

On the grammatical substance of colour words

Alina Villalva (University of Lisbon)

Portuguese colour words - those that refer a colour and those that refer the colour of something or someone - surface either as nouns or as adjectives. These colour words are lexically ambiguous with regard to their grammatical category. Syntax may disambiguate them (and it often does), but that disambiguation doesn't provide enough information to understand the relationship that holds between colour nouns and colour adjectives.

In this paper, I will explore different avenues (i.e. etymology, semantics, grammatical properties and contemporary usage) to discuss the following issues:

- From a lexical point of view, are colour nouns and colour adjectives one sole entity or two different items?
- If colour nouns and colour adjectives differ, how do they differ? Which of them comes first? How is the other one obtained?
- If they are one only lexical item, how can its different outputs be generated?
- Do diachronic and synchronic analyses converge or diverge?

The words under consideration will be *amarelo* 'yellow', *azul* 'blue', *laranja* 'orange', *verde* 'green' and *vermelho* 'red'. I will present a brief description on colour words gender and number variation, within the general picture for gender and number variation in Portuguese. This description will justify why I will just consider non-feminine singular colour nouns. In the second section, I will show the results of a colour word query on two Portuguese textual databases. They indicate that colour adjectives are more salient than colour nouns, but nor the statistics nor the record of early usage make significant contributions to the discussion of the above-

mentioned questions. The following section offers a survey of colour words in a contemporary and some early dictionaries, that, again, is unable to provide more definitive hints.

Given the accumulation of inconclusive observations, I decided to take a look at compound colour words. These compounds have a hybrid nature due to the fact that they are words built upon a phrasal sequence. Colour adjective compounds are particularly interesting. In fact, they originate in a noun phrase, which means that they are obtained by conversion. But not straightforward - the scrutiny of usage data will allow us to conclude that, irrespective of the gender and number values of their syntactic antecedent, most of these colour compound adjectives will be masculine and singular, just like simple colour nouns. Based on this, I will claim that these compound adjectives are generated by conversion of a compound colour noun, which is produced by conversion of a noun phrase headed by a simple colour noun.

The same kind of reasoning will then be applied to simple colour words.

1. Gender and number variation

Portuguese adjectives and nouns have a quite similar morphological behaviour as far as gender and number variation are concerned; yet they are not exactly identical (see Villalva 2008 for details).

Portuguese nouns must have a gender value (either masculine, feminine or underspecified). When they refer animate entities, the gender value largely corresponds to the male/female distinction (cf. 1a), but there are some mismatches, such as underspecified nouns (cf. 1b), masculine nouns that refer female beings (cf. 1c) and feminine nouns that refer both genders, as well as the opposite (cf. 1d). As to the gender of inanimate nouns, it is fully unmotivated (cf. 1e):

- | | | | |
|-----|----|---|---|
| (1) | a. | <i>aluno</i> _{masc} | ‘male pupil’ |
| | | <i>aluna</i> _{fem} | ‘female pupil’ |
| | b. | <i>(um/uma) estudante</i> _{masc/fem} | ‘(a _{male} /a _{female}) student’ |
| | c. | <i>mulherão</i> _{masc} | ‘kind of woman’ |
| | d. | <i>testemunha</i> _{fem} | ‘(male/female) witness’ |
| | | <i>cônjuge</i> _{masc} | ‘(male/female) spouse’ |
| | e. | <i>sofá</i> _{masc} | ‘couch’ |
| | | <i>poltrona</i> _{fem} | ‘armchair’ |

All nouns do require a gender value, but only a subset is prone to gender variation. In fact, variation is restricted to animate nouns, although it admits many different realizations (cf. 2a). Furthermore, it is not compulsory, not even within this subgroup of nouns, so a large number of animate nouns are invariable (cf. 2b):

- | | | | |
|-----|----|--|----------------------------------|
| (2) | a. | <i>aluno</i> _{masc} / <i>aluna</i> _{fem} | ‘male pupil’/‘female pupil’ |
| | | <i>doutor</i> _{masc} / <i>doutora</i> _{fem} | ‘male doctor’/ ‘female doctor’ |
| | | <i>estudante</i> _{masc} / <i>estudante</i> _{fem} | ‘male student’/ ‘female student’ |
| | | <i>galo</i> _{masc} / <i>galinha</i> _{fem} | ‘rooster’/‘hen’ |
| | | <i>profeta</i> _{masc} / <i>profetisa</i> _{fem} | ‘male prophet’/‘female prophet’ |
| | | [<i>águia</i> _{fem} - <i>macho</i> _{masc}] _{fem} | ‘male eagle’ |
| | | [<i>crocodilo</i> _{masc} - <i>fêmea</i> _{fem}] _{masc} | ‘female crocodile’ |
| | | <i>homem</i> _{masc} / <i>mulher</i> _{fem} | ‘man’/‘woman’ |
| | | <i>soldado</i> _{masc / fem} | ‘(male / female) soldier’ |
| | b. | <i>criança</i> _{fem} | ‘child’ |
| | | <i>indivíduo</i> _{masc} | ‘individual’ |

Colour nouns, which are inanimate nouns, are quite systematic wrt to gender. In fact, although they spread for different thematic classes, their value is systematically masculine⁵⁷:

- | | | |
|-----|----|---------------------------------|
| (3) | -o | <i>amarelo</i> _{masc} |
| | | <i>vermelho</i> _{masc} |
| | -a | <i>laranja</i> _{masc} |
| | -e | <i>verde</i> _{masc} |
| | -∅ | <i>azul</i> _{masc} |

As to adjectives, generally speaking, they may either be variable or invariable for gender, a capacity that is lexically determined, usually for diachronic reasons, and it is also related to their thematic affiliation:

- | | | | |
|-----|----|---------------------------------|--|
| (4) | a. | <i>seco</i> _{masc} | ‘X _{masc} dry’ |
| | | <i>seca</i> _{fem} | ‘X _{fem} dry’ |
| | b. | <i>leve</i> _{masc/fem} | ‘X _{masc} / X _{fem} light’ |

⁵⁷ The case of *laranja* is particularly revealing, since it corresponds to a conversion of the name of a fruit, which is feminine (i.e. *a laranja* ‘the orange’).

Colour adjectives accommodate quite easily to this state of affairs, which means that the set corresponding (3) gets a bit more crowded, once we add those forms that can only be adjectives (which are marked in grey):

- (5)
- | | |
|----|-----------------|
| -o | <i>amarelo</i> |
| | <i>vermelho</i> |
| -a | <i>amarela</i> |
| | <i>vermelha</i> |
| | <i>laranja</i> |
| -e | <i>verde</i> |
| -Ø | <i>azul</i> |

Number variation is customary for both nouns and adjectives, and it is systematically obtained by adding the suffix *-s*. There are some morphophonological processes operating on *-Ø* stems, especially when they have a final lateral consonant. They are triggered by the need to fulfil the thematic index position for phonetic reasons, related to the resulting sequence of consonants:

- (6)
- | | | |
|----|---|--|
| a. | <i>mar</i> ‘sea’ | <i>mar+s</i> → <i>mares</i> ‘seas’ |
| | <i>anel</i> ‘ring’ | <i>anel+s</i> → <i>anéis</i> ‘rings’ |
| b. | <i>capaz</i> ‘capable X _{sing} ’ | <i>capaz+s</i> → <i>capazes</i> ‘capable X _{pl} ’ |
| | <i>normal</i> ‘normal X _{sing} ’ | <i>normal+s</i> → <i>normais</i> ‘normal X _{pl} ’ |

Some nouns (e.g. *lápiz* ‘pencil’, in (7a)) and some adjectives (e.g. *simples* ‘simple’, in (7b)) are ambiguous regarding to number, due to the fact that their root ends with a fricative consonant. Consequently, in these cases, the adjunction of the plural suffix, although occurring, yields no visible result:

- (7)
- | | |
|----|---|
| a. | <i>tenho aqui um lápiz bem afiado</i> ‘I have a well sharpened pencil here’ |
| | <i>tenho aqui uns lápiz bem afiados</i> - cf. <i>lápiz+s=lápiz</i> |
| b. | <i>este motor é muito simples</i> ‘this engine is very simple’ |
| | <i>estes motores são muito simples</i> - cf. <i>simples+s=simples</i> |

A major restriction for number inflection concerns mass nouns such as *água* ‘water’. Usually, this kind of nouns can only occur in the singular form, even though a plural form is fully available (cf. 8a). Arguably, mass nouns that may occur in the plural form (cf. *águas*) are no longer mass nouns (cf. 8b):

- (8) a. *velhice* ‘old age’ vs. [?]*velhices*
 b. *Traga-me duas águas com gás* ‘bring me two (bottles of) sparkling waters’
As águas sulfurosas são terapêuticas ‘sulphur waters are therapeutic’

Being intrinsically mass nouns, colour nouns resist pluralisation, and when they allow it, plural forms refer to different sorts of the colour that the noun refers, usually hues or shades of that colour:

- (9) *O vermelho é uma cor quente* ‘red is a hot colour’
Neste quadro, os vermelhos (=tons de vermelho) predominam ‘in this painting, reds (=hues of red) are predominant’

Thus, Portuguese colour words can either be inanimate masculine mass nouns, if singular, or count nouns if plural, or adjectives that can be gender variable or invariable. They have a quite straightforward lexical behaviour, since they split over all thematic classes and they inflect for number as all other nouns and adjectives:

(10)

	MASCULINE NOUNS		ADJECTIVES	
	SINGULAR (MASS NOUNS)	PLURAL (COUNT NOUNS)	SINGULAR	PLURAL
-o	<i>amarelo</i> <i>vermelho</i>	<i>amarelos</i> <i>vermelhos</i>	<i>amarelo</i> _m <i>vermelho</i> _m	<i>amarelos</i> _m <i>vermelhos</i> _m
-a	<i>laranja</i>	<i>laranjas</i>	<i>amarela</i> _f <i>vermelha</i> _f <i>laranja</i> _{m/f}	<i>amarelas</i> _f <i>vermelhas</i> _f <i>laranjas</i> _{m/f}
-e	<i>verde</i>	<i>verdes</i>	<i>verde</i> _{m/f}	<i>verdes</i> _{m/f}
-Ø	<i>azul</i>	<i>azuis</i>	<i>azul</i> _{m/f}	<i>azuis</i> _{m/f}

Apart from the feminine form of gender variable adjectives (i.e. *amarela(s)*, *vermelha(s)*), all colour words are lexically ambiguous wrt to syntactic category. This is why we need to take a look at colour words in context.

2. Colour words in Portuguese contemporary and in early usage

In this section I will consider the real usage of Portuguese colour words. A search of colour nouns (tagged CN) on the *Corpus de Referência do Português Contemporâneo* (=CRPC) yields a set of hits similar to the following:

- (11) *azul_CN, vermelho_CN e amarelo_CN vão dominar sinalética da Expo-98*
 ‘blue, red and yellow will dominate Expo-98’s labelling’

This set of results had to be manually filtered, since they included a large number of cases that can't be treated as colour nouns. Feminine and plural forms, for instance, were systematically excluded, since, as it has been established in the previous section, they are not colour names. This is clearly the case in the following circumstance: politics and sports make a very extensive usage of colour words to identify organizations and their members. These hits (often used in the plural) were excluded because they don't refer to a colour:

- (12) **Verdes** - Denominação de um curioso partido político que tem por hábito [...] aliar-se aos **vermelhos**
 'Greens - Name of a curious political party that usually makes alliances with the reds'

Secondly, the query line 'colour name_CN' (e.g. *verde_CN*) brought hits that should have been tagged as adjectives. They occur as nouns merely superficially, due to an ellipsis of the antecedent:

- (13) *foi uma chuva de [...] cartões: ao todo, nove [cartões] amarelos_CN e um [cartão] vermelho_CN* 'it was a rain shower of [...] cards: nine yellow [cards] and a red [card], in total'
uma outra [cerâmica] vermelha_CN e grosseira 'another red and rough [china]'
devia estar a pensar nos [sacos] vermelhos_CN '[he] was probably thinking of red [bags]'

A particular case of wrong tagging involves the sequence *bué*, which is a loan of African origin, nowadays often used by younger speakers. In *Infopedia* it is listed as an adverb; in CRPC it is often tagged as a noun, but it also gets an adjective tag and even other, more exotic, options do show up. In Portuguese, *bué* is mostly used as an adverb (a proxy of *muito* 'very') and it is often accompanied by *de* or *da*, which, at first sight, look like prepositions (with the contraction of a definite article in the second case). Eventually, *bué de* or *bué da* may be analysed as adjectival quantifiers, which means that *vermelhos*, in the following example, is wrongly tagged as a noun:

- (14) *o gajo está sempre janado, meu, os olhos bué da vermelhos_CN*
 'the guy is always high, dude, his eyes are very red'

Thus, after filtering, the number of singular colour nouns is roughly as in (15). Notice that, apart from *laranja* - which clearly seems to be far less frequent, these figures look quite balanced:

(15)

NOUNS	<i>amarelo</i>	<i>azul</i>	<i>laranja</i>	<i>verde</i>	<i>vermelho</i>	TOTAL
SINGULAR	1.031	1.584	3	1.426	1.534	5.578

As for adjectives⁵⁸, no filtering was applied, although they were not all tested for their semantics. Figures in (16) show that they largely outnumber their nominal counterparts, and that singular forms are always more frequent, as well as masculine, if both gender forms are available. *Laranja* is again the least frequent case, but the distance to the other is less impressive:

(16)

ADJECTIVES		<i>amarelo</i>	<i>azul</i>	<i>laranja</i>	<i>verde</i>	<i>vermelho</i>	TOTAL
SINGULAR	MASCULINE	2.156	4.271	3.486	7.339	2.932	24.147
	FEMININE	2.018				1.945	
PLURAL	MASCULINE	1.172	2.113	2	4.407	1.507	10.431
	FEMININE	409				824	
TOTAL		5.755	6.384	3.488	11.746	7.208	34.581

So, apparently, colour adjectives are far more salient for speakers than colour nouns, but does this salience have any bearing in the questions that were initially raised? It doesn't seem to be the case. The meaning of this numbers, if any, remains unclear and the observation of colour words in syntactic contexts surely helps to disambiguate their specific uses, but it doesn't make any significant contribution to the discussion of the issues initially stated.

Regarding early usage, the output of a search on the *Corpus do Português* (=CdP) produces a set of quite similar results: *amarelo*, *azul*, *verde* and *vermelho*⁵⁹ are well documented since the late 14th century or the early 15th. Apparently,

⁵⁸ In Portuguese, all colour adjectives occur in a post-nominal position. Prenominal position is clearly unacceptable (cf. *saia azul* vs **azul saia*).

⁵⁹ *Laranja* was not scrutinized, since most hits concern the fruit, not the colour. Probably, the early colour name had a derivative form, *laranjado* (cf. *cama tambem de verde e ouro muito rica e de boa emuenção forrada de laranjado* – CdP: *Mosteiro de Guadalupe e Portugal*, 16th century); or *alaranjado* (cf. *capa lombarda de cetim alaranjado* – CdP: João de Barros (1553) *Décadas da Ásia. Década segunda*. As a colour name, *laranja* is probably the outcome of an ellipsis from the expression *cor (de) laranja* and a subsequent conversion that entails a gender change. *Laranja*, the fruit, is a feminine noun; *laranja*, the colour name, is a masculine noun.

adjectives are again more commonly used than nouns, but it is impossible to draw any conclusion from that fact.

NOUNS	ADJECTIVES
<i>as penas rosadas cõ collar amarelo esplandecente</i> CdP: <i>Orto do Esposo</i> (<1438)	<i>caeu en terra amarelo et todo banado en sange</i> CdP: <i>Cronica Troyana</i> (1388)
<i>tabernaculo muy nobremëte obrado a ouro e azul</i> CdP: <i>Crónica Geral de Espanha de 1344</i>	<i>seis Capaz de veludo azul</i> CdP: <i>Crónica de D. Fernando</i> (1431-1443)
<i>Donde lhe vem o verde na folha de fora e o branco de dentro</i> CdP: João de Lucena (1600) <i>Historia da vida do Padre S. Francisco Xavier</i>	<i>deitarõno en hûa herva verde</i> CdP: <i>Crónica de D. Fernando</i> (1431-1443)
<i>no peyto trazia hum synal de vermelho mais luzente que as estrelas</i> CdP: <i>Cronica de Portugal</i> (1419)	<i>hûu boy branco e outro vermelho</i> CdP: <i>Crónica Geral de Espanha de 1344</i>

3. Colour words in Portuguese dictionaries

The salience of colour adjectives found in CRPC parallels what can be found in Portuguese contemporary dictionaries, although this is not explicitly stated. *Infopedia*, an online reference dictionary for contemporary Portuguese, for instance, treats *azul*, *laranja*, *verde* and *vermelho*, firstly, as adjectives, and then as nouns (cf. 17); *amarelo* is presented the other way round (cf. 18). No clarification is offered wrt the relationship between adjective and noun in each case:

<p>(17) azul adjetivo de 2 géneros 1. da cor do céu sem nuvens; cerúleo nome masculino 1. cor do arco-íris semelhante à do céu sem nuvens 2. qualquer gradação desta cor</p> <p>laranja adjetivo invariável que é da cor característica deste fruto nome masculino cor resultante da adição de vermelho e amarelo</p>	<p>vermelho adjetivo 1. da cor do sangue nome masculino 1. cor do sangue</p> <p>verde adjetivo de 2 géneros 1. da cor da erva; da cor resultante da mistura do azul com o amarelo nome masculino 1. uma das cores do arco-íris, situada entre o amarelo e o azul</p>
---	--

(18) **amarelo**

nome masculino

1. cor que, no espetro solar, se situa entre o verde e o laranja, semelhante à cor do limão maduro, da gema do ovo ou do ouro

adjetivo

1. que tem a cor do limão, da gema do ovo ou do ouro

Is there any grounded reason to present these words differently? Under the influence of the haphazard semantic description of these entries (which either make reference to the ‘solar spectrum’ or to the ‘rainbow’, or to metaphors such as ‘lemon’, ‘yolk’, ‘gold’, ‘cloudless sky’, ‘grass’ and ‘blood’ and obviously ‘orange’, and even to processes, like the ‘mix of blue and yellow’ or the ‘sum of red and yellow’), we might conclude that word class is presented randomly. Although no justification is given, we might presume that the order of presentation is related to the ancestry of the colour words in the Portuguese lexicon. However, etymology doesn’t seem to be of much help either. According to Coromines, *amarillo* (an obvious cognate of *amarelo*) has probably originated in the Hispanic Latin adjective *AMARELLUS*, a diminutive of the adjective *AMARUS*; and *bermejo* (the Spanish cognate of *vermelho*) originated in the Latin diminutive noun *VERMICULUS*. So, an etymological approach would justify exactly the opposite result - not the output that is in fact presented by *Infopedia*.

A look up of a selection of early dictionaries may help to shed some light. In Cardoso (1559), *amarelo* is clearly registered as an adjective (cf. *amarela cousa*), but the other three (i.e. *azul*, *verde*, *vermelho*) are not, although they are translated into Latin adjectives. There is no mention to *laranja* as a colour word in this dictionary. Barbosa (1611) has a more systematic approach: all these colour words are presented as colour names, although they are adjectives trapped inside the expression *cor X* ‘colour X’ (cf. *cor amarela*, *cor azul*, *cor laranjada*, *cor verde*, *cor vermelha*); but, in this dictionary, each entry, except for *laranja*, is dominated by the adjectival item (cf. *amarela cousa*, *azul cousa*, *verde cousa*, *vermelha cousa*), again as translations of Latin colour adjectives. Pereira (1647) resumes Cardoso’s lesson, but he adds some more Latin equivalents. Bluteau (1712-28), who first presents colours as nouns,

supplies a considerable amount of information on how they may be obtained, and used (*laranja* is again differently treated, appearing as a ‘caste’ of ‘red’, but still a colour name, i.e. *cor de [...] laranja*). Then, in some cases, Bluteau also makes a register of colour adjectives, but this is quite non-systematically achieved. Folqman (1755) resumes the pre-Bluteau tradition that favours colour adjectives, either by showing that they are gender variable forms (cf. *amarello, a; vermelho, a*) or by explicitly stating the word class, for the first time in this set of early dictionaries (cf. *azul, adj; verde, adj*). There is no room for *laranja* in this dictionary.

Moraes (1789) tries a different solution, since he conveys two separate entries for some of these colour words: a noun entry and an adjective entry are available for *azul* and *verde*, but *amarelo* and *vermelho* are just entitled to the adjective entry. *Laranja*, as usual, is not a colour word, but the derived adjective *laranjado* re-entries in the word list. Finally, Figueiredo (1899) concentrates all the information in a single entry: *azul, verde* and *vermelho* are presented as adjectives and then, as nouns; *amarelo* is merely considered as an adjective and *laranja* is absent.

Despite the interest that a thorough analysis of all this lexicographic information⁶⁰ may have for the study of colour words, the fact is that they don’t add much in terms of categorization. No systematic and coherent picture emerges from the survey of early or contemporary dictionaries.

Cardoso 1559	Amarela cousa. Pallidus (a. um).	Azul. Caeruleus(a. um).		Verde. viridis(e).	Vermelho. Ruber(a. ũ). rubicūsus.
Barbosa 1611	Amarela cousa. Pallidus, a, um s.v. cor Cor amarela. Color pallidus.	Azul cousa. Caeruleus, a, eum. s.v. cor Cor azul. Caeruleus color. [...] aliàs cor do ceo, ou cor celeste.	s.v. cor Cor laranjada. <i>Luteus color.</i>	Verde cousa. Viridis s.v. cor Cor verde. Color viridis, color herbeus	Vermelha cousa. Ruber, rubra, rubrum. s.v cor Cor vermelha. Color ruber

⁶⁰ Early dictionary considered in this paper are available at CLP.

Pereira 1647	Amarella cousa. Pallidus, a, um. Luteus, a, um	Azul. Caeruleus, a, um. Thalastinus, a, ü.		Verde cousa. Viridis, & de. Glaucus, a, um.	Vermelha cousa. Ruber, bra, brum. Rubiumdus, a, um.
Bluteau 1712-28	amarello. Cor entre branco, & vermelho	azúl. Derivase do Arabico Zul, que segundo algüs, val o mesmo, que em Latim Caeruleus. [...] Cousa azul. Caeruleus, a, um.	s.v. vermelho [...] tres castas de vermelho [...] A segunda, que tem sua parte de amarello, como a cor de fogo, & a de Laranja s.v. Pintarroxo a cor da garganta [...] não he vermelha, mas alaranjada	verde. A cor, que a natureza dà às hervas, arvores, & c.	vermelho. Cor. Vermelho. Cousa de cor vermelha. Ruber, bra, brum
Folqman 1755	amarello, a, Flavus, a, um. [...]	azul, adj. Coeruleus		verde, adj. Viridis	vermelho, a, Ruber, bra, brum.
Moraes 1789	amaréllo, adj. Da còr da gemma de ovo, do ouro [...]	azúl, s. m. Tinta azul azúl, adj. Còr da massa extrahida do anil	laranjádo, adj. De còr de laranja.	vèrde, s.m. Huma das cores principaes vèrde, adj. Da còr do verde	vermêlho, adj. Còr do rosto corado com vergonha
Figueiredo 1899	amarelo adj. Que tem a cor do oiro, da gema de ovo, [...]	azul adj. Que tem uma das cores do espectro solar M. A cor azul.		verde adj. Que é da cor ordinária da erva e das fóllhas das árvores. M. A cor verde.	vermelho adj. Muito encarnado; rubro. M. A cor vermelha.

4. Compound colour words

Since focusing on colour words by themselves seems to reach no satisfactory answers, the focus will now be put on compound colour words, since they provide interesting insights to the understanding of the word class of their colour constituents.

The following lists include all the compounds (formed upon the five colour words initially established) that occur in Bluteau (cf. 19), in *Infopedia* (cf. 20) and in CRPC (cf. 21).

(19)

<i>amarelo gualde</i>	<i>azul ultramarino</i>	<i>verde lirio</i>
<i>amarelo tostado</i>	<i>verde azulado</i>	<i>verde mar</i>
<i>azul celeste</i>	<i>verde bexiga</i>	<i>verde montanha</i>
<i>azul ferrete</i>	<i>verde crè</i>	<i>verde negro</i>
<i>azul turqui</i>	<i>verde desmayado</i>	<i>verde terra</i>

(20)

<i>amarelo-claro</i>	<i>azul-cobalto</i>	<i>verde-água</i>	<i>verde-mar</i>
<i>amarelo-canário</i>	<i>azul-escuro</i>	<i>verde-amarelado</i>	<i>verde-musgo</i>
<i>amarelo-escuro</i>	<i>azul-ferrete</i>	<i>verde-azeitona</i>	<i>verde-negro</i>
<i>amarelo-pálido</i>	<i>azul-marinho</i>	<i>verde-claro</i>	<i>verde-velho</i>
<i>amarelo-topázio</i>	<i>azul-pavão</i>	<i>verde-escuro</i>	<i>vermelho escuro</i>
<i>amarelo-torrado</i>	<i>azul-turquesa</i>	<i>verde-esmeralda</i>	<i>vermelho-</i>
<i>azul celeste</i>	<i>azul-violeta</i>	<i>verde-gaio</i>	<i>amarelado</i>
<i>azul-bebé</i>	<i>verde alface</i>	<i>verde-garrafa</i>	<i>vermelho-púrpura</i>
<i>azul-claro</i>	<i>verde crè</i>	<i>verde-limão</i>	<i>vermelho-tijolo</i>
			<i>vermelho-vivo</i>

(21)

<i>amarelo-açafrão</i>	<i>azul-céu</i>	<i>verde-arbóreo</i>	<i>verde-salva</i>
<i>amarelo-</i>	<i>azul-cinza</i>	<i>verde-azebre</i>	<i>verde-seco</i>
<i>acastanhado</i>	<i>azul-cinzento</i>	<i>verde-azeitona</i>	<i>verde-seda</i>
<i>amarelo-água</i>	<i>azul-clarinho</i>	<i>verde-azul</i>	<i>verde-sujo</i>
<i>amarelo-alaranjado</i>	<i>azul-claro</i>	<i>verde-azul-claro</i>	<i>verde-tília</i>
<i>amarelo-avermelhado</i>	<i>azul-cobalto</i>	<i>verde-azulado</i>	<i>verdes-Tirol</i>
<i>amarelo-baço</i>	<i>azul-cobre</i>	<i>verde-azulada-</i>	<i>verde-transparente</i>
<i>amarelo-banana</i>	<i>azul-de-metileno</i>	<i>escura</i>	<i>verde-tropa</i>
<i>amarelo-canário</i>	<i>azul-do-céu</i>	<i>verde-bandeira</i>	<i>verde-turquesa</i>
<i>amarelo-castanho</i>	<i>azul-eléctrico</i>	<i>verde-bosque</i>	<i>verde-verde</i>
<i>amarelo-cera</i>	<i>azul-esbranquiçado</i>	<i>verde-branco</i>	<i>verde-verde-verde</i>
<i>amarelo-chá</i>	<i>azul-escuro</i>	<i>verde-bronze</i>	<i>verde-vivo</i>
<i>amarelo-claro</i>	<i>azul-esverdeado</i>	<i>verde-cadáver</i>	<i>vermelho-abrunho</i>
<i>amarelo-de-ovo</i>	<i>azul-faísca</i>	<i>verde-caqui</i>	<i>vermelho-</i>
<i>amarelo-desbotado</i>	<i>azul-ferrete</i>	<i>verde-castanho</i>	<i>acastanhado</i>
<i>amarelo-doce</i>	<i>azul-forte</i>	<i>verde-castanho-triste</i>	<i>vermelho-</i>
<i>amarelo-dourado</i>	<i>azul-gelo</i>	<i>verde-cinza</i>	<i>acinzentada</i>
<i>amarelo-encarniçado</i>	<i>azul-ganga</i>	<i>verde-cinzento</i>	<i>vermelho-alaranjado</i>
<i>amarelo-escuro</i>	<i>azul-glauco</i>	<i>verde-claro</i>	<i>vermelho-alizarina</i>
<i>amarelo-esverdeado</i>	<i>azul-índigo</i>	<i>verde-cobre</i>	<i>vermelho-amarelado</i>
<i>amarelo-ferrugem</i>	<i>azul-liso</i>	<i>verde-couve</i>	<i>vermelho-amarelo</i>
<i>amarelo-forte</i>	<i>azul-marinheiro</i>	<i>verde-de-prata</i>	<i>vermelho-arroxado</i>
<i>amarelo-</i>	<i>azul-marinho</i>	<i>verde-dourado</i>	<i>vermelho-arroxeadado</i>
<i>fosforescente</i>	<i>azul-metálico</i>	<i>verde-ecologista</i>	<i>vermelho-azul</i>
<i>amarelo-gema</i>	<i>azul-miosótis</i>	<i>verde-erva</i>	<i>vermelho-azulado</i>
<i>amarelo-império</i>	<i>azul-negro</i>	<i>verde-esbatido</i>	<i>vermelho-baço</i>
<i>amarelo-</i>	<i>azul-negrusco</i>	<i>verde-escarlata</i>	<i>vermelho-barro</i>
<i>incandescente</i>	<i>azul nocturno</i>	<i>verde-escuro</i>	<i>vermelho-bordeaux</i>

<i>amarelo-junquilha</i>	<i>azul-noite</i>	<i>verde-esmeralda</i>	<i>vermelho-branco</i>
<i>amarelo-laranja</i>	<i>azul-oceano</i>	<i>verde-esperança</i>	<i>vermelho-brilhante</i>
<i>amarelo-limão</i>	<i>azul-pálido</i>	<i>verde-florescente</i>	<i>vermelho-cádmio</i>
<i>amarelo-manteiga</i>	<i>azul-pardo</i>	<i>verde-fosco</i>	<i>vermelho-carmim</i>
<i>amarelo-metalizado</i>	<i>azul-pavão</i>	<i>verde-franco</i>	<i>vermelho-castanho</i>
<i>amarelo-ministério</i>	<i>azul-pervinca</i>	<i>verde-fresca</i>	<i>vermelho-cereja</i>
<i>amarelo-mostarda</i>	<i>azul-petróleo</i>	<i>verde-gaio</i>	<i>vermelho-chama</i>
<i>amarelo-negro</i>	<i>azul-piscina</i>	<i>verde-garrafa</i>	<i>vermelho-claro</i>
<i>amarelo-ocre</i>	<i>azul-prata</i>	<i>verde-gótico</i>	<i>vermelho-coca-cola</i>
<i>amarelo-ouro</i>	<i>azul-prateado</i>	<i>verde-grande</i>	<i>vermelho-desbotado</i>
<i>amarelo-palha</i>	<i>azul-profundo</i>	<i>verde-Guiné</i>	<i>vermelho-distância</i>
<i>amarelo-pálido</i>	<i>azul-prussiano</i>	<i>verde-lagarta</i>	<i>vermelho-dourado</i>
<i>amarelo-pardo</i>	<i>azul-roxa</i>	<i>verde-lama</i>	<i>vermelho-escarlata</i>
<i>amarelo-primavera</i>	<i>azul-safira</i>	<i>verde-lima</i>	<i>vermelho-escuro</i>
<i>amarelo-púrpura</i>	<i>azul-sulfato</i>	<i>verde-limoso</i>	<i>vermelho-esverdeado</i>
<i>amarelo-queimado</i>	<i>azul-tinta</i>	<i>verde-maçã</i>	<i>vermelho-força</i>
<i>amarelo-saramago</i>	<i>azul-transparente</i>	<i>verde-malva</i>	<i>vermelho-forte</i>
<i>amarelo-semáforo</i>	<i>azul-turquesa</i>	<i>verde-mar</i>	<i>vermelho-granada</i>
<i>amarelos-sépia</i>	<i>azul-ultramarino</i>	<i>verde-marinho</i>	<i>vermelho-intensivo</i>
<i>amarelo-sujo</i>	<i>azul-verde</i>	<i>verde-mate</i>	<i>vermelho-laranja</i>
<i>amarelo-terrosa</i>	<i>azul-vermelho</i>	<i>verde-musgo</i>	<i>vermelho-mínio</i>
<i>amarelo-torrado</i>	<i>azul-violácea</i>	<i>verde-negro</i>	<i>vermelho-negro</i>
<i>amarelo-tostado</i>	<i>azul-violeta</i>	<i>verde-noite</i>	<i>vermelho-ocre</i>
<i>amarelo-união</i>	<i>azul-vivo</i>	<i>verde-ocre</i>	<i>vermelho-pálido</i>
<i>amarelo-vivo</i>	<i>verde-acastanhado</i>	<i>verde-oliva</i>	<i>vermelho-púrpura</i>
<i>azul da Prússia</i>	<i>verde-ácido</i>	<i>verde-palha</i>	<i>vermelho-rosa</i>
<i>azul-acinzentado</i>	<i>verde-acinzentado</i>	<i>verde-pálido</i>	<i>vermelho-rosado</i>
<i>azul-água</i>	<i>verde-aço</i>	<i>verde-papagaio</i>	<i>vermelho-rubro</i>
<i>azul-aguado</i>	<i>verde-água</i>	<i>verde-pastel</i>	<i>vermelho-ruço</i>
<i>azul-amarelo</i>	<i>verde-alface</i>	<i>verde-pinheiro</i>	<i>vermelho-sangue</i>
<i>azul-anil</i>	<i>verde-alfazema</i>	<i>verde-pinho</i>	<i>vermelho-sanguíneo</i>
<i>azul-arroxeadado</i>	<i>verde-alga</i>	<i>verde-rã</i>	<i>vermelho-terracto</i>
<i>azul-bebé</i>	<i>verde-alvo</i>	<i>verde-relva</i>	<i>vermelho-tijolo</i>
<i>azul-berrante</i>	<i>verde-amarelado</i>	<i>verde-rosa</i>	<i>vermelho-tomate</i>
<i>azul-brilhante</i>	<i>verde-amêndoa</i>	<i>verde-rubro</i>	<i>vermelho-verde</i>
<i>azul-celeste</i>	<i>verde-anil</i>	<i>verde-salsa</i>	<i>vermelho-violáceo</i>
<i>azul celeste claro</i>			<i>vermelho-vivo</i>
<i>azul-celestial</i>			
<i>azul-cerâmica</i>			

Notice that from Bluteau's fifteen dictionary registers, ten fail to be acknowledged by *Infopedia* (cf. 22a) and eight have no register in CRPC (cf. 22b). Also notice that very few of the compounds that are listed in *Infopedia* fail to occur in CRPC. Exceptions are those in (22c).

- (22) a. *amarelo gualde*
amarelo tostado
azul turqui
azul ultramarino
verde azulado
verde bexiga

verde desmayado
verde lírio
verde montanha
verde terra
- b. *amarelo gualde*

azul turqui

verde bexiga
verde-cré
verde-desmayado
verde-lírio
verde montanha
verde terra
- c. *amarelo-topázio* ‘topaze yellow’
verde-cré ‘golden green’
verde-limão ‘lemon green’
verde-velho ‘old green’
vermelho-púrpura ‘purple red’

The number of compounds found in CRPC (i.e. 254) shows that reality goes indeed much beyond imagination, but this is a very promising result, since their morphosyntactic behaviour will probably reveal the grammatical nature of their constituents, unbiased as far as grammatical prescriptions are concerned. Most of the compounds actually used by speakers (maybe only once or twice) are not in the dictionary – they show up if they are needed and no dictionary register is felt to be required.

The categorization of these compounds is certainly one of their most interesting aspects. In *Infopedia*, different possibilities are offered: they can be listed as nouns and then as adjectives, as in (23a), (23b) and (23c); or as adjectives first and nouns afterwards, as in (23d), (23e) and (23f); or as adjectives and nouns, at the same time (cf. 23g) and (23h) and, finally, just as nouns (23i). This classification looks as arbitrary as the classification of simple colour words (cf. (17) and (18)). The largest group favours the precedence of adjectives (cf. 23d-h), but the percentage of the opposite group (cf. 23a-c) is not low.

(23) a.	amarelo-canário nome masculino. tom amarelo-claro ligeiramente esverdeado, como o de certos canários adjetivo. que apresenta esse tom	<i>amarelo-claro</i> <i>amarelo-escuro</i> <i>amarelo-pálido</i> <i>amarelo-torrado</i>	39,0%
b.	azul-claro nome masculino. tom claro do azul adjetivo. que apresenta esse tom	<i>azul-escuro</i> <i>azul-marinho</i> <i>azul-cobalto</i>	
c.	azul-bebé nome masculino. tom de azul que tende para o branco; azul-claro adjetivo invariável. que apresenta esse tom	<i>azul-ferrete</i> <i>azul-pavão</i> <i>azul-turquesa</i> <i>azul-violeta</i>	
d.	amarelo-topázio adjetivo. da cor do topázio nome masculino. 1. uma das cores do espectro solar. 2. cor do topázio	<i>verde-água</i> <i>verde-amarelado</i> <i>verde-azeitona</i>	58,3%
e.	verde-claro adjetivo. que apresenta uma tonalidade clara de verde nome masculino. essa tonalidade	<i>verde-escuro</i> <i>verde-esmeralda</i> <i>verde-gaio</i> <i>verde-garrafa</i>	
f.	vermelho escuro adjetivo invariável. que apresenta uma tonalidade escura de vermelho nome masculino. essa tonalidade	<i>verde-limão</i> <i>verde-negro</i> <i>verde-mar</i> <i>verde-musgo</i>	
g.	verde alface adjetivo invariável, nome masculino. que ou tom que é verde-claro vivo característico da alface	<i>verde -velho</i> <i>vermelho-amarelado</i> <i>vermelho-púrpura</i>	
h.	azul celeste adjetivo de 2 géneros, nome masculino. azul da cor do céu	<i>vermelho-tijolo</i> <i>vermelho-vivo</i>	
i.	verde cré nome masculino. verde tirante a ouro		2,7%

Furthermore, regarding, it is worth mentioning that the order of the listing of major categories for these colour compounds in *Infopedia* is quite inconsistent with the order followed for simple colour nouns:

(24)

N, ADJ	ADJ, N	
amarelo <i>amarelo-canário</i> <i>amarelo-claro</i> <i>amarelo-escuro</i> <i>amarelo-pálido</i> <i>amarelo-torrado</i> <i>azul cobalto</i> <i>azul ferrete</i> <i>azul pavão</i> <i>azul turquesa</i> <i>azul violeta</i>	<i>amarelo-topázio</i> azul <i>azul-celeste</i> verde <i>verde-água</i> <i>verde-alface</i> <i>verde amarelado</i> <i>verde escuro</i> <i>verde esmeralda</i> <i>verde gaio</i> <i>verde garrafa</i> <i>verde negro</i>	<i>verde velho</i> <i>verde azeitona</i> <i>verde limao</i> <i>verde mar</i> <i>verde musgo</i> vermelho <i>vermelho amarelado</i> <i>vermelho-escuro</i> <i>vermelho-púrpura</i> <i>vermelho-tijolo</i> <i>vermelho vivo</i>

Gender variation for compound colour adjectives is also problematic. Apparently, if the compound includes one variable constituent, it will be presented as a variable adjective (cf. (23a), (23b), (23d) and (23e)), but the classification of *vermelho-escuro* (cf. (23f)) as an invariable adjective is absolutely puzzling, since its two constituents are variable and also because *amarelo-escuro*, which is structurally identical, is classified as a variable adjective. Eventually, this is a mistake since, for instance, *castanho-escuro* and *cinzento-escuro* follow the model of *amarelo-escuro*. The distinction between ‘adjetivo invariável’ (in (23c) and (23g)) and ‘adjetivo de 2 géneros’ (in 23h) is equally an undesirable background noise, since they mean the same.

5. Variation in compounds – the results

The following data was extracted from CRPC, on the basis of a query line formed by one of the above colour words followed by a dash⁶¹ (-) and any other word (e.g. ‘*amarelo-**’). The search yielded more than one thousand hits, but, once again, some of the results had to be excluded, either for structural or for semantic reasons. In the first case, I have eliminated all coordinated structures (cf. 25), mainly

⁶¹ The orthography of this type of word compounds requires the use of a dash to connect both words (cf. VOP).

formed by two colour words, since these compounds are exocentric and the identification of the compound was needed for the upcoming discussion.

- (25) *cubinhos com duas cores: um vermelho-amarelo, outro azul-amarelo e um terceiro azul-vermelho*
'two-coloured little cubes: the first is red-yellow, the other is blue-yellow and a third one is blue-red'

Hits that were excluded for semantic reasons include, like for simple colour words, colour compounds that are not used to refer a colour, such as the following:

- (26) *coligação vermelha-verde* 'red-green coalition'
clube vermelho-branco 'red-white team'

The subcorpus thus obtained is formed by 875 different compounds. Their frequency will not be taken into account, since, generally, it is individually always very low.

Before looking at the results, some remarks on tagging are advisable. Tagging of a corpus such as CRPC is largely obtained automatically, which is a risky but inevitable operation. In the case of colour words, this automatic tagging is even more risky, since, as we have seen before (cf. (10)), the form of adjectives and nouns is virtually the same. Wrong tagging was identified above, in the case of simple colour words (cf. (12) and (13)), but it seriously increases when we consider compound colour words.

The first problem concerns the appearance of bizarre categories, given their syntactic context. In fact, some colour compounds are tagged as verb forms other than past participle, infinitive and gerund forms (cf. 27a), but these can also be found (cf. (27b), (27c), (27d) and (27e)). Some other colour compounds are tagged as proper names (cf. (27f)), and even as cardinals (cf. (27g)) and roman numerals (cf. 27h), although the last two are quite scarcely found.

- (27) a. **V = Verb**
barra [vermelha-escura]_V
b. **PPT = Past Participle (in compound tenses)**
O Rabisgato [...] tem os bagos dum [verde-claro]_{PPT} de certos licores

- c. **PPA = Past Participle (not in compound tenses)**
o cloro é um gás [amarelo-esverdeado]_{PPA}
de aspecto límpido e cor [amarelo-pálido]_{PPA}
- d. **INF = Infinitive**
tons intensos de [verde-mar]_{INF}
- e. **GER = Gerunds**
viu -lhe no olho [amarelo-castanho]_{GER} o círculo vermelho da orla contrair -se
- f. **PNM = Part of Name**
entrecortada [...] pelo [verde-escuro]_{PNM} dos montados
- g. **CARD = Cardinals**
os tons «camel», [verde-bronze]_{CARD} e preto predominam
- h. **DTGR = Roman Numerals**
dez classificadores de um horrível [verde-caqui]_{DTGR}

These hits (94 in total) could be disregarded, and the search would still be representative, but hits tagged as nouns or adjectives are also problematic, which means that the results are very unreliable:

- (28) a. **Nouns wrongly tagged as adjectives**
os tons característicos são o verde e o azul-acinzentado, [...] o [vermelho-tijolo]_{ADJ}
- b. **Adjectives wrongly tagged as nouns**
grãos cor [vermelho-alaranjada]_N

Another frequently mistaken tagging concerns sequences that involve two potential colour words, as in (29), for very understandable reasons: if one colour word is difficult to tag, a sequence of two has to be even harder to classify. But, in this case, the first word is not a colour word, hence the sequence is not a compound colour word:

- (29) *Só alguns espíritos mais observadores se aperceberam de um pormenor poético: **uma rosa_ADJ vermelha_CN** também lançada pelo tratador*
 ‘only a few attentive spirits realized the poetic detail: a red rose also thrown by the keeper’
*um sumo de **laranja_ADJ vermelho_CN** ‘a red orange juice’*

Mentioning this kind of mistaken tagging has a purpose, which is not to evaluate negatively the CRPC’s tagger. In fact, colour compounds are difficult to tag because they have to be morpho-syntactically analysed and that functionality is probably not there. Hence, CRPC’s POS results will be disregarded.

Thus, search results that I will now present have been manually tagged, both globally and internally. Morphosyntactic compounds, such as these colour

compounds are lexical items with a phrasal basis. So, I have pinpointed the word class of each constituent and the word class of the compound, according to its syntactic context. I have also annotated gender and number values for the antecedents of compound adjectives and gender and number values for each compound constituent. In total, this subcorpus includes 316 nouns and 559 adjectives. Unexpectedly, the asymmetry between nouns and adjectives is much more discrete now than what it was with simple colour words (cf. (15) and (16)):

(30)

nouns	singular	300
	plural	16
adjectives		559

This complex morphosyntactic tagging allows to set different configurations and it also allows to detect predominant grammatical behaviours. This is a summary of the survey of noun colour compounds:

- (31)
- a. $[[X]_{\text{singular colour word}} [Y]_{\text{singular adjective}}]_{\text{singular masculine noun}}$ (144/300)
 e.g. *amarelo-claro* ‘light yellow’, *azul-celeste* ‘sky blue’, *laranja-escuro* ‘dark orange’, *verde-pálido* ‘pale green’, *vermelho-vivo* ‘vivid red’
- b. $[[X]_{\text{singular colour word}} [Y]_{\text{singular noun}}]_{\text{singular masculine noun}}$ (127/300)
 e.g. *abrunho* ‘apricot’, *açafrão* ‘saffron’, *limão* ‘lemon’, *miosótis* ‘myositis’, *musgo* ‘moss’, *saramago* ‘wild radish’
alfaca ‘lettuce’, *alfazema* ‘lavender’, *amêndoa* ‘almond’, *banana* ‘banana’, *cereja* ‘cherry’, *coca-cola* ‘Coca-Cola’, *couve* ‘kale’, *lima* ‘lime’, *maçã* ‘apple’, *salva* ‘sage’, *violeta* ‘violet’
bronze ‘bronze’, *cádmio* ‘cadmium’, *cobalto* ‘cobalt’, *cobre* ‘copper’, *indigo* ‘indigo’, *minio* ‘minium’, *ocre* ‘ochre’, *ouro* ‘gold’, *petróleo* ‘petrol’, *sulfato* ‘sulphate’
água ‘water’, *alizarina* ‘alizarin’, *ardósia* ‘slate’, *esmeralda* ‘emerald’, *ferrugem* ‘rust’, *safira* ‘sapphire’
- c. $[[X]_{\text{singular colour word}} [Y]_{\text{singular colour word}}]_{\text{singular masculine noun}}$ (33/300)
 e.g. *azul-turquesa* ‘turquoise blue’, *verde-negro* ‘black green’, *vermelho-escarlate* ‘scarlet red’
- d. $[[X]_{\text{singular colour word}} [Y]_{\text{plural adjective}}]_{\text{plural masculine noun}}$ (1/16)
 e.g. *amarelo-acastanhados* ‘brown yellow’
- e. $[[X]_{\text{plural colour word}} [Y]_{\text{plural adjective}}]_{\text{plural masculine noun}}$ (7/16)
 e.g. *amarelos-doces* ‘sweet yellow’, *azuis-berrantes* ‘shouting blue’, *azuis-celestiais* ‘sky blue’
- f. $[[X]_{\text{plural colour word}} [Y]_{\text{singular noun}}]_{\text{plural masculine noun}}$ (7/16)
 e.g. *amarelos-ouro* ‘gold yellow’, *azuis-turquesa* ‘turquoise blue’, *verdes-cinza* ‘ash green’, *vermelhos-rosa* ‘rose red’
- g. $[[X]_{\text{plural colour word}} [Y]_{\text{plural noun}}]_{\text{plural masculine noun}}$ (1/16)
 e.g. *verdes-algas* ‘algae green’

The analysis of singular nouns must focus on the second constituent, which can either be an adjective (cf. (31a)) or a noun, referring to a fruit, a vegetable, a mineral, or anything else, in fact (cf. 31b). The third type (cf. (31c)), with a colour word as its second constituent is not a real type – I’ve isolated them because their analysis as those in (31a) or as those in (31b) is equally available. From a semantic point of view, all these compounds are hyponyms of the first constituent: *amarelo-claro* is a kind of *amarelo*; *vermelho-coca-cola* is a kind of *vermelho*. So, if these nouns are headed by a colour word that is modified either by an adjective or by a noun, the colour word has to be a noun - if in *vermelho-coca-cola*, *vermelho* was to be an adjective, then it would be a kind of *coca-cola*, which is semantically inadequate, if not for other reasons, merely because the feminine noun *coca-cola* would require the feminine adjective *vermelha*, but also because colour adjectives can not precede nouns in Portuguese (see footnote 2). Compounds in (31c) also have a colour head noun and their modifier can either be an adjective (e.g. *verde-negro* ‘black green’), similarly to those in (31a), or a noun (e.g. *vermelho-escarlata*), like those in (31b).

Plural nouns are obviously of secondary importance, simply because there are only a few of them.

Adjectives offer much more interesting information. In this case, I will set three different subtypes: agreement configurations, ‘disagreement’ configurations and irrelevant context. In the first subtype (cf. 32), one of the compound constituents (the first, the second or both) explicitly agrees with its antecedent:

- | | | |
|---------|--|---|
| (32) a. | MASCULINE SINGULAR ANTECEDENT
e.g. <i>açúcar amarelo</i> _{MASC SING} - <i>escuro</i> _{MASC SING}
<i>cabelo amarelo</i> _{MASC SING} - <i>manteiga</i> _{NOUN SING}
<i>vestido azul</i> _{INV SING} - <i>claro</i> _{MASC SING} | (144/225)
‘dark yellow sugar’
‘butter yellow hair’
‘light blue dress’ |
| b. | FEMININE SINGULAR ANTECEDENT
e.g. <i>camisa de popelina amarela</i> _{FEM SING} - <i>chá</i> _{NOUN}
<i>coloração amarela</i> _{FEM SING} - <i>esverdeada</i> _{FEM SING}
<i>camisa azul</i> _{INV SING} - <i>clara</i> _{FEM SING} | (69/201)
‘tea yellow poplin shirt’
‘greenish yellow colour’
‘light blue shirt’ |
| c. | FEMININE PLURAL ANTECEDENT
e.g. <i>ondas azuis</i> _{INV PLU} - <i>escuras</i> _{FEM PLU}
<i>t-shirts azuis</i> _{INV PLU} - <i>ganga</i> _{NOUN} | (13/79)
‘dark blue waves’
‘denim blue t-shirts’ |
| d. | MASCULINE PLURAL ANTECEDENT
e.g. <i>olhos azuis</i> _{INV PLU} - <i>claros</i> _{MASC PLU} | (8/53)
‘light blue eyes’ |

In the second subtype (cf. 33), there is no agreement between one or both of the compound constituents and its antecedent. There are no cases of ‘disagreement’ with masculine singular antecedents, which will be a relevant issue for the analysis below:

- (33) a. FEMININE SINGULAR ANTECEDENT (76/201)
 e.g. *auréola amarelo* MASC SING -*dourada* FEM SING ‘golden yellow halo’
cabeleira amarelo MASC SING -*fosforescente* INV SING ‘phosphorescent yellow hair’
camisola verde INV SING -*claro* MASC SING ‘light green sweater’
casaca vermelho MASC SING -*escuro* MASC SING ‘dark red coat’
tonalidade vermelho MASC SING -*sangue* NOUN ‘blood red hue’
- b. FEMININE PLURAL ANTECEDENT (66/79)
 e.g. *flores amarelo* MASC SING -*esverdeadas* FEM PLU ‘greenish yellow flowers’
letras azul INV SING -*celestes* INV PLU ‘sky blue letters’
manchas azul INV SING -*celeste* INV SING ‘sky blue spots’
calças azul INV SING -*escuro* MASC SING ‘dark blue trousers’
folhas verde INV SING -*escuras* FEM PLU ‘dark green leaves’
boinas vermelho MASC SING -*escuro* MASC SING ‘dark red berets’
unhas vermelho MASC SING -*sangue* NOUN ‘blood red nails’
- c. MASCULINE PLURAL ANTECEDENT (45/53)
 e.g. *tons amarelo* MASC SING -*escuros* MASC PLU ‘dar yellow hues’
panos amarelo MASC SING -*torrado* MASC SING ‘toasted yellow clths’
livrinhos azul INV SING -*bebé* NOUN ‘baby blue booklets’
poentes laranja MASC SING -*púrpura* NOUN ‘purple orange sunsets’
uniformes verde INV SING -*escuros* MASC PLU ‘dark green uniform’
paralelepípedos verde INV SING -*seco* MASC SING ‘dry green parallelepiped’
tufos verde INV SING -*mate* INV SING ‘mate green wisp’
óculos vermelho MASC SING -*azul* INV SING ‘blue red glasses’
rolos vermelho MASC SING -*rosa* NOUN ‘rose red hair pin’
- d. MASCULINE SINGULAR ANTECEDENT (0/225)

Finally, the third subtype comprises all the cases that involve singular antecedents and gender invariable colour names, invariable modifier adjectives or modifier nouns that do not agree with the colour head (cf. 34). If there is any kind of agreement, in these cases, it is not visible:

- (34) a. MASCULINE SINGULAR ANTECEDENT (81/225)
 e.g. *véu azul* INV SING -*água* NOUN ‘water blue veil’
céu azul INV SING -*celeste* INV SING ‘sky blue sky’
- b. FEMININE SINGULAR ANTECEDENT (57/201)
 e.g. *malinha azul* INV SING -*bebé* NOUN ‘baby blue bag’
cúpula azul INV SING -*celeste* INV SING ‘sky blue dome’

In sum, most of these colour compounds show agreement features (circa 41,9%), but the antecedent is predominantly a masculine singular noun (circa 61,5%). This observation must be related to the fact that there are no masculine singular antecedents when agreement fails to occur and it is still relevant to note that the ‘disagreeing’ forms within the compound are masculine and singular (cf. *boinas vermelho-escuro* ‘dark-red berets’). A similar observation is offered by the fact that compounds that have invariable forms (cf. *véu azul-água* ‘water blue veil’, *malinha azul-bebé* ‘baby blue little bag’) always have singular antecedents. This is a summary of the counting:

(35)

	ANTECEDENT				TOTAL
	SINGULAR		PLURAL		
	MASC	FEM	MASC	FEM	
AGREEMENT STRUCTURES	144	69	8	13	234
‘DISAGREEMENT’ STRUCTURES	-----	76	45	66	187
IRRELEVANT CONTEXT	81	57	-----	-----	138
	225	202	53	79	559

The corollary of all these observations seems to be that there is a tendency to assign unmarked gender and number values to these compounds. In other words, most of these compounds will have a masculine singular form, irrespective of the gender and number values of the antecedent. Notice that only 90 colour compounds exhibit full agreement with non-masculine singular antecedents (circa 16%) and probably, native speakers would accept their non-agreeing counterparts:

(36) *coloração amarela-esverdeada* vs. *coloração amarelo-esverdeado*
olhos azuis-claros vs. *olhos azul-claro*

We will now leave this line of reasoning in stand-by, to find out what is going on inside colour compounds that surface as adjectives. All the cases under consideration here are formed by a colour word plus a modifier that can be either an adjective or a noun. Apparently, the structure of these compounds is similar to the noun compounds described above (cf. 31). So, if a modifier follows the colour word, then the colour word must be a noun, but if this is so (i.e. if the colour word, which is the head of the compound, is a noun), how does the compound come to be an adjective? The hypothesis that seems more plausible is that these compound

adjectives are initially generated as nouns and then they are re-categorised as adjectives.

(37) $[[\textit{amarelo}]_{\text{NOUN}} [\textit{escuro}]_{\text{ADJ}}]_{\text{NOUN}} \rightarrow [[[\textit{amarelo}]_{\text{NOUN}} [\textit{escuro}]_{\text{ADJ}}]_{\text{NOUN}}]_{\text{ADJ}}$

Conversion processes often have a ‘freezing’ effect wrt to morpho-syntactic variation. This may be the reason why adjectival colour compounds tend to occur in the masculine-singular form (which is the form of the base noun) and tend to elope agreement with its antecedent. These adjectives are gender and number invariable forms:

(38) $[\textit{casaca}]_{\text{FEM SING}} [\textit{vermelho}_{\text{MASC SING}} -\textit{escuro}_{\text{MASC SING}}]_{\text{INV ADJ}}$
 $[\textit{boinas}]_{\text{FEM PLU}} [\textit{vermelho}_{\text{MASC SING}} -\textit{escuro}_{\text{MASC SING}}]_{\text{INV ADJ}}$
 $[\textit{panos}]_{\text{MASC PLU}} [\textit{amarelo}_{\text{MASC SING}} -\textit{torrado}_{\text{MASC SING}}]_{\text{INV ADJ}}$

Finally, we need to account for the cases of compound colour adjectives that do not behave as invariable adjectives, since they exhibit partial (cf. 39a) or full (cf. 39b) agreement features.

(39) a. $[\textit{auréola}]_{\text{FEM SING}} [\textit{amarelo-dourad}]_{\text{FEM SING}} \textit{a}_{\text{FEM SING}}$
 $[\textit{flores}]_{\text{FEM PLU}} [\textit{amarelo-esverdead}]_{\text{FEM PLU}} \textit{as}_{\text{FEM PLU}}$
 $[\textit{tons}]_{\text{MASC PLU}} [\textit{amarelo-escur}]_{\text{MASC PLU}} \textit{os}_{\text{FEM PLU}}$
 b. $[\textit{coloração}]_{\text{FEM SING}} [\textit{amarela}]_{\text{FEM SING}} [\textit{esverdeada}]_{\text{FEM SING}}$
 $[\textit{ondas}]_{\text{FEM PLU}} [\textit{azuis}]_{\text{INV PLU}} [\textit{escuras}]_{\text{FEM PLU}}$
 $[\textit{olhos}]_{\text{MASC PLU}} [\textit{azuis}]_{\text{INV PLU}} [\textit{claros}]_{\text{MASC PLU}}$

Certainly, partial and full ‘agreement’ cases have different implications. Partial agreement, which affects the last constituent, is probably due to a more consolidated stage of lexicalization. Frozen forms, such as those in (38) are the direct output of conversion; adjectives that show gender and number agreement at the right-hand periphery are lexicalised forms that have ‘lost’ their original internal structure. Full agreement, on the other hand, seems to reveal a completely different word-formation strategy. In this case, we may consider that the compound is formed upon a simple colour adjective by coordination, which is a plausible explanation for cases such as *coloração amarela esverdeada* ‘greenish yellow colouring’.

Notice that a few cases of partial agreement come from compounds that have a modifier noun (e.g. *camisa amarela-chá*, *t-shirts azuis-ganga*). The coordination hypothesis to explain these cases is simply inadequate: sequences of nouns and adjectives can't be coordinated, nor can sequences of adjectives, unless they have cumulative reading, which is not the case. No other good explanation seems to be available, which may suggest that these cases are occasional deviations.

6. Closing remarks

Portuguese colour words, such as *amarelo*, *azul*, *laranja*, *verde* and *vermelho*, are more frequently used as adjectives, although they may also occur as nouns. Lexicographic descriptions, from early to contemporary dictionaries also tend to privilege the description of these words as adjectives, although descriptions are quite inconsistent. Etymology can't be of much help either, since some colour words originate in nouns (e.g. *azul*) and other originate in adjectives (e.g. *amarelo*). Thus, the analysis of colour words doesn't provide enough information to decide about their word class status.

Compound colour words present a quite interesting set of properties that may help to better understand the former. Compound colour nouns are formed by a simple colour noun and a modifier (an adjective or a noun). Like simple colour nouns, they are typically masculine nouns. Compound colour adjectives provide some more interesting information. In general, these adjectives take the masculine singular form even when they have other kinds of antecedents, which may be explained as an outcome of conversion, the word formation process used to produce them. If this is the case, compound colour adjectives are obtained from compound colour nouns, which are also obtained by conversion of a noun phrase headed by a colour noun.

So, those questions that were initially put can now be replied. The observation of the grammatical behaviour of simple and compound colour words supports the hypothesis that colour nouns and colour adjectives are different lexical items, and also that adjectives are obtained by conversion of colour nouns.

The small set of compound colour adjectives that exhibits agreement on the rightmost constituent may suggest that other colour word formation strategies may apply, although marginally for the time being.

References

- Barbosa, A. (1611) *Dictionarium Lusitanico Latinum*. Bracharae: Fructuosi Laurentij de Basto.
- Bluteau, R. (1712-1728) *Vocabulario Portuguez e Latino* [...]. Coimbra-Lisboa: Collegio das Artes da Companhia de Jesu.
- Cardoso, J. (1569-1570) *Dictionarium Latinolusitanicum*. Conimbricae: Joan Barrerius.
- CdP = Davies, Mark & Michael Ferreira (2006-) *Corpus do Português: 45 million words, 1300s-1900s*. [www.corpusdoportugues.org] (01-02-2016).
- Corominas, J. & J. Pascual (1992) *Diccionario Crítico Etimológico Castellano e Hispánico*. Madrid: Gredos.
- CLP - *Corpus Lexicográfico do Português* [clp.dlc.ua.pt/DICIweb/] (01-02-2016).
- CRPC - *Corpus de Referência do Português Contemporâneo* [alfclul.clul.ul.pt/CQPweb/].
- Figueiredo, C. de (1913) *Novo Dicionário da Língua Portuguesa*. Porto: Typ. da Empr. Litter. e Typographyca.
- Folqman, C. (1755) *Diccionario Portuguez, e Latino*. Lisboa : na Officina de Miguel Manescal da Costa
- Infopédia - Dicionário da Língua Portuguesa com Acordo Ortográfico* [em linha]. Porto: Porto Editora, 2003-2016. [www.infopedia.pt/dicionarios/lingua-portuguesa/dicionário] (01-02-2016).
- Morais e Silva, A. (1789) *Diccionario da Lingua Portuguesa*. Lisboa: na Typographia Lacerdina. [archive.org/details/diccionariodalino1mora & archive.org/stream/diccionariodalino2mora] (01-02-2016).
- Pereira, B. (1647): *Thesouro da lingua portugueza*. Eborae : Typographia Academiae.
- Villalva, A. (2008). *Morfologia do Português*. Lisboa: Universidade Aberta.
- VOP – *Vocabulário Ortográfico do Português*.
[http://www.portaldalinguaportuguesa.org/vop.html] (01-02-2016).