wittgenstein was never satisfied with the solution given in the *Remarks on Logical Form*, so that view can only be treated as a transition to the position expressed in *Remarks on Colour* and in *Philosophical Investigations*.

**Third stage:** (*Remarks on Colours* and *Philosophical Investigations*) Here Wittgenstein is presenting his final interpretation of sentence A by referring to the concept of language games. He believes, that to know the meaning of "red" amounts to be able to play certain language game. However, the only full fledged example of a language game with colours he gives is some kind of "show-me game" with goes as follows: There is a challenging person X who expect the responding person Y to show what according to X is red. But the person Y does not know what X has in mind. She can find it out by demanding from X: "You show me red", etc. The ability of two persons to go on with the game serves as a criterion of arriving at the meaning of "red". The "show me" game is a very instructing example indeed, but one example is certainly not enough, given the fact that colour terms are used in numerous contexts beside standing with somebody else in front of a colourful object pointing at it.

It seems to me that in *Remarks on Colour* Wittgenstein treats as equivalent four ways of speaking of colours:

- (1) colour phenomena
- (2) colour concepts
- (3) logical rules governing the use of colour terms
- (4) language games establishing the meaning of colour terms

So, firstly, colour phenomena amount to colour concepts: In his critical note to Goethe's *Farbenlehre* Wittgenstein points out that the phenomenological analysis attempted by Goethe is in fact a conceptual analysis. Secondly, our colour concepts can be boiled down to logical rules governing the uses of colour terms. For in-

stance, Wittgenstein assumes the definition of saturated colour to be "the saturated colour X cannot be sometimes brighter sometimes darker than a saturated colour Y" and goes on to say that this definition amounts to saying that it makes no sense to speak that X is sometimes brighter sometimes darker than Y, and this is a logical statement. And thirdly, colour concepts amount to language games according to well known statement, that not knowing how to play a game amounts to not possessing the corresponding concept. I couldn't find any explicit equation of (3) and (4) but it seems to follow quite clearly.

So the final position on the nature of colours taken by Wittgenstein may be summarised in three points:

- (1) Colour terms denote certain complex concepts;
- (2) The relations between colour concepts are determined by logical rules;
- (3) The ultimate source of such rules are language games.

Now, let's consider three objections to Wittgenstein's view:

1. Wittgenstein approvingly quotes German painter Phillip Otto Runge who argues against Goethe that imagining bluish orange has the same contradictory character as speaking about south-west north wind. Runge, however, is wrong when he makes these two impossible descriptions equal. The nonsensical character of the first one is contingent on properties of physical things including our sensory apparatus. We will (probably) never see anything like bluish orange but still we can think of it. But the south-west north wind is conceptually impossible which means that we cannot even think of it. The difference is grounded in the fact that the terms "south" and "north" get an interpretation in certain model. The interpretation is such, that they are both (1) separate and (2) exhausting together certain domain established by the model. Nothing like this is true about colours. There is no domain which blue and orange would fill. The ability to think about bluish orange is based on the ability of keeping in mind two different images, each referring to different domain. Both this ability and the contradictory character of resulting descriptions are contingent on properties of real world (very much like Ramsey wanted) and not on properties of semantic model.

2. There is tension between conceptual character of colours and our inability to set out criteria of colour classification. In order to get such criteria we would have to use some standards. And in fact such standards are being used by paint manufacturers or designers. A special role is played here by pure colours, because every colour is represented as a mixture of them (good example of such standard is the RGB coding of colours for computer displays). But Wittgenstein's argues that pure colours are entirely irrelevant to our perception of colours. According to him they are of no use for painters and might be "at most good for designers". So on this theory, colours are concepts, but there is no orienting points helping to navigate through their conceptual structure. There may even be no conceptual structure at all.

3. In light of above mentioned difficulties it seems logical for Wittgenstein to resort to language games. The advantage of this concept is that it does not commit one to deal with classification problem, nor it implies worry about properties of corresponding semantic models. In particular, it is quite understandable from language game perspective why the concepts of pure colours are so artificial. They

simply do not play any tractable role in our regular language games. Language games serve to describe linguistic practice, they are not semantic theories. For something which is so unimportant for daily practice, like pure colours, there is simply no relevant language games.

Also the shortcomings of the general conception of language games have their consequences for the plausibility of Wittgenstein's theory of colour concepts:

(1) As I have mentioned already, the only example of language game Wittgenstein gives is the "show me" game which is far from satisfactory, although the example itself is correct. It may turn out that the reason of there being few examples of language games with colours is simply the fact that there is no such games.

(2) No matter how much the very concept of game is based on family resemblance, it does not justify such mysterious Wittgenstein's phrases like "[Painter] does not picture colours, he plays certain game". What game?

(3) One cannot formulate any criterion of mastering a language game apart from using a linguistic expression properly. However this is circular, since proper usage is defined as a competent participation in a language game. What does it mean "to master", or "to be competent", or "to be able to play" Wittgenstein does not explain, or I am unaware he does.

### 3. Back to Analytic Theories of Colours

We see it often believed that Wittgenstein's therapeutic arguments puts end to many analytic projects, including conceptual analyses and standard semantic models. I have tried to defend opposite strategy, however: by pointing at shortcomings of Wittgenstein's position on colours I am coming back to untimely abandoned analytic problems concerning colours. The most important of them are:

- (1) Do colour terms denote colour concepts?
- (2) Do colour concepts grasp some properties of colourful things? (if (1))
- (3) What properties (property) of things is grasped by colour concepts? (if (1) and (2))

Let's take a look at typical answer to these questions:

**(i) Physicalism.** This is the most radical of uniform treatments of colours. Here colour terms are considered substitutes for more adequate descriptions of physical properties like light frequency, composition of light beams, molecular structure of light reflecting and light absorbing objects. The shortcomings of physicalism are obvious though:

(a) Physicalism cannot provide any correspondence rules between phenomenal and physical descriptions of colours. There are many regularities of our colour perception that painters are clearly competent to deal with and that do not have any convincing physical interpretation.

(b) Even if the difficulty indicated in (a) was just a matter of refining our physical and physiological theories of colour, there would still be the puzzling fact, that we have long had non-physical commonsensical or phenomenological explanations (still doing fine) without waiting for the physicists. Can the explanatory power of these explanations simply disappear after physics takes over (assuming it ever does)? This question arises, because if physics ever took over it would not be like one theory incorporating another (like Einsteinian theory incorporating Newtonian one), but rather like one theory replacing another.

(c) Physicalism cannot explain how come, that some contents (colours) which are phenomenal substitutes of actual physical properties nevertheless give animals and humans a tool for reliable empirical generalisations.

(d) Well known today is the phenomenon of metamerism (things that reflect light of different wave length nevertheless appear to human subjects as having the same colour). The physiological basis of metamerism is known. The perceived colour depends namely on the proportion of light waves of different lengths rather than on the absolute wave length. However, this does not count as physicalist explanation unless we have purely naturalistic theory of human cognitive apparatus, which we do not have. Taking into account the proportions between three different types retina irritation is a matter of computation and not a matter of physical cause and effects. But the question whether computational effects in our brains can be reduced to physical processes is still open.

The failure of physicalism to meet quite modest epistemological and ontological condition has produced a bunch of alternative interpretations of the nature of colours. Below I comment on just a few of them, thus putting even more difficulties to the already big stock.

**(ii) Theories accepting double meaning of colour names.** Metamerism not only was used as an argument against physicalism about colours but also as a reason to ascribe colour purely subjective experience (Harding, 1990, 1993). It seems indeed that the perception of colours depends much more on the properties of the cognitive apparatus than on the intrinsic property of perceived things. However, Rosenthal (2000) undermines the argument. He says that claiming that colours do not have explanation in terms of physical properties of things does not amount to maintaining that they are subjective. Colours may simply have broader physical basis than any intrinsic properties of things. Later on in this paper I am going to make use of this suggestion. Another argument against jumping to subjectivist conclusion has been advanced by Ross (2000). According to him subjectivism is unable to explain the role colours play in locating objects in physical space.

Rosenthal tries to do justice to all the arguments and finally proposes to ascribe colour names double meaning. On one meaning they refer to internal phenomenal space, on the other they refer to properties of physical things. The solution, however, is highly counterintuitive. There is nothing in commonsensical using of colour names that entitle us to think they have double meanings. Natural languages (unless there are counterexamples, which I am unaware of) treat colours uniformly and in an intentional manner, always attaching colours to certain objects—imaginary or real ones. It means that if there is some duplicity about colours it is

not based on difference between phenomenal space and properties of objects, but rather on the difference between two kinds of objects. In both cases however (and this Rosenthal admits) colours are intentional and spatially located. So there is still reasons to search for a uniform semantics for colour names and a uniform theory of colour perception.

**(iii) Colours as special kinds of concepts.** In the theory of conceptual spaces proposed recently by Gärdenfors (2000) colours are represented as regions in a four-dimensional space with co-ordinates being light frequencies. The whole colour space is limited by values corresponding to RED, GREEN, BLUE, BLACK, WHITE. According to Gärdenfors there are two kinds of conceptual spaces, the first one has co-ordinates that do not show themselves independently in experience. Regions in such spaces stand for **properties** and colours are among them. This in a sense do justice to Wittgenstein's remark about irrelevance of pure colours for our experience. The other kind of spaces has independently manifestable dimensions, and the regions in such spaces stand for **concepts** in proper sense.

Gärdenfors theory is formally interesting, because it gives properties and concepts uniform interpretation. However, the theory has some similarities to physicalism as it builds the abstract spaces for colours on the basis of physical properties of environment (light wave lengths), while neglecting the properties of physiological mechanisms. Well, not quite neglecting, since the functioning of perceptual mechanisms is in a sense incorporated into the construction of conceptual spaces, such features like the number of dimensions, the size and shape of regions, etc. are certainly contingent on human (or animal) physiology. And this is no doubt a merit of Gärdenfors theory. My worry is that what we have two sets of properties here: one related to nervous system and another related to perceived things and that there is no way to treat them uniformly as an ontological basis for colours. Since Gärdenfors is more interested in solving cognitive problems, he doesn't need to worry about ontology. His theory works does well anyway (there is more and more good research being made within the framework proposed by Gärdenfors). I shall be arguing, however, for incorporating the properties of physiological apparatus into extended physical basis of colours instead of mapping colours onto properties of abstract conceptual spaces.

**(iv) Prototypic experiences of colours.** The Polish linguist Wierzbicka (1999) takes another path in explaining the nature of colour, namely by invoking numerous linguistic data to the effect of situational prototypes being sources of meanings of colour names. The following reasoning is typical for this author: She starts quoting Wittgenstein who wonders why is it so that of all colours only white cannot be transparent. Wierzbicka offers an explanation: "The thought, that white is non-transparent thus being an barrier for vision is consistent with the image of snow covering the ground and keeping it from sight" (1999, 422). In much the same way Wierzbicka explains the prototypical basis of other colours.

Her strategy is not convincing. No matter how accurate and linguistically grounded are the symbolic links she draws, we do not arrive at any theory of colours. Even if Wierzbicka tells us important things about how we got the colour

names we are using, she does not offer any account for conceptualisation of colour. The range and cognitive relevance of the concept blue, for instance, certainly does not depend on an association between the name "blue" and the look of a cloudless sky.

**(v) Colours as Berkeleyan ideas.** In her book on perception Park (1983) advances arguments for reviving Berkeley's interpretation of colours as ideas in mind. She claims that this untimely abandoned conceptualisation of colours provides satisfactory solution to Hempel's paradox of confirmation. The reasoning well known to philosophers of science and epistemologists goes like this: Let's take a sentence "All ravens are black", The sentence has the same truth conditions as sentences "For all x, if x is a raven, then x is black" or "If x is not black, then x is not raven". So the class of black ravens is compatible with classes of non-black non-ravens, black non-ravens, etc. So any experience telling us something about non-black non-ravens may count as part of confirmation of the sentence "All ravens are black". This is highly counterintuitive and shows a major trap of empirical generalisation.

However, Park argues, the conclusion might be avoided if we agreed to treat black as a property of ravens' shapes and not of ravens. Having detached colours from colour objects Park leads us back to Berkeley's theory. Colours are to be perceptual ideas, which get attached to perceived shapes and disappear when the perception is gone. It is quite different from talking about colourful objects or a paint in an artist's box. Here we are looking for some complex physical properties of things that may guarantee that there will be ideas of some kind in our perception, say green. Obviously, the argument goes, the paint manufacturers do not produce any green or red but only substances of certain chemical properties that produce desirable ideas in human minds.

Why not to be happy with this solution? There are at least two reasons:

(a) Park presupposes that there are two separate mechanism: one for detecting shapes and the other detecting colours. She ignores the fact, however, that it is mainly colours (apart from brightness) that help to discern shapes.

(b) Park allows for ambiguity of colour names which is a position I have already argued against in my criticism of Rosenthal position.

#### 4. A Perspective of Semantics for Colour Names

Maybe the source of problems is elsewhere and comes from typical shortcomings of extensional semantics, which underlies the proposals discussed above. It is the difficulty of giving the colour names extensional interpretation that has lead some philosophers to allow for double meaning of colour names. There seem to be a natural way out, which is allowing for an intensional semantics for colour names—possible worlds semantics (PW semantic) for example.

As we know, in order to establish such a semantics we need to assume:

(1) two classes of elementary objects: possible worlds and their inhabitants;

(2) Two functions, one mapping objects onto sets of possible worlds, the other mapping possible worlds onto classes of objects. The first function gives us the meaning of an expressions like “This tree is small”, the other gives us the extension of the constituent expression “small tree”.

This structure can be applied to expressions like “This apple is yellow”, but there are some discouraging features of possible worlds semantics that makes one doubt if this is the right path. The objection has been famously raised by Putnam in his construction of artificial properties *cat\** and *mat\** and showing that the sentences “The *cat\** is on the *mat\**” may have —on possible worlds semantics—the same meaning as “The cat is on the mat”. Putnam's conclusion is that PW semantics allows for too many properties. Should we then abandon this semantics? I do not think so, although arguments I can offer are rather remarks about inconveniences of Putnam's position than a defence of PW semantics.

(1) What does it mean “too many properties”? Isn't it so, that our categorisations of the world objects are always somehow superfluous? Doesn't it simply reflect the fact that the universe of our concepts is open? The reason of the abundance and openness is that we do not know if properties like *cat\** and *mat\** will or will not present some epistemic value in the future, although they seem superfluous now.

(2) There is tricky part in Putnam's reasoning, because for his new definitions he uses descriptions of certain situations. But we do not define ordinary cats and mats by describing situations! So Putnam assumes that “cat” and “*cat\**” are not synonymous. But are they really? We do not know, because there is no equivalence of definitions.

(3) Putnam's resorting to situations in definitions are compatible with his own referential semantics. Putnam assumes something like transferring situational references to other speakers. But doing this we are never free from using conceptual constructs. We cannot do without them, because nobody can see a tiger or yellow colour in exactly the same situations I did. Situations are not identical but **similar** and this similarity must be represented in communication by means of certain conceptual framework.

So if the perspectives for PW semantics are still open we may at least try some suitable amendments. Let's introduce two notions:

(1) **Second order differences between objects**. Let's imagine a scene on the seashore. It is so misty that one cannot tell see from sky by their colours. “Look”—I say—“The see ad the sky are merging”. What is my remark based on? I certainly don't first see separately the see and the sky and then see them merging. I obviously resort to some previous knowledge and I make judgement about what **should** be there and is not. If I hadn't been prepared to see the line dividing the see and the sky, I wouldn't have been surprised. So the expectation of colour difference was based on knowledge of some other difference. Now, let's reverse the reasoning: Maybe colours are to reflect and amplify **other**, probably more important differences. On this interpretation colours would be such differences between objects that can symbolise or amplify other sets of intrinsic, but not so easily accessible (more cognitive resources have to be used) differences. This might be responsible for the fact that colours are always intentional—object oriented—no matter how interesting the interplay between the colours themselves is (some green spot may

look different when a yellow patch is placed next to it, but we do not say that the new green is a property of a couple green-yellow or of some other complex entity; we keep saying that there is a green spot—now slightly changed—in front of us).

(2) **Property basis.** Objects exemplify properties and properties are **of** objects, but it does not mean that every property is ascribed to an object on the same ontological basis. The basis can be broader or narrower. The scope can be expressed in terms of relational properties, because obviously colours depend on number of relations, like ratio of constituent wave lengths, neighbouring colours, texture, etc. Now some properties, which depend on a perceiver, like “tall”, are given semantic interpretation in PW semantics. The meaning of corresponding sentences are interpreted as sets of possible worlds. However for more complicated property bases we may need to assume more complicated language and allow for sets of sets of possible worlds (sets of possible worlds for the constituent relational property and sets of sets for the resulting properties—colours in this case).

**To conclude:** What I am suggesting implies certain amendment to standard PW semantics. The question if such idea is productive, formally feasible and economical remains open. What I wanted to show is that there are names for some properties like colours, that need special semantics treatment, because of their wide ontological basis. Whether this special treatment means a modified PW semantics or some other formalism is naturally opened to discussion. In particular, there seem to be certain trade-off between semantic and ontological goal. On the one hand there are semantic reasons to assume wider ontological basis for some properties. In order to give colour names uniform interpretation we are forced to allow for new modalities to be described as sets of sets of possible worlds. On the other hand such solution has very unpleasant ontological cost, namely, it implies the existence of a new kind of higher order objects (this consequence was suggested to me by Vincent Hendricks). It is not certain if we can really deal with such ontology. The situation of conceptual trade-off is certainly not fortunate but it seems inevitable unless we find some ingenious escape from object-property-word conceptual schema for both semantics and ontology.

## References

1. Allaire, E. B., (1996) Tractatus 6.3751, w: "LW. Critical Assessment" vol.1 *From the Notebooks to Philosophical Grammar. The Construction and Dismantling of the Tractatus*, ed. S. Shanker, London.
2. Austin, J., (1996) Wittgenstein's Solutions to the Color Exclusion Problem, "LW. Critical Assessment" vol.1 *From the Notebooks to Philosophical Grammar. The Construction and Dismantling of the Tractatus*, ed. S. Shanker, London.
3. Goethe, J. W. (1981), *Wybór pism estetycznych*, red. W. Namowicz, Warszawa.
4. Hardin, C. L., (1990) Color and Illusion, in: *Mind and Cognition: A Reader*, ed. W. Lycan, Oxford and Cambridge, Mass.
5. Hardin, C. L., (1993) *Color for Philosophers: Unweaving the Rainbow*, Indianapolis.
6. Park, D. (1983) *The Problems of Perception*, Oxford.
7. Roosenthal, D. (2000) Color, Mental Location and the Visual Field, *Consciousness and Cognition*, IX.

8. Ross, P., (2000) The Location Problem for Subjectivism, *Consciousness and Cognition* IX.
9. Wierzbicka, A. (1999), Znaczenie nazw kolorów i uniwersalia widzenia, w: A. Wierzbicka, *Język – umysł – kultura*, Warszawa.
10. Wittgenstein, L. (1974) *Tractatus logico-philosophicus*, London.
11. Wittgenstein, L. (1977) *Bemerkungen über die Farben / Remarks on colours*, Basil Blackwell.

---

Visit us: <http://www.itb-forum.pl>