

Crowdsourcing Word-Color Associations

M. Lafourcade¹, N. Le Brun², V. Zampa³

¹LIRMM, Montpellier, France

²Imagin@t, Lunel, France

³LIDILEM, Grenoble, France

Abstract. In Natural Language Processing and semantic analysis in particular, color information may be important in order to properly process textual information (word sense disambiguation, and indexing). More specifically, knowing what colors are generally associated with what terms is crucial information. In this paper, we explore how crowdsourcing through a game with a purpose (GWAP) can be an adequate strategy to collect such lexico-semantic data.

Keywords: Word Color Associations, Lexical Network, Crowdsourcing

1. Introduction

Color information is important in our daily life and may be of interest in the context of Natural Language Processing. Although, this is beyond the scope of this paper, there are strong connections between colors and emotions. However, providing information about word-colors associations to a system dedicated to the semantic analysis of texts, in addition to other traditional knowledge (hypernym, parts of, semantic role, etc.) could greatly improve system performance. Association between word and color could be made for abstract nouns related to emotions (like fear, anger, danger, hope, ...) but in a more straightforward way for concrete nouns (like sky, lion, snow, sea, ...)

There have been many psychological studies about colors and their association to words or their impact on communication (either verbal or not). Mohammad (2011a) explains that color is a crucial aspect for the successful delivery of information, whether it is in marketing a commercial product (Sable and Akcay, 2010), in web design (Meier, 1988; Pribadi et al., 1990), or in information visualization (Christ, 1975; Card et al., 1999). Color is definitely a strong bearer of sense.

But the relation between colors and senses, for many researchers in psychology/sociology depends on a number of factors. Luscher (1969), a psychotherapist known for inventing the eponymous color test, a tool to measure a person's psychophysical state based on his or her color preferences, argues that the choice of one color by a person is related to (or even translates) his emotional status. Child et al. (1968), and Ou et al. (2011) show that people of different ages and genders may have different color preferences. Moreover, some colors may have lexical variation according to languages and culture (for example Hungarian and Turkish have two words for *red*).

There is a very lively debate as to whether the associations between color and direction were independent of age, gender, or nationality. This might be the case, for example, for the relation between *red* and *danger*, since the red may be related to *blood/fire* regardless of other factors. Berlin and Kay (1969) argued that differences could be organized into a coherent hierarchy, and that there is a limited number of *universal* basic color terms used in various cultures. Berlin and Kay based their analysis on a comparison of color words in 20 languages from around the world, but their findings have been discussed a lot. In the same way, several expressions using colors have the same meaning in different languages especially when they are culturally close, which is hardly a surprise. For example, *dark thoughts* in English and *idées noires* are roughly equivalent, as well as *to see red* and *voir rouge*.

Many studies, mostly in English, have been undertaken to determine relations between colors and words or colors and emotions, etc. Most of those studies of psycholinguistics are undertaken in a classical way, and their raw results are generally modest in size and not freely available. Furthermore, as previously mentioned, it is extremely delicate (and probably unwise) to translate directly the result of

such studies from one language to another. Finally, we can say that there is no definite consensus on the universality of word-color associations for abstract or feeling words-

Mohammad (2011b) conducted experiments using eleven colors and showed that more 30% of the terms have a strong color association (for a lexicon of everyday words, not including specialized domains). About 33% of thesaurus categories (like Roget) have color associations, and abstract terms are associated with colors almost as often as physical entities do, mostly by metaphor.

Again, Nijdam (2010) compares different models on relation between color and emotion and proposes a correspondence between emotions and colors. He concludes that some models about color-meaning may have some overlap but they also show a great amount of vagueness, certainly because the color is dependant of personal/cultural situation. In summary, the acquisition of data on associations between words and colors is hampered by high variability, (even restricted to a given language). The more abstract is the term, the higher the variation is to be anticipated.

Gözde et al. (2011) make a resource that contains information about the association of words and colors in English. They made a short selection of 200 words, a subset of words used in Grefenstette (2005), and compare annotations made with Amazon Mechanical Turk service (10 annotators – 11 colors) with three automatic methods: image analysis, language models on web data and similarity between words and colors (using LSA).

At the present time, as far as we know, Gözde et al. (2011) are the only ones who tried to make a resource about association of colors and words in English. Such resources are very rare in other languages, and especially it doesn't exist in French language. But as noted by Grefenstette (2005) what people generally know about things (concrete things, but even abstract concepts) includes as an important component their typical color. That is why, this information is not usually shown in dictionary definitions or lexical resources, although it would be useful for various computer applications, and in NLP in particular.

We think that color information can be very helpful in the context of automated lexical disambiguation (Word Sense Disambiguation - WSD). For example, the French word *tissu* is polysemous and can mean either *fabric* or *living tissues*. If in a given text, *tissu* is associated to *bleu* (blue), the color information can help to choose (or to find) the right meaning or at least eliminate the wrong ones. Another computer application for which color information would be helpful is one that allows you to find a momentarily forgotten word but "on the Tip of the Tongue" (Lafourcade 2012 and Joubert 2012). The color information is very often crucial for finding tip of the tongue terms.

The objective of this paper is to present how to efficiently produce a (French) resource that lists associations between colors and words. For this purpose, we implemented a GWAP (game with a purpose as dubbed by L. von Ahn. (2006)) named ColorIt through which the player is asked to tell the color he spontaneously associates with a given term. (accessible at the following url <http://www.jeuxdemots.org/colorit.php>). At first we show how information on color is a valuable asset in the context of WSD, and then we detail the features of the game as well as some quantitative and qualitative results.

2. Word-Color Association and WSD

One of the major advantages of a lexical resource containing associations between words and colors is to use it in the context of WSD. In the given text, a reference to a color can be a crucial information to select the proper meaning of a term. Of course, for some words the color information is not applicable (*no color*), but *no color* can be useful to select the appropriate sense of a polysemous word in the context of WSD. Some terms may have a wide range of colors (like *car*), or several colors at the same time (black and white for a *penguin*, many but not all for a *rainbow*).

Without claiming to be exhaustive, we illustrate the relevance of the color information in word sense disambiguation by the following French words (meanings that could be eliminated with the color information are marked *, the ? indicates that ambiguity remains):

lit blanc	lit > couche : blanc, rouge, ... (bed) (almost any color) * lit>rivière (river bed)
regard bleu	regard> expression (expression, look) * regard>trou (manhole, trap)
trompe grise	trompe>éléphant (trunk) *? trompe>instrument de musique (horn) * trompe>reproduction (sexual organ)
bas noir	bas> sous-vêtement (stoking) * bas>partie inférieure (bottom)
manteau rouge	manteau>vêtement * manteau>géologie
ours blanc	ours>animal (bear) * ours>imprimerie (printing)
piste blanche/ verte / rouge	piste>ski (skiing) piste>chemin (trail)
langue blanche / rose	langue>organe (tongue) * langue>langage (language)
sucre blanc	sucre>substance (sugar) *? sucre>électricité (lustre terminal) (ambiguous but improbable)
tableau noir	tableau>école ? tableau>oeuvre d'art (improbable)
robe rose / bleu	robe>vêtement (dress) * robe>pelage (animal)
robe blanche /jaune	? ambiguous for some colors
feuille blanche	feuille>papier (piece of paper) (most probable) *? feuille>arbre (leaf) (improbable)
rouge / jaune / vert	? ambiguous for some colors
savon noir / rose / blanc	savon>savonnette (soap) savon>réprimande (scolding)
brioche dorée	brioche>viennoiserie *? brioche>ventre (paunch, gut)
canapé bleu	canapé>meuble (sofa) * canapé>petit four (canapé) ? ambiguous for some colors
culotte bleu / rouge/ blanche	culotte>sous vêtement (underpants) * culotte>boucherie (rump)
motif fleuri	motif>dessin (pattern) motif> raison (motive)
papillon multicolore	papillon>insecte (butterfly) *papillon>nage (butterfly stroke) *papillon>contravention (ticket)
costume anthracite	costume>tenu (suit) (most probable) *? costume>déguisement (outfit)

Table 1. The relevance of color-information in word sense disambiguation (WSD)

Some word meanings can be mostly eliminated when a color is provided. In some cases, some ambiguity ~~hold~~ remains, but a level of uncertainty can be computed.

2. ColorIt, a Game for Collecting Colors

The goal of ColorIt is to collect spontaneous associations between colors and words, colors assigned to concrete terms as well as those symbolically and subjectively associated to verbs or nouns denoting abstract entities.

2.1 Game Principles

A word is presented to the user (see Figure 2 in annex) along with a choice of colors. The user is invited to click on the color he associates to the displayed word. It is possible to associate *no color* (if color is not applicable) or to pass (if the player does not know the word for instance). Passing is not penalized. As an alternative to the color palette, a text field is (optionally) proposed in which experienced player can enter *several* colors (as free text).

Once the choice is done, a notification indicates the score along with the answers of other players for the same word. The score depends on the adequacy of the response of the player and the color distributions already assigned to the word *via* other answers. There are two types of answers: specify one or several colors or choose *no color/not applicable*. If the player's choice corresponds to the other votes, then he gets some points (max 50), otherwise he loses some points (max 100). The amount of points is related to the normal distribution of colors related to the word in the lexical network. If a color is mildly associated (compared to others), the player earns a lots of points. If the color is very strongly or weakly associated, the player does not gain many points. The associated color weights are increased by one for each vote. A newly introduced color is not rewarded (0 vote before), but will be rewarded afterward; it enters the lexical network with a weight of 1.

For example (very simplified considering the ~~actual~~ current distribution), suppose the term *elephant* has the following colors with weights: *gray/24*, *white/5*. If the player selects *grey* he will win 24 points and *grey* is set to 25. If the player selects *no color* he loses 29, but the *no color* relation is introduced in the lexical network with a weight of 1 for this term.

This system encourages players to play honestly by selecting the proper colors, but also to try to be original when possible. The assumption, that has been empirically verified and is rather common to all GWAP, is that a very large majority of people will play honestly. A mistake may be costly in points. A player who tries to sabotage the data by systematically answering incorrectly is easily detected and his answers are not taken into account. Every 100 earned points, the player's level increases by one. The number of characters allowed in the free text zone is proportional to the player's level.

Settings that make the game exciting are the possibility to compare one's answers with other players and the increased in level, since the higher the level, the higher one can enter words in the text field free. Most of the players choose to enter several colors instead of clicking on only one.

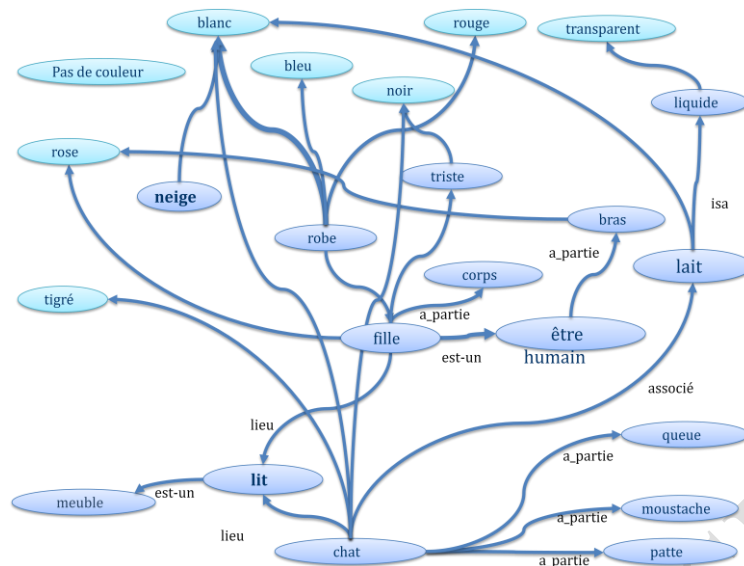


Fig 1. The JDM lexical network contains word meanings (refinements) linked to the main term. For example, *avocat* is linked to *avocat>fruit* (avocado) and *avocat>justice* (lawyer). Collecting information that is specific for some meanings and not others is of crucial importance if the resource has to be used in Word Sense Disambiguation. Colors are amongst the information looked for.

ColorIt is a quite challenging game, and we received very positive feedback from players. Beside some terms whose color is obvious (*snow*, *night*, *coal*, etc.) and no lend itself to interpretation, some others may have a wide range of possible colors either objective for concrete nouns (like *flower*, *car*, etc.) or subjective for abstract concepts (such as *anger*, *sadness*, etc.) or no color at all (*increase*, *subjectivity*, *forwardness*, etc.). Choose a color that is believed to be the most relevant is not an easy task and the confrontation with the other answers may generate excitement, suspense, and surprise ...either good or bad ! Although the purpose of this game is to collect lexico-semantic data, no specific linguistic knowledge is required to play, and the data collected are of good quality.

2.2. Behind the Scene

The JDM project (Lafourcade, 2007 and Chamberlain, 2013) aims at building a very large lexical and semantic network through games (GWAP, Games With A Purpose) and crowdsourcing. Such a network is composed of lexical items as nodes (vertices), and relations between nodes. Relations are typed, oriented and weighted. The lexical network for French has been made freely available to the community, so as to be an interesting resource to work with. The resource contains over 300 000 terms and 5 millions relations.

The JDM lexical network (illustrated in Figure 1) contains terms, word meanings and various semantic information (like person, living being, abstract, concrete, artefact, ...). Relations are weighted (the most obvious or frequent having the higher value). Some relations may have a negative weight to represent an interesting impossibility or exception (like for example *ostrish --can:-100-- fly*). Some impossible colors can be represented with such an approach.

The word meanings are connected to the main word with a refinement relation. For example; *frégate* (frigate) is connected to *frégate>navire* (frigate ship) and *frégate>oiseau* (frigate bird). The term *frégate>navire* is itself refined as *frégate>navire>moderne* (modern frigate ship) and *frégate>navire>ancien* (ancient frigate ship). Each refinement may itself be connected to several other terms in the network.

In this network, the information about color existed until now only through the *characteristics* relation, thus mixed with other information. Now it is the subject of a specific relation that the game COLORIT enriches consistently. A heuristic algorithm was designed to have a reasonable probability of finding a

term for which the color information is meaningful. Indeed, we avoid to propose too many terms for which the color information would be irrelevant, so as not to deter players.

The words which are displayed to the player are selected *via* the algorithm as follows: We select randomly, in the *JeuxDeMots* lexical network, a word that has at least one color relation, and we propose, through a virtual coin tossing, either the term or one of its linked neighbors in the network. This causes a rapid spread of color information across the network. A term, that has been repeatedly tagged as *without color* will be removed from the list of selectable neighbors. In the same way, the refinements of "colored" words quickly become "colored" themselves.

The assumption behind this *propagation algorithm* is that there is more chance to select a term eligible for color information if it is linked to a term having already a color characteristic. A completely random selection would be counterproductive because most of the time a color characteristic would not be applicable. In the JDM lexical network, we estimate by sampling that only about 10% of the approximately 300.000 terms (30,000 out of over 300,000) are "color eligible".

If a response provided by the player is in the lexical network but is not known as a color/appearance, it is then proposed and will be validated (or invalidated) as such by an administrator. But if the proposed term does not exist at all in the network, it must first be integrated in the network before being characterized as a color or appearance (again by an administrator). The Diko interface of the JDM lexical network is a contributory tool for validation of proposals (<http://www.jeuxdemots.org/diko.php>).

The propagation algorithm has been bootstrapped with the colors proposed in the game: *white, grey, white, red, dark red, orange, yellow, light green, green, blue, light blue, purple, pink, beige, brown*. Before the creation of COLORIT, these colors were associated with many terms through other relations than color.

3. Data Collected and Evaluation

After 3 months (from August to November 2013), more than 32,000 word-color(s) associations have been created and more than 15,000 eligible terms were provided with color information. So, on average, one term is associated with 2.2 colors. The number of votes exceeded 320,000. From the figures above, we estimate that half of the terms in the JDM lexical network that are color eligible (15000 out of over 30.000) have been "colorized"

As soon as a color is associated to a term, the network is updated with this color or appearance. For example, *eau* (water) is associated with different blues (*light blue* (26), *blue* (9), *blue lagoon* (1), *turquoise* (1), etc.) many colors (*greenish* (4), *green* (2), etc.) and different appearances (*cloudy* (20), *transparent* (9+9), *crystal clear* (4), *colorless* (4), etc.). With the possibility to add a new color or appearance *via* the free text field, the network has been completed with names of specific colors like *yellow carthusian, tangerine*, etc. or appearances like *striped, translucent*, etc. or with new combination of colors. The color list is available in annex. At the present time, more than 700 terms refers to colors or appearance.

The data collected were evaluated by hand, on two random samples of around 500 terms, totalizing around 2,500 term-color associations. The first sample S_C was taken from the terms having at least one color associated, the second sample S_{NC} amongst those with at least one vote *no color*. Terms in S_C may have some *no color* associations and conversely, terms in S_{NC} might have some colors associations. We asked volunteers to examine randomly some word-color associations and tell whether they were correct or not (either as color or *no color*).

The evaluation approach, where people are asked to judge word-color associations (closed task) is an opposite process compared to the acquisition where people are asked to produce associations (open task). However, this evaluation protocol does not inform on the color combinations that may have been omitted.

	Proper colors	Wrong colors	Proper no color	Wrong no color	Total
# votes	13308	366	708	81	14463
%	92 %	2,5 %	5 %	0,5 %	100

Table 2. Percentages of correctly and incorrectly assigned colors

The evaluation by the volunteers covered about 10% of term-color associations generated by the players (acquisition). The collected data seem to contain a quite low percentage of wrong associations (see Table 2).

Furthermore, closer inspection of the lexical network shows that wrong associations are assigned a low weight compared to the correct associations (less votes). For example, for 1 vote as *blue* for *sun*, we have more than 100 votes as *yellow*. In addition, the examination of the ambiguous (polysemous) words that have been characterized with color reveals that the information of color is usually adequate and could then be useful for discrimination in WSD.

Term related to	Politic (2702 terms)	Zoologie (3191 terms)	Arts (2815 terms)	Emotions (304 terms)	Plants (4481 terms)
agreement	0.55	0.65	0.71	0.28	0.76

Table 3 Kappa between players according to semantic field (class) to which the word belongs. The number in parentheses indicates the number of terms related to this semantic field in the lexical network.

Table 3 presents the evaluation of the agreement between players based on some topics (semantic fields) of the suggested words. The agreement value for a given term is the mean of the Cohen kappa between players associations taken in pairs for this term. The agreement value for a class is the mean of the agreements of the terms of this class (those that have been played). Some terms may relate to more than one class. This is for the semantic field related to plants that the agreement between the players is the highest. It appears that many players refer to external sources (like Wikipedia) to find out what color(s) must be assigned to an exotic or unknown plant or animal. The weakest agreement is that related to the semantic field of emotions, which is easily explained by the high degree of subjectivity of corresponding associations (and this although our experiment has been conducted only in French, and with a relatively homogenous population consisting of 70% women between 30 and 50). On the semantic field of the arts, there are many color names expressed through specific vocabulary, but once decrypted finally have quite a few variations..

4. Conclusion

We have designed a simple but challenging and effective game for collecting data about word-color associations. In a few months more than 15 000 terms were associated with one or more colors, and about 4,000 were tagged as irrelevant for color information. Many colors and appearances have been introduced in the JeuxDeMots lexical network *via* the game ColorIt. Currently, as far as we know, the data collected are the first French language resource for this type of associations. Furthermore, this resource is freely available dynamically (at <http://www.jeuxdemots.org/colorit.php?action=list>). One of the challenges in the game design was to achieve a propagation algorithm allowing the selection of eligible terms for color information, so as to avoid boring players in offering too many terms to which it is not appropriate to assign a color.

We had several surprise with the data collected with ColorIt. First, choose "no color" was originally designed to be a loophole for players when nothing suitable. But it soon became clear that the feature "colorless" was precious to eliminate certain meanings when polysemy. Indeed, in the context of word sense disambiguation, selection of meaning by the indications of color is more often done by elimination than by selection. One can explain this, at least partially, by the polysemy metaphorically: the abstract sense (without color) often comes from the concrete term.

Secondly, a side effect of the project was to add a large number of color names and appearances that were beforehand unknown in the lexical network. In addition to newly introduced true colors, the network now contains numerous terms related to the visual aspect, like *translucent*, *stripped*, *blurred*, *scratched*, *spotted*, etc.

To go further, another perspective might be to evaluate the correlation between color and feelings associated with certain words, especially abstract nouns. Once we have a sufficient number of words-colors associations, the comparison with other resources already available on the polarity, or feelings / emotions associated might be conducted without much difficulty, without forgetting their use in the context of word sense disambiguation. Moreover, a full study of the impact of the color information obtained by crowd sourcing in WSD tasks is still to be conducted.

5. References

1. von Ahn, L. (2006). *Games with a purpose*. Computer, 39(6):92–94, 2006.
2. Berlin, B. & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
3. Card, S., Mackinlay, J. and Shneiderman, B., (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann
4. Chamberlain, J., K. Fort, U. Kruschwitz, M. Lafourcade and M. Poesio (2013) *Using Games to Create Language Resources: Successes and Limitations of the Approach*. Theory and Applications of Natural Language Processing. Gurevych, Iryna; Kim, Jungi (Eds.), Springer, ISBN 978-3-642-35084-9, 2013, 42 p.
5. Child, I. L., Hansen, J.A., and Hornbeck, F.W. (1968). *Age and sex differences in children's color preferences*. Child Development, 39 (1):237–247.
6. Christ, R. (1975). *Review and analysis of color coding research for visual displays*. Human Factors: The Journal of the Human Factors and Ergonomics Society, 17:542–570.
7. Grefenstette, G. (2005). *The color of things: Towards the automatic acquisition of information for a descriptive dictionary*. Revue Française de Linguistique Appliquée 2 83–94
8. Joubert, A.; M. Lafourcade, M. (2012) *A new dynamic approach for lexical networks evaluation*. In proc of the Eight International Conference on Language Resources and Evaluation (LREC'12), Istanbul, Turkey, 23–25h May 2012.
9. Kaya, N. & Epps, H. (2004). *Relationship between color and emotion : a study of college students*. Academic journal article from *College Student Journal*, Vol. 38, No. 3.
10. Lafourcade, M. (2007). *Making people play for Lexical Acquisition*. In Proc. SNLP 2007, 7th Symposium on Natural Language Processing. Pattaya, Thailande, 13–15 December 2007, 8 p.
11. Lafourcade, M. Joubert, A. (2012) *Increasing long tail in weighted lexical networks*. In proc of Cognitive Aspects of the Lexicon (CogAlex-III), COLING, Mumbai, India, December 2012.
12. Luscher, M. (1969). *The Luscher Color Test*. Random House, New York, New York.
13. Meier, B. (1988). Ace: a color expert system for user interface design. In Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software, UIST '88, pages 117–128, New York, NY, USA. ACM.
14. Nijdam, N A. (2010). Mapping Emotion to Color, Human Media Interaction Human Media Interaction. University of Twente, the Netherlands. Available at http://hmi.ewi.utwente.nl/verslagen/capita_selecta/CS_Nijdam_Niels.pdf
15. Ou, L-C., Luo, M.R, Sun, P-L., Hu, N-C., and Chen, H-S. (2011). Age effects on colour emotion, preference, and harmony. Color Research and Application.
16. Ozbal, G., Strapparava, C., Mihalcea, R. and Pighin, D. (2011). *A Comparison of Unsupervised Methods to Associate Colors with Words*. Affective Computing and Intelligent Interaction. Lecture Notes in Computer Science. Volume 6975, 2011, pp 42-51.
17. Pribadi, N, S., Wadlow, M.G., and Boyarski, D. (1990). *The use of color in computer interfaces: Preliminary research*. Information Technology Center, Carnegie Mellon University, 49 p.
18. Sable, P. and Akcay, O. (2010). *Color: Cross cultural marketing perspectives as to what governs our response to it*. Proceedings of ASBBS, vol 17:1, -Las vegas, CA. February 2010, pages 950–954.
19. Saif, M. (2011a). Colourful language: measuring word-colour associations. Proceeding CMCL '11. Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics. 97-106. <http://www.aclweb.org/anthology-new/W/W11/W11-0611.pdf>
20. Saif, M.. (2011b). Even the Abstract have Colour : Consensus in Word–Colour. Associations. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technology. <http://www.aclweb.org/anthology-new/P/P11/P11-2064.pdf>

6. Annexes

6.1. List of Color Names Collected with ColorIt

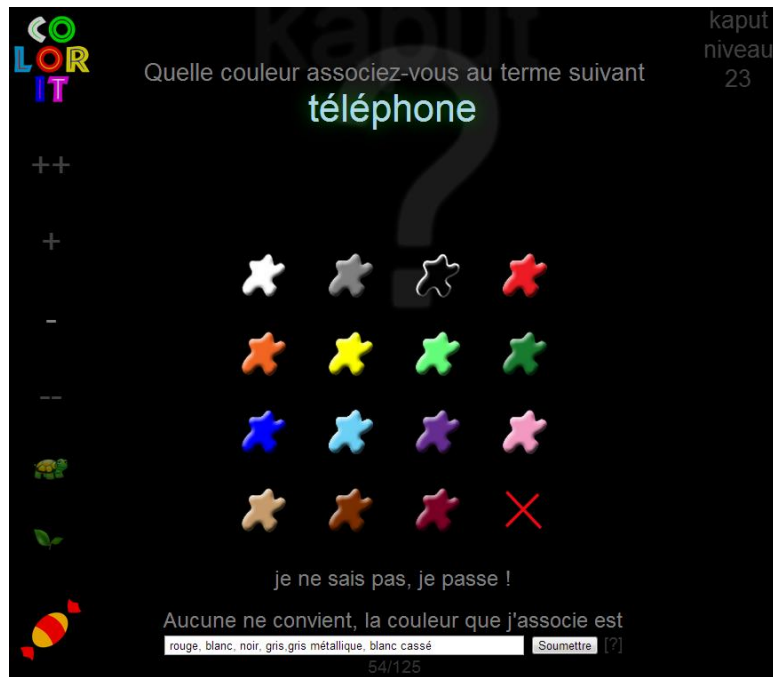
The color list below (available at <http://www.jeuxdemots.org/colorit.php?action=colorlist>) contains color names that are included in the JDM lexical network. Many of them have been introduced by players of ColorIt. Some terms are more appearances or visual aspects like *à carreaux*, *tigré*, *translucide*, ... Some color terms are color combinations as proposed by players, like *noir et blanc*, *blanc et rouge*, ... In the list below, terms denoting aspects rather than colors are in bold.

6.1. Screenshot

Figures 2a and 2b (below) are screenshots of ColorIt game. The game can be played on web browser and/or mobile phone.



Fig 2a and 2b. A typical screenshot of a ColorIt Game. The current game is to be played on the word *serviette* (which is polysemous: *towel*, *napkin*, *briefcase*). In the free text zone, the player can write several colors up to 15 characters. The notification on the right, concerns the previous game on the word *lavabo* (*washbasin*) where the user won 15 points by choosing *blanc* (*white*). The turtle icon on the left allows to increase/reduce the number of animals names. The same for plant names below. The candy allows getting only common words.



6.2. Some Various Word-Color Associations

<p>éléphant : gris:49 ; rose:19 ; blanc:2 ; gris clair:1 ; gris bleu:1 ; bleuté:1</p> <p>mort : noir:64 ; violet:20 ; feu:10 ; blanc:2 ; rouge:2 ; sans couleur:1 ; gris:1</p> <p>sang : rouge:50 ; rouge clair:16 ; bleu:9 ; rouge foncé:1 ; noir:1 ; vermillon:1 ; écarlate:1 ; cramoisi:1</p> <p>rouge : rouge:39 ; rouge brique:14 ; rouge vermillon:13 ; carmin:8 ; pâle:6 ; fluo:6 ; pourpre:3 ; cramoisi:2 ; écarlate:2 ; blanc:1 ; bordeaux:1 ; rouge sang:1 ; vermillon:1 ; rouge foncé:1 ; rouge cerise:1 ; auburn:1</p> <p>ciel : bleu:20 ; bleu clair:15 ; bleu ciel:13 ; gris:4 ; azur:4 ; rouge:4 ; noir:4 ; blanc:4 ; sombre:4 ; azur>bleu:4 ; clair:4 ; blanchâtre:1 ; ciel:1 ; bleu foncé:1 ; bleu azur:1 ; orangé:1 ; grisâtre:1</p> <p>soleil : jaune:53 ; rouge:14 ; brillant:9 ; orange:3 ; vert clair:2 ; orangé:1</p> <p>chat : blanc:14 ; noir:13 ; tigré:9 ; gris:8 ; roux>couleur:8 ; roux:7 ; noir et blanc:7 ; beige:6 ; brun:1 ; marron et blanc:1 ; rayé:1 ; marron:1 ; jaune:1 ; tacheté:1 ; blanc et noir:1 ; gris clair:1</p> <p>banane : jaune:36 ; blanc:20 ; verte:6</p> <p>sucré : blanc:42 ; roux:13 ; brun:11 ; naturel:10 ; beige:2 ; marron:1</p>	<p>corps: rose:46 ; naturel:8 ; beige:8 ; sans couleur:5 ; chair:5 ; blanc:5 ; rouge:4 ; noir:2 ; gris:2 ; pâle:2 ; jaune:2 ; rayé:2 ; tacheté:2 ; marron:2 ; brun:2 ; cuivré:1 ; doré:1 ; hâlé:1 ; bronzé:1 ; roux:1 ; noir et blanc:1 ; fauve:1</p> <p>gilet de sauvetage : jaune:38 ; orange:8 ; rouge:8</p> <p>marbre blanc : blanc:50</p> <p>disque noir: noir:55</p> <p>aile de corbeau: noir:39 ; noir profond:18 ; bleuté:18</p> <p>iceberg : blanc:61</p> <p>deuil : noir:49</p> <p>papier : blanc:51 ; à carreaux:12 ; pastel:1 ; rose:1 ; bleu clair:1 ; bleu:1 ; jaune:1 ; gris:1 ; blanc cassé:1 ; écru:1 ; bleu pâle:1</p> <p>vert : vert:16 ; céladon:9 ; vert olive:8 ; bleu-vert:8 ; émeraude:5 ; pâle:5 ; turquoise:5 ; bleu:4 ; vert clair:4 ; olive:4 ; bouteille>couleur:4 ; jade:4 ; clair:4 ; fluo:4 ; foncé:4 ; brillant:4 ; verdâtre:1 ; vert sapin:1 ; vert foncé:1 ; kaki:1 ; vert d'eau:1 ; verdouillasse:1 ; bleu vert:1 ; vert tendre:1 ; verdasse:1 ; vert-de-gris:1 ; vert bouteille:1</p> <p>assiette : blanc:39 ; blanc>couleur:20 ; liseré:20 ; bleu:3 ; jaune:2 ; blanche:2 ; rose:2 ; rouge:2 ; à fleurs:2 ; à pois:2 ; gris:2 ; vert:1 ; mauve:1 ; pastel:1 ; noir:1 ; beige:1</p> <p>lapin : blanc:43 ; gris:12 ; noir:11</p>
---	---

Table 3. Sample of colors associated to terms as collected with CorlorIt

trompe>éléphant (trunk)	grey, light grey
trompe>instrument de musique (horn)	golden, white
trompe>reproduction (sexual organ)	no color, pink, purple
bas> sous-vêtement (stoking)	black, grey, any color
bas>partie inférieure (bottom)	no color
manteau>vêtement	any color, but mostly red, black, grey, blue, brown
manteau>géologie	no specific color
ours>animal (bear)	white, brown
ours>imprimerie (printing)	no color
piste>ski (skiing)	white, blue, red, green, red
piste>chemin (trail)	sandy, brown, earth
langue>organe (tongue)	pink, white

langue>langage (language)	no color
sucré>substance (sugar)	white, brown
sucré>électricité (lustre terminal)	no color, grey, black
tableau>école	black, white
tableau>oeuvre d'art	any color, low probability of white, black
robe>vêtement (dress)	any color, mostly red, black, white, blue
robe>pelage (animal)	skink, brown, light brown, white, black
feuille>papier (piece of paper)	mostly white, light blue
feuille>arbre (leaf)	green, yellow, but no blue, white
savon>savonnette (soap)	pink, white, greenish
savon>réprimande (scolding)	no color
brioche>viennoiserie	yellow, brown, <i>doré</i>
brioche>ventre (paunch, gut)	skin, reddish,
canapé>meuble (sofa)	black, white, brown
canapé>petit four (canapé)	many colors but no blue, no grey
culotte>sous vêtement (underpants)	any color : white, red, pink, black, ...
culotte>boucherie (rump)	no particular color, chair
motif>dessin (pattern)	any color
motif>raison (motive)	no color
papillon>insecte (butterfly)	any color, white, yellow, brown, blue
papillon>nage (butterfly stroke)	no color
papillon>contravention (ticket)	color, green, pale green
costume>tendue (suit) (most probable)	most probable colors : black, grey, dark grey, dark blue
costume>déguisement (outfit)	any color : red, yellow, etc

Table 4. Sample of colors associated to specific terms refinements (meanings) as collected with ColorIt.