

# Unweaving The Lexical Rainbow:

## Grounding Linguistic Creativity in Perceptual Semantics

Tony Veale<sup>1</sup>, Khalid Al-Najjar<sup>2</sup>

<sup>1</sup>School of Computer Science and Informatics, University College Dublin, Belfield D4, Ireland.

<sup>2</sup>School of Computer Science, University of Helsinki, Finland.

### Abstract

The challenge of linguistic creativity is to use words in a way that is novel and striking and even whimsical, to convey meanings that remain stubbornly grounded in the very same world of familiar experiences as serves to anchor the most literal and unimaginative language. The challenge remains unmet by systems that merely shuttle or arrange words to achieve novel arrangements without concern as to how those arrangements are to spur the processes of meaning construction in a reader. In this paper we explore a problem of lexical invention that cannot be solved without an explicit model of the perceptual grounding of language: the invention of apt new names for colours. To solve this problem we shall call upon the notion of a *linguistic readymade*, a phrase that is wrenched from its original context of use to be given new meaning and new resonance in new settings. To ensure that our linguistic readymades, which owe a great deal to Marcel Duchamp's notion of *found art*, are anchored in a consensus model of perception, we introduce the notion of a lexicalized colour stereotype.

### Call me but [X], and I'll be new baptized

*What's in a name? that which we call a rose  
By any other name would smell as sweet;*

-- Juliet, in William Shakespeare's *Romeo and Juliet*

Shakespeare wrote that a rose by any other name would smell just as sweet. From a chemical perspective he was certainly correct: a rose retains all of its olfactory qualities no matter what we choose to call it. Yet as a talented poet, Shakespeare often exploited the power of words to evoke fond memories, to arouse the imaginations and to stir the emotions of his audience. It is certainly true that the word "*rose*" obtains its warm associations and poetic resonance from its perceptual qualities – its deep red color, silky texture and sweet fragrance – but it is surely just as true that this flower would not be so beloved of poets if its established name were a lexical eyesore like "*goreweed*", "*bloodwort*", "*thorn gore*," "*prickstem*" or "*turdblossom*."

Names are important. We choose them not just to serve as unique identifiers, but as evocative signs that are more than mere symbols. Steve Jobs chose the name "*Apple*"

for his new technology venture to exploit the wholesome familiarity of its conventional meaning, a ubiquitous fruit that is seen as natural, attractive and unthreatening. *Apple Corp.* continues to make good use of this naming motif in its products, ranging from the Apple GS (nicknamed the *Granny Smith*) to the Apple *Macintosh* (a type of apple) to the Apple *Newton* (referencing both the popular myth of Isaac Newton and the falling apple that inspired him, and a fruit-filled cookie that is popular with children). The technology company *Sun Microsystems* chose its name to be a signifier of light, solidity and power, while *Oracle* chose its name to evoke all that is wise and knowledgeable. *Cisco* is evocative of the freedoms one associates with the company's home city, San Francisco, while *Google* has benefited from seeing its name go from being a noun (a static thing) to a verb (a dynamic action). A good name cannot save a bad product, but it can help to make a good product great. Conversely, a poor choice of name can only add to the woes of a weak product. Though there are surely many reasons for the failure of Microsoft's "*Zune*", the fact that so many who care to speak of it can only remember the product as *Microsoft's answer to the iPod* suggests that its name was a big part of the problem.

We also use names to divide up the colour spectrum into shareable bundles of perceptual experiences. We all know what is meant by the words "*red*" or "*green*" but we also appreciate that such simple names subsume a wealth of possible tones and tints. Insofar as each color variant has its own uses, it deserves its own name. The [Pantone](#) company, a provider of colour palettes to industry, uses functional alphanumeric names for its many variations. Poets are more evocative, and anchor their chosen names in our shared experiences of a shared physical world. So when, in the *Iliad*, Homer describes the colour of morning light with the epithet *rosy-fingered dawn*, he succeeds in conveying a very specific shade of red by grounding his description in the familiar colour stereotype of the *rose*. A lexical stereotype is any lexicalized idea that can evoke a range of qualities, perceptual or otherwise. But one must be careful when using such dense descriptors. Homer's frequent use of the epithet "*wine-dark sea*" has led many a scholar to the edge of rational explanation, to question not just Homer's visual sense (he is traditionally believed

to have been blind, if indeed he was a single individual), but also ancient nautical conditions (e.g. to posit red tides, dense with rust-hued algae) and even the colour of ancient Greek wine (dark blue, perhaps, if heavily diluted with alkaline water). Yet the simplest answer is that which does not ask us to question our colour stereotypes: Homer really did mean to imply that the sea – at dusk, under an auspicious red sky – looked as dark and *red* as red wine.

With creativity we aim to be fresh and original, yet it is familiarity that lies at the heart of creativity. Conversely, it is obviousness, not familiarity, that is the antithesis of creativity, for to be creative one must knowingly exploit familiar ideas in non-obvious ways. Indeed, psychologists have long argued that a grounding in familiar stereotypes should guide the appreciation of new ideas, leading Giora *et al.* (2004) to advance, and empirically verify, the theory of *Optimal Innovation*. This theory argues that novelty is, in itself, neither sufficient for creativity nor a reliable benchmark of creativity. For Giora, an optimal innovation is any novel turn that contains the recognizable seeds of its familiar origins, as when a witty phrase is seen as a clever variation on a familiar expression, or a novel name can be decomposed into familiar elements. A colour name such as *Jealous Monster*, for a shade of green, would be an optimal innovation in this sense if it is appreciated as a variation on Shakespeare's *Green ey'd monster, jealousy*. So too are technology names that knowingly borrow – in the fashion of *Apple Corp.* – from the world of fruit. Thus, *BlackBerry* and the *Raspberry Pi* each nod to *Apple Corp.* while emphasizing their berry-like *petiteness*.

For a modern connoisseur of colours and colour names, a paintshop catalogue proves to be a more diverse source of evocative names than a book of verse. After all, paint manufacturers have a vested interest in selling more than emulsified RGB codes. So like poets, paint makers craft names that are dense in emotion and poetic resonance, to sell an entire colour “experience” to aspirational buyers. Why else name a paint colour *Soho Loft* or *Eton Mist*? The colour spectrum is free, and available to anyone with eyes, while paint makers all have access to much the same technologies. But names add value that can make a colour desirable, allowing manufacturers to sell feelings in a can. Paint catalogues are thus filled with colour names such as *Mocha Cream, Oyster Shell, Harvest Sun, Toffee Crunch, Vintage Plum* and *Almond Butter*, each a name that can stir the appetite as much as the imagination. Paint makers compete to find the most marketable names for what are virtually the same RGB codes, so that one maker's *Pale Liqueur* is another's *Baked Biscotti* or *Crème Caramel*.

Our colour preferences serve as superficial expressions of deeper personality traits, or at least we feel this to be so when we stake out claims to favorite colours or ask others about theirs. On Twitter, an automated bot that generates a random RGB code and a corresponding colour swatch every hour has attracted almost 30,000 human followers. The outputs of this Twitterbot, named *@everycolorbot*, are frequently favorited and re-tweeted, not because users are drawn to specific RGB hexcodes, but because of what

the corresponding colours say about their own aesthetics. Similarly, the website [colourlovers.com](http://colourlovers.com) invites its users to express their *loves* for (i.e., to vote for) specific colours and RGB codes. Users of the site may also invent their own names for specific codes, and cluster these codes into recommended palettes. Rather like a vast paint catalogue, the site is a trove of insightful data on the creative naming strategies we humans use to lexicalize our favorite hues.

In this paper we seek to automate the creative task of inventing new names for specific colours and RGB codes. The task is interesting not just because humans find it so, or because name invention is a creative industry in itself; rather, the task interests us here primarily because it offers us a framework to explore issues of perceptual grounding in linguistic creativity. Much like *@everycolorbot*, our solution is implemented as an autonomous bot on Twitter. Yet this new Twitterbot is not a mere generator of random RGB codes, but an inventor of meaningful, perceptually-grounded names for its chosen colours. These names are grounded via a large inventory of colour stereotypes, and this database of stereotypes constitutes a reusable result of this research that we make available to others. To ensure that all names are semantically and syntactically well-formed as linguistic constructs, we also exploit the notion of a *linguistic readymade*, a Duchampian idea in art in which something – a physical object or even a phrase – is taken from its conventional context of use and placed in a new context that gives it new meaning and new relevance.

### **The memory be green, and that it us befitted**

There is both a science and an art to creative naming (see Keller, 2003), for though we want our new names to seem effortlessly apt, their creation often requires considerable amounts of search, filtering, evaluation and refinement. So while inspiration can arise from almost any source, a small number of reliable generative strategies dominate. Punning, for instance, is popular as a naming strategy for non-essential services or products that exude informality. Puns thus proliferate in the names of pet shops and pet services (e.g., *Indiana Bones and the Temple of Groom, Hairy Pop-Ins*), hair salons (*Curl Up & Dye*), casual food emporia (*Thai Me Up, Jurassic Pork, I Feel Like Crêpe, Custard's Last Stand, Tequila Mockingbird*) or any small business that relies on a memorable hook to direct future footfall (*Lawn Order, Sew It Seams, Sofa So Good*). As innovations, punning names are optimal in the sense of Giora *et al.* (2004), insofar as they ground themselves in the cozy familiarity of an idiom (“*so far, so good*”) or a popular TV show (“*Law and Order*”) or a film (“*Indiana Jones and the Temple of Doom*”) and give their audience the thrill of recognition when first they encounter them. Computational creativity has had notable successes with punning (Binsted and Ritchie, 1997; Hempelmann, 2008), leading Özbal and Strappava (2012) to obtain promising results for a pun-based automated naming system. With tongue placed firmly in anesthetized cheek, these authors suggest that the punning name *Fatal Extraction* might be used to add humour to a dentist's advertisement, or that a

vendor of cruise holidays might find use for a slogan like *Tomorrow is Another Bay* (though not *Die Another Bay*).

Newly invented names may often take the form of new words, or *neologisms*. One especially productive strategy for neologism creation is the *portmanteau* word, or formal blend, in which a new word is stitched together from the lexical clippings of two others. A good *Frankenword* (the word is itself a portmanteau of “Frankenstein” + “word”) will contain identifiable components of both ingredients, as in “*spork*” (“spoon”+“fork”), “*brunch*” (“breakfast” + “lunch”) or “*digerati*” (“digital”+ “literati”). Veale (2006) presents an automated approach to harvesting neologistic portmanteaux from Wikipedia and for assigning plausible interpretations using the site’s link topology. For instance, as the *Feminazi* Wikipedia page links to that of *feminist* and *Nazi*, and each denotes a kind of person, a “*Feminazi*” is assumed to be a formal blend of a *feminist* and a *Nazi*. Butnariu and Veale (2006) later describe a system, named *Gastronaut*, that invents and evaluates its own neologistic portmanteaux, by combining morphemes of Greek origin (e.g. “*gastro-*”, “*-naut*”) to which it assigns lexical glosses (e.g. “*gastro-*”→*food*, “*-naut*”→*traveller|explorer*). As this system can propose a phrasal gloss for each portmanteau it invents (e.g. proposing “*food traveller*” for *gastronaut*), it uses the presence of this phrase on the Web to validate the linguistic usefulness of the corresponding neologism.

Özbal and Strappava (2012) use a portmanteau strategy to propose salient names for products and their qualities; e.g., their system proposes “*Televisun*” for an extra-bright television, as *sun* is an oft-used stereotype for brightness. Smith *et al.* (2014) present a semi-automatic collaborative portmanteau creator, called *Nehovah*, that uses synonyms of the input words in its formal blends, as well as relevant phrases gleaned from sites such as [www.thetoptens.com](http://www.thetoptens.com). This diversity of lexical sources allows *Nehovah* to invent portmanteau words that do not contain clippings from *any* of its inputs, but to clip words that are nonetheless salient. Özbal and Strappava also use word associations in their formal blends, to propose names such as *Eat*alian (“Eat” + “Italian”) and *Pastarant* (“Pasta” + “Restaurant”) for Italian eateries, the first of which names a real restaurant.

Creative naming, like modern art, is often a matter of wholesale appropriation: we reuse an existing product that is not itself original, but use it in a new context that makes it fresh again. Consider the name *Fifty Shades of Grey* for a hair salon that aims to imbue dye jobs with sex appeal, or the name *The Master and Margherita* for a pizzeria. The movie *The Usual Suspects* takes its striking title from an immensely quotable line from the movie *Casablanca*, the film *Pretty Woman* takes its title from a song by Roy Orbison, while the movie *American Pie* is named after a song by Don McLean. Veale (2012) refers to this kind of appropriation as a *linguistic readymade*, after the *found art* movement launched by Marcel Duchamp in 1917 with his *Fountain* – a signed urinal exhibited as a work of art.

Veale (2011,2012) generalizes this approach to creative text appropriation into a computational paradigm named CIR: *Creative Information Retrieval*. CIR is based on the

observation that much of what is deemed creative in language is either a wholesale reuse of existing linguistic forms – *linguistic readymades* – or a coherent patchwork of modified readymades. CIR provides a non-literal query language to permit creative systems to retrieve suitable readymades with appropriate meanings from a corpus of text fragments such as the Google n-grams (Brants and Franz, 2006). For example, the CIR query operator *@Adj* matches any word/idea that is stereotypically associated with the property *Adj*, and so the query “*@cold @cold*” retrieves bigrams whose first and second words denote a stereotype of coldness, such as “*robot fish*” or “*January snow*”. The retrieved phrases may never have been used figuratively in their original contexts of use, but they can now be re-used to evocatively convey coldness in novel witticisms, similes and epithets. Veale (2012) uses CIR as a flexible middleware layer in a robust model of affective metaphor interpretation and generation that also combines metaphors to generate poetry. Veale (2012) uses CIR in a generative model of irony, to invent ironic similes such as “*as threatening as a wet whisper*” and “*as strong as a cardboard tank*”). A key advantage of using linguistic readymades for automated invention – perhaps the single biggest reason to exploit readymades – is that, as phrases, their syntactic and semantic well-formedness has already been well-attested in the outputs of human authors.

We exploit CIR middleware here as a means of finding readymade colour names in the Google n-grams. That is, we seek out attested phrases that may evocatively suggest a colour, regardless of whether these phrases were ever used to name a colour in any of their original contexts of use (which, of course, an n-gram model cannot tell us). We use a large inventory of lexicalized colour stereotypes to permit CIR to find these candidate phrases, and employ a mapping from stereotypes to RGB hexcodes to derive a composite colour from their individual colour ingredients. Having established a mapping from colour readymades to colour codes, a perceptual Twitterbot can then creatively name the colours it wishes to showcase in its tweets.

## If Snow Be White

CIR offers users a range of non-literal query operators, of which *@* is perhaps the most useful for metaphor retrieval but also the most knowledge-dependent. For *@* is only as useful as its stock of stereotypical associations – such as that *fridges*, *winter*, *fish* and *ice* are each *cold* or that *suns*, *flames*, *ovens* and *deserts* are all *hot* – will allow. Veale (2013) outlines a semi-automated approach to acquiring these associations from similes found on the Web, such as “*as hot an oven*” and “*as cold as winter*”. While a number of these similes identify popular colour stereotypes, such as that lemons are yellow (“*as yellow as a lemon*”), night is black, grass is green and snow is almost always white, we require a considerably more substantial inventory of colour stereotypes if we are going to extract a diversity of readymade colour names from the Google n-grams.

Basic colour words like “*red*” and “*blue*” are often used as simple, descriptive adjectives, while more subtle hues

call for longer adjectival forms. For example, hyphenated compounds, such as “*cherry-red*” and “*nut-brown*”, are commonplace in English and easily harvested from Web texts or from large databases of Web n-grams. Consider the following matches for the CIR query “*^noun - red*” in the Google 3-grams (*^noun* matches any noun):

<i>blood</i>	- red	(3-gram frequency: 57,932)
<i>ruby</i>	- red	(3-gram frequency: 16,366)
<i>cherry</i>	- red	(3-gram frequency: 15,667)
<i>rose</i>	- red	(3-gram frequency: 14,513)
<i>brick</i>	- red	(3-gram frequency: 11,676)
<i>flame</i>	- red	(3-gram frequency: 2,874)
<i>coral</i>	- red	(3-gram frequency: 2,371)

Each of the nouns in the modifier-first position above denotes a familiar stereotype of *redness*. But the 3-grams also provide problematic matches, such as the following:

<i>tallahassee</i>	- red	(3-gram frequency: 172,082)
<i>lemon</i>	- red	(3-gram frequency: 5,486)
<i>mahogany</i>	- red	(3-gram frequency: 1,029)

*Tallahassee*, a place name, does not denote a stereotype of redness in the same way as e.g., the place name *Mars*. Rather, it is a conventionalized name for a specific shade of red, while *lemons* have no association at all with *red* in the popular imagination. *Lemon-red* most likely denotes a blend then, of *red* and *lemon-yellow*, rather than the name of a stereotypical source of redness. It takes knowledge of the world to distinguish such n-grams – undesirable *near misses* – from the desirable *hits* of earlier n-gram matches.

We broaden our n-gram retrieval net by using the CIR query “*^noun - ^colour*”, where *^noun* matches any noun and where *^colour* matches any member of the set {*red, blue, green, yellow, orange, brown, purple, black, white, grey, pink*}. To keep the hits, such as *coral-red*, and to discard the misses, such as *lemon-red*, we must manually filter all retrieved matches. Since our aim is to construct a high-quality resource with extensive reuse value, manual filtering is a good investment of effort. We think it better to construct a perfect resource with manual effort than to design a one-off machine learning algorithm that would do the job imperfectly yet take longer to implement and test. A day of manual effort yields a filtered set of 801 compound adjectives, ranging from *acid-green* to *zinc-white* with hues such as *sulfur-yellow, tandoori-red* and *whale-blue* in between. But a more arduous task awaits.

We must now assign a representative RGB code to each colour stereotype. For instance, we assign #E53134 to *tandoori-red* but #FD5E53 to *sunset-red*. This mapping of colour stereotypes to colour hexcodes provides the perceptual grounding for each stereotype and so must be performed with great care. The [encycolorpedia](#) website and others are used to explore possible RGB codes for each stereotype, and human judgment is used in each case in the selection of the most apt colour code. We use RGB as a coding system for its popularity and simplicity, as

RGB codes can later be converted into one’s preferred coding scheme, such as LAB (see Hunter, 1948), whose dimensions offer a better of model of human perception. The result of this manual effort is a map that associates each of our 801 colour stereotypes with an apt RGB code.

## And summer's green all girded up in sheaves

These lexicalized stereotypes are the building blocks with which we can build novel colour names. Conversely, they are the identifiable signifiers of colour that we can use to recognize the potential of arbitrary readymades to suggest and name specific colours. As noted earlier, we choose to view the invention of colour names as a *readymade art* task, in which coherent, existing phrases are ripped from their original contexts of use – where they are unlikely to name a colour – and given new life as apt colour names.

For CIR purposes, we construct the ad-hoc set *^stereo* to hold the names of all of our colour stereotypes, from *acid* to *zucchini*. The simple CIR query “*^stereo ^stereo*” can now retrieve all bigram phrases from the Google n-grams in which both modifier and head suggest a colour. Consider the matching bigram “*chocolate espresso*” (freq = 2,548). As the stereotype *chocolate-brown* maps to the RGB code #7B3F00, and the stereotype *espresso-black* maps to #393536, a creative system can infer that the colour named by “*chocolate espresso*” will have an RGB code that sits somewhere on the line connecting #7B3F00 to #393536 in RGB space. Veale (2011) demonstrates how phrases like “*chocolate espresso*” are retrieved from the Google n-grams because the stereotypes for *chocolate* and *espresso* have shared properties, such as *smooth* and *dark*, allowing a system named the [Jigsaw Bard](#) to invent the simile “*as smooth and dark as a chocolate espresso*.” In effect, what we aim to achieve here is the generation of novel similes that have discernible perceptual groundings.

The CIR query “*^stereo ^stereo*” retrieves 5,841 bigram phrases from the Google 2-grams, from “*lemon tree*” (frequency=3,236”) and “*honey mustard*” (freq=3,120) to “*Brick Park*” (freq=40) and “*Bear Shadow*” (freq=40). When this query is applied to the Google 1-grams – by splitting complex unigrams into their lexical parts – an additional 5,666 unigram readymades are found, ranging from “*honeymoon*” (frequency=2,410,981, which may be interpreted as a pale blend of *honey-yellow* and *moon-white*) to “*firemelon*” (freq=200, perhaps naming a blend of *fire-red* and *melon-orange*). The least frequent names also tend to be the most enigmatic. Consider “*braincloud*” (freq=201), which suggests a striking name for a shade of gray, or “*demonmilk*”, “*coralstar*” and “*bananadragon*”. These seem to have been crafted by another person in another context to name some idea or thing; now they can be used again, this time to provocatively name a colour.

These readymades are not manually filtered for quality, and so, as CIR cannot disambiguate word-senses in n-grams, it may retrieve phrases that use colour stereotypes in non-stereotypical senses. For instance, CIR retrieves “*Holly Hunter*” (an actress, but also a potential blend of

holly-red and hunter-green) and “*Tiger Woods*” (a famous golfer, but also, potentially, a tawny blend of tiger-orange and wood-brown). Recall that the ultimate artistic value of a readymade lies in its ability to be re-interpreted with a new meaning or a new resonance. An orange-brown colour named *Tiger Woods* would be not just apt then, but humorously apt, and we should embrace this serendipity.

Each readymade can be assigned a potential RGB code at its moment of retrieval, by employing a parameterized mixture model to the RGB codes of its lexical ingredients. For a readymade like “*chocolate espresso*”, whose words denote nearby points in RGB space, we can simply split the difference and average the colours, so that chocolate espresso is a mix of 50% chocolate-brown (#7B3F00) and 50% espresso-black (#393536). When these components denote more distant colours/codes, it is necessary to bring linguistic and perceptual intuition to bear on them. For instance, we can expect “*chocolate forest*” (freq=153) to denote a different hue than “*forest chocolate*” (freq=170). The rules of compounding suggest that “*forest chocolate*” denotes a kind of chocolate, and that its colour should be perceived as a brown hue. In contrast, as “*chocolate*” is a modifier, not a head, in “*chocolate forest*”, we expect this name to denote some variation of (forest) green. As such, *forest chocolate* should contain as much *forest-green* as one can put into it while keeping it an identifiable brown, while *chocolate forest* should contain as much *chocolate-brown* as is possible while achieving a green hue overall.

The assignment of colours to readymade phrases is one side of the coin, of which the naming task is the flip side. Given an RGB color code, a creative naming system must assign an apt and original name to this code. This is the specific task that we focus on in this paper.

### O, speak again, bright angel!

Suppose one wanted a creative Twitterbot to respond to the postings of another bot, such as @everycolorbot. In this case, our responsive bot could await new tweets from @everycolorbot, extract the RGB code from each, and generate a catchy name for this colour to tweet as an apt response. Alternately, our bot could invent its own names for much loved colours on [colorlovers.com](http://colorlovers.com), to compete with names already invented by human users of the site.

Suppose our CC bot is given the RGB code #FCF9F0, a code which corresponds to a very pale yellow hue and which, on [colorlovers.com](http://colorlovers.com) has received 69 loves (and the name “*vanilla ice cream*” from one of the site’s users). Locating #FCF9F0 on an RGB colour wheel (Jennings, 2003), we consider this to be the dominant colour in an *analogous* colour scheme (see Pentak, 2010) in which the dominant color sits between two other colours, #FCF3F0 and #F9FCF0 on the colour wheel. We refer to #FCF3F0 and #F9FCF0 as *analogous* colours of our given colour, #FCF9F0. We choose to use an analogous colour scheme because it allows us to find adjacent colours that match well and which are often found together in nature. We then use these two analogous colours to find a readymade

name for the dominant colour it brackets on the wheel, one that is perceptually and linguistically appropriate.

For each analogous colour, our system seeks out the most appropriate colour stereotype. But first, we convert all relevant RGB codes into the equivalent CIE LAB code (Sharma 2003:29-32). The CIE LAB space is perceptually uniform, so any change  $\delta$  in a CIELAB code induces a uniform change  $\delta'$  in the perceptibility of the equivalent colour. The *Delta E CIE76* distance function can now be used to measure the distance between a given colour and that associated with any colour stereotype term. Thus, for instance, the *Delta E CIE76* distance between #FCF3F0 and *seashell-white* (#FFF5EE) is 2.17, while the distance between #F9FCF0 and *pearl-white* (#F7FBEF) is 0.55. As it happens, these two stereotypes – *seashell-white* and *pearl-white* – are the closest available colour stereotypes for the analogous colour pair #FCF3F0 and #F9FCF0.

Multiple readymades may each combine the words “*pearl*” and “*seashell*” in various ways. But as neither of the unigrams *pearlseashell* or *seashellpearl* is attested in the Google 1-grams, the system cannot choose a solid compound for a name. But the Google 2-grams do attest to the bigrams “*pearl seashell*” (freq=1,383) and “*seashell pearl*” (freq=5,633), and also attest to the plural bigram “*seashell pearls*” (freq=421). To maximize its chances of choosing a phrase that is semantically and syntactically well-formed, the system most prefers to choose attested unigram names, as these are most likely to have been coined as names; if it cannot find an attested unigram, it prefers a plural bigram, such as “*seashell pearls*”, as these are more likely to have been coined as a *modifier:head* construction; if it cannot find an attested plural bigram, it settles for the most frequent bigram (e.g. *seashell pearl*). In this case, it opts for the plural bigram “*seashell pearls*” and chooses its singular form, “*seashell pearl*” as a name.

A glance through any paint catalogue reveals that the most popular paint names are those that appeal to our love of nature, to our appetites, or to our aspirations. So paint names often use naming elements that denote a natural kind (*tree, pearl, forest, sea, etc.*), a food or drink (*toffee, butter, almond, espresso, etc.*) or a distinctive culture or place (*China, Persian, etc.*). So words such as *tandoori* and *kangaroo* tick two boxes at once. We may filter our readymade names by their adherence to this scheme, and choose only those phrases that use a colour stereotype that suggests a natural kind, food, drink, culture or place. The *Thesaurus Rex* Web service of Veale and Li (2013) can be used to provide fine-grained categorizations of colour stereotypes (such as *kangaroo, butter, pearl, etc.*) and to filter possible readymades by the categories they evoke. The filter employed by a naming system determines its aesthetic sensibility, and different systems may exhibit different aesthetic senses. One can imagine a system that prefers poetic names, smutty names, provocative names (e.g. *cocainestar* for a whiteish hue) or fantastic names (e.g. *alienbrain* for a gray-green hue). In the following experiments, our system employs the *paintshop-friendly natural-animal-food-drink-culture* filter described above.

## Beauty doth varnish age, as if new-born

To evaluate the quality and aptness of the readymade phrases that we repurpose as attractive new colour names, we compare these automatic names to those assigned by humans on the website [ColourLovers.com](http://ColourLovers.com). We download the [top 100,000](#) colour codes from this site, ranked from most to least *loves*; the mean number of *loves* per colour code is 13, while each code has at least one *love* and just one human-assigned name (as the site does not permit multiple names for the same RGB code). For each RGB code our automated naming system seeks out the most apt readymade name it can find. To ensure a good perceptual match between each code and its new name, a threshold distance of 14 is chosen for use with the *Delta E CIE76* distance function, which measures Euclidean distance in the CIELAB space. Thus, the CIELAB code of any colour stereotype (such as *pearl-white*) will only match the CIELAB equivalent of an analogous RGB code (such as #F7FBEB) if their Euclidean distance in CIELAB space is 14 or less. We choose a maximum of 14 empirically, so as to impose tight control on colour matching while allowing every colour code to be assigned at least one readymade.

We automatically identify the most apt readymade for each of the 100,000 downloaded colour codes, using the preferential approach to n-gram selection outlined in the previous section. Of the 100,000 assigned names, 2587 are selected as paintshop-style names using the aforementioned *natural-animal-food-drink-culture* filter. It is this subset of readymade names that we focus on here for purposes of empirical evaluation. The mean number of *loves* for each of the named colours on [ColourLovers.com](http://ColourLovers.com) is 2.188. For each of the 2,587 machine-generated names, we determine the name assigned to the corresponding RGB colour by users of [ColourLovers.com](http://ColourLovers.com). This allows us to construct a set of 2,587 triples, each comprising an RGB code, a human-assigned name and a name invented (via a repurposed readymade) by a machine.

We used these triples to pose comparison questions to human judges recruited via the crowd-sourcing platform [CrowdFlower.com](http://CrowdFlower.com). For each triple, a visual sample of the colour and a pair of names, one human-generated *and* one machine-generated, were put before the judges, who were asked to take a moment to imagine the colour being used. The ordering of both names was randomly selected on a case-by-case basis, so that the human-generated name was listed first in ~50% of cases, and the machine-generated name was listed first in the other ~50% of cases. In all cases, judges were *not* told of the origin of either name. Each judge was paid a small sum to answer 4 questions:

1. Which name is more descriptive of the colour shown?
2. Which name do you prefer for this colour?
3. Which name seems the most creative for this colour?
4. Why did you answer these questions the way you did?

The fourth question is a source of qualitative responses that may, in future work, offer useful insights into the factors that shape the appreciation of names. Judges were

timed on their responses, and those that spent less than 10 seconds presenting their answers for any colour were classified as *scammers* and discarded. We required that each question be answered by 5 non-scanning judges to be trusted for evaluation, and thus, we obtained 12,608 trusted judgments in all that contributed to the evaluation, and 5,040 untrusted judgments that were instead ignored.

A total of \$220 was allocated to the experiment, which was terminated after these funds were exhausted and 940 judges had been paid to contribute to the task. At this point, 1578 out of 2587 colours had received five trusted judgments for each of their questions, and so it is on the collected judgments for these 1578 colours that we base our evaluation. Tallying the individual judgments per question, we see that 70.4% of individual judgments for *most descriptive name* (Q1) favored the machine; that 70.2% of individual judgments for *most preferred name* (Q2) favored the machine; and that 69.1% of individual judgments for *most creative name* (Q3) favoured the machine. Similarly, when we tally the majority judgment for each question under each colour – the choice picked by three or more judges – we see that for just 354 (23%) of the 1578 colours, a majority of judges deemed the human-assigned name for a given colour to be more descriptive than that assigned by the machine. The results for the next two questions, Q2: *which name do you prefer?* and Q3: *which name is most creative?*, are very much in line with those of the first question. Only for 355 colours does a majority of the five human judges for a given colour prefer the human-assigned name over that assigned by the machine, and only for 357 colours does a majority of judges consider the human-assigned name to be more creative than the machine-assigned name. This consistent breakdown of approx. 3-to-1 in favour of the machine suggests that machine-assigned readymade names can be more than competitive with human names.

However, the surprising consistency of these results also suggests that the human judges are really only offering *one* opinion for all three of the binary questions that they are asked. It seems that judges, who are asked to ponder the possible users of a colour before answering the questions that follow, apparently favour a given name for a colour and *then* follow through with much the same answer for all three questions. Indeed, when we calculate the rate of agreement across all questions, we find that judges choose the same name for at least two of the three questions in 93% of cases, and choose the same name for all three of the questions (that is, most descriptive, most preferred and most creative) in 91% of cases. These agreement statistics suggest that most human judges see these questions as paraphrases of each other. Though it can aid our understanding of the mechanics of linguistic creativity to try and tease apart the related notions of descriptive adequacy, personal preference and creative appreciation, these three notions now appear to be too tightly interwound to effectively separate them, at least within the same experimental task.

## Let our bloody colours wave!

A Twitterbot named [@HueHueBot](#) has been constructed (by the second author, [Khalid Al-Najjar](#)) to showcase the perceptually-anchored creativity of this readymade-based approach to colour-name invention. An example tweet of this bot, with attached colour sample, is shown in Fig. 1.



Figure 1. A tweet with both RGB hexcode and apt name.

[@HueHueBot](#) exploits colour stereotypes and Google n-grams in the manner described in previous sections. But this inventory of colour stereotypes and their RGB codes can be reused by other Twitterbots that exhibit their own colour aesthetics and linguistic framing preferences. To this end, we gave the stereotype lexicon and a large stock of relevant n-grams to students as resources to be used for a course project on computational linguistic creativity. Students were asked to build colour-naming Twitterbots which might invent and name their own colours, or name the colour codes generated by [@everycolorbot](#). The bots that ensued demonstrate a variety of possible approaches to naming and to the linguistic framing of those names.

[@ColorCritics](#) frames its outputs as though it as an art critic that specializes in colour, and thus, in addition to offering to name colours generated by [@everycolorbot](#), it critiques the palette choices of this bot. [@ColorCritics](#) expresses a preference for unigram names, of which examples include [TandooriTikka](#), [PukePuke](#) and [FireSky](#). [@WorldIsColored](#) mimics the bravura personality of Stan Lee, a famous creator of comic book superheroes, and thus expresses a preference for colour names that use alliteration (a much-loved ploy of Lee's). Its alliterative colour names, such as [BlueberryBlush](#), are framed in the language of superhero comics, such as in this tweet: "May be coloring my costume as BLUEBERRY BLUSH was not a very good idea! RT. [@everycolorbot](#): 0xdd4fc3".

[@ColorMixALot](#) combines 2-gram phrases to generate complex colour names that run to three and four words. Example colour names include [tree frog bile yellow](#) and [moonlight coral pink](#). The Twitterbot [@DrunkCircuit](#) adopts the persona of a bored worker at an IT company, and so its tweets drip with ennui and bitterness. Examples include the sarcastic riposte to [@everycolorbot](#) in Fig. 2.

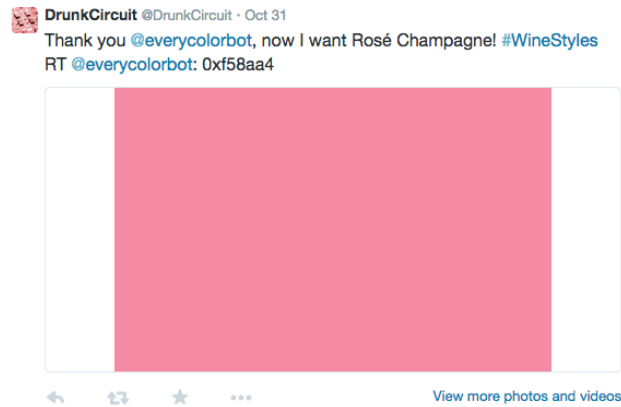


Figure 2. A sarcastic response to another colour bot: "thank you [@everycolorbot](#), now I want Rosé Champagne [#WineStyles](#) [@everycolorbot](#): 0xf58aa4"

Like [@HueHueBot](#), [@DrunkCircuit](#) locates the category into which a new name fits best (using Wikipedia's hierarchy of topic categories), and then tailors its tweets to exploit this information. Thus, a name that denotes a kind of wine (as in Fig. 2) is affixed with the hashtag [#WineStyles](#), while the name [Almond Crust](#) is used to anchor a tweet that insults the company canteen ("Looks just like the Almond Crust in the canteen today. Yuck! RT [@everycolorbot](#): 0xd3ba8f").

[@haraweq](#) is a colour-naming hybrid that combines elements of two popular Twitterbots, [@everycolorbot](#) and [@metaphorminute](#). The latter is a bot by Darius Kazemi that invents random metaphor-like tweets, such as "an evacuation is a mainframe: evergreen yet slicked." In this vein, [@haraweq](#) coins colour similes, such as "a location like a dusty taxicab RT [@everycolorbot](#): 0xf4ec24." It uses Wikipedia to determine e.g. that a taxicab is a location, and uses the Google n-grams to find specific combinations such as "[dusty taxicab](#)", which it interprets as a blend of *taxicab-yellow* and *dust-brown*.

[@AwesomeColorBot](#) also tailors its tweets to suit the category of a name, to produce [outputs](#) like that of Fig. 3.



Figure 3. A tweet with a colour, a name, and an attitude.

So the most interesting colour bots do more than just invent new colour names; they find a context to motivate a new name, and then frame a tweet as an intelligent – or at least a human-like – response to this context. There is a lesson here for computational linguistic creativity. A new turn of phrase can only be considered creative in a context for which it is non-obvious and apt, and to the extent that it exercises the imagination of the reader. The imagination may take flight on the wings of whimsy, but the most compelling flights into the new and the original remain stubbornly grounded in the realm of familiar experiences.

## Acknowledgements

This research was facilitated by a travel mission funded by the University of Helsinki and by the EC coordination action PROSECCO (*Promoting the Scientific Exploration of Computational Creativity*; [PROSECCO-network.EU](http://PROSECCO-network.EU)). The authors would like to thank Prof. Hannu Toivonen at the University of Helsinki and all the students of the 2014 UH class on Computational Linguistic Creativity (whose bots [@ColorCritics](https://twitter.com/ColorCritics), [@WorldIsColored](https://twitter.com/WorldIsColored), [@ColorMixALot](https://twitter.com/ColorMixALot), [@DrunkCircuit](https://twitter.com/DrunkCircuit), [@haraweg](https://twitter.com/haraweg), and [@AwesomeColorBot](https://twitter.com/AwesomeColorBot) are discussed here; please do check them out on Twitter). We also wish to express our gratitude to Mike Cook of Imperial College, London, who suggested the idea of a colour-naming Twitterbot, and to Hyesook Kim, who painstakingly mapped colour stereotypes to RGB codes.

## References

Thorsten Brants and Alex Franz. (2006). Web 1T 5-gram database, Version 1. *Linguistic Data Consortium*.

Kim Binsted and Graeme Ritchie. (1997). Computational Rules for Generating Punning Riddles. *HUMOR, the International Journal of Humor Research*, 10(1):25-76.

Christian F. Hempelmann. (2008). Computational Humor: Beyond the pun. In Victor Raskin (ed.), *The Primer of Humor Research*. Berlin: Mouton de Gruyter.

Rachel Giora, Ofer Fein, Jonathan Ganzi, Natalie Alkessassy Levi, Hadas Sabah. (2004). Weapons of Mass Distraction: Optimal Innovation and Pleasure Ratings. *Metaphor and Symbol* 19(2):115-141.

Richard S. Hunter (1948). Photoelectric Color-Difference Meter. *Journal of the Optimal Society of America* 38 (7): 661.

Simon Jennings (2003). *Artist's Color Manual: The Complete Guide to Working With Color*. Chronicle Books.

Kevin L. Keller. (2003). *Strategic brand management: building, measuring and managing brand equity*. Prentice Hall.

Gozde Özbal and Carlo Strapparava. (2012). A computational approach to automatize creative naming. In Proceedings of the 50<sup>th</sup> annual meeting of the Association of Computational Linguistics (ACL-2012), Jeju Island, Korea.

Stephen Pentak. (2010). Analogous Color Scheme. *Design Basics* (8th edition.). Australia: Cengage.

Gaurav Sharma. (2003). *Digital Color Imaging Handbook* (1.7.2 edition). CRC Press.

Michael R. Smith, Ryan S. Hintze and Dan Ventura. (2014). Nehovah: A Neologism Creator Nomen Ipsum. *Proceedings of the 5<sup>th</sup> International Conference on Computational Creativity*, Ljubljana, Slovenia.

Tony Veale. (2006). Tracking the Lexical Zeitgeist with Wikipedia and WordNet. In Proceedings of ECAI'2006, the 17<sup>th</sup> European Conference on Artificial Intelligence, Trento, Italy.

Tony Veale and Cristine Butnariu. (2006). Exploring Linguistic Creativity via Predictive Lexicology. At the ECAI'2006 Joint International Workshop on Computational Creativity. Italy.

Tony Veale. (2011). Creative Language Retrieval: A Robust Hybrid of Information Retrieval and Linguistic Creativity. In Proceedings of ACL'2011, the 49<sup>th</sup> Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, Oregon.

Tony Veale. (2012). *Exploding the Creativity Myth: The Computational Foundations of Linguistic Creativity*. London: Bloomsbury Academic.

Tony Veale. (2013). Linguistic Readymades and Creative Re-use. *Transactions of the SDPS: Journal of Integrated Design and Process Science*, 17(4):37-51.