

CORPOS E CORES: COLORINDO A DESCRIÇÃO DA LÍNGUA PORTUGUESA

Cláudia Freitas¹
Diana Santos²
Rosário Silva³

RESUMO

Na senda de nossos estudos anteriores, apresentamos aqui um quadro das diferenças de uso entre as cores nas variantes brasileira e portuguesa do português, com ênfase nos verbos. Após uma panorâmica dos recursos utilizados e da filosofia e prática de anotação subjacente, apresentamos as pesquisas sobre palavras de cor e seus argumentos, de uma forma didática e convidando os leitores a repetir as buscas ou a efetuarem outras no nosso material. No artigo apresentamos também o corte-e-costura, a ferramenta de apoio à anotação do projeto AC/DC, e o tipo de regras utilizadas. Focando a nossa atenção na classe gramatical dos verbos, concluímos que estes tendem a perder o seu significado originalmente colorido e a especializarem-se em sentidos muitas vezes metafóricos ou restritos a conotações ou mesmo expressões idiomáticas. É além disso nesse sentido, principalmente, que observamos diferentes usos nas duas variantes.

PALAVRAS-CHAVE: Linguística com corpos, cores, português, variantes do português, verbos

ABSTRACT

Following up previous studies on the subject of colours in Portuguese and possible differences between the two variants, we consider in detail the verbs conveying colour in this paper. After presenting the resources and the underlying annotation philosophy and practice, we submit to the reader a set of searches on colour verbs and their arguments, showing how they can be replicated or modified for further study. We also briefly present *corte-e-costura*, the annotation tool in the AC/DC project and the corresponding rules. Our focus here being the verbs, we concluded that they tend to lose their originally visual import and specialize in metaphorical meanings, connotatively laden expressions, or even idioms, and that this trend is the main reason for differences among the two varieties investigated.

KEYWORDS: Corpus linguistics, colours, Portuguese, Portuguese varieties, verbs

1. Introdução

A questão das palavras de cor, além do seu interesse intrínseco, linguístico e filosófico (repare-se que além de ilustrarem a atribuição de propriedades (e não apenas ações ou objetos), são prototipicamente aplicadas a referentes concretos), tem uma particularidade que a torna especialmente interessante para estudos com corpos: são suficientemente comuns para serem estudadas quantitativamente. E além de comuns, ou talvez por isso, são encontradas em muitíssimas expressões mais ou menos convencionais, em que perdem o seu significado básico e exprimem outros valores bem afastados do campo visual, tal como atitude, conotações, assim como refletem (ou podem refletir) usos e artefatos antigos (*passadeira vermelha/tapete vermelho, papel pardo, páginas amarelas*).

Com efeito, do ponto de vista de frequência, alguns adjetivos de cor são dos adjetivos mais frequentes, mas interessa saber até que ponto o são porque a cor é

¹ Doutora, professora do Departamento de Letras da PUC-Rio e colaboradora da Linguateca. Email: claudiafreitas@gmail.com

² Doutora, professora associada do departamento de línguas, literaturas, e culturas europeias da Faculdade de Letras da Universidade de Oslo, líder da Linguateca. d.s.m.santos@ilos.uio.no

³ Mestre, colaboradora da Linguateca. Email: mrosariomsilva@sapo.pt

fundamental na língua portuguesa, ou o são porque tomaram muitos mais objetivos e sentidos além do prototípico (como é o caso de palavras como *gente* ou *coisa*, que se tornaram genéricos, ou *pé* ou *mão* que estão no caminho da gramaticalização em locuções como *ao pé de* (variante de Portugal) ou *à mão*).

Por outro lado, é possível que o número distinto de expressões convencionais com uma dada palavra de cor ateste a idade dessa cor na língua, e/ou a produtividade do uso dessa forma, assim como é possível que ateste diferenças ou aumento de sentidos.

Gill Philip (2011) ilustra de forma muito clara todo o espectro de questões semânticas na língua, desde colocação, expressões idiomáticas e metaforização até ao uso criativo da língua, usando a cor nos corpos.

Como será referido na nossa resenha sobre trabalhos anteriores, na secção 1.1, a cor tem também já uma longa história de tratamento linguístico e estudo, mesmo baseado em corpos. Neste trabalho, damos continuidade à exploração das cores em corpos da língua portuguesa, ao mesmo tempo que, aproveitando as possibilidades de pesquisa em corpos das variantes do Brasil e de Portugal, contrastamos também o uso nas duas variantes.

De maneira geral, este trabalho pode ser entendido como uma continuação de Santos et al. (2011), que, ao apresentar dados que sugerem um maior uso das cores na variante portuguesa, levanta questões sobre o uso das cores no Brasil e em Portugal.

Partindo de uma investigação sobre a distribuição das categorias gramaticais nas duas variantes, comparamos, em um primeiro momento, a distribuição das cores por categoria gramatical. Além disso, tirando proveito de corpos ricamente anotados com informação linguística, exploramos a distribuição das cores por categoria sintática (e por variante). O objetivo do presente trabalho é duplo: por um lado, buscamos perceber como se relacionam cor e categoria gramatical nas variantes PT e PB; por outro, tentamos prover uma descrição detalhada do comportamento dos verbos de cor.

1.1 A cor nos estudos linguísticos

O estudo das cores ocupa um papel importante no debate sobre universalismo e relativismo, interessando a diferentes áreas do conhecimento. Como possui aspectos biológicos e linguísticos, é natural que o campo das cores seja de especial relevância nos estudos da/sobre a linguagem (Deutscher, 2010). No entanto, como notam Santos et al. (2011), boa parte dos estudos sobre a cor, defendendo a universalidade da conceptualização da cor (Berlin & Kay, 1969; Rosch, 1975) ou, pelo contrário, refutando-a (Wierzbicka, 1990), tomam como base experiências com informantes. Como também aponta Lucy (1997), as pesquisas sobre as cores têm se concentrado na comparação entre as línguas, em um refinamento da tipologia e no reforço de argumentos de base biológica, enquanto relativamente pouco tem sido feito para melhorar a qualidade da descrição linguística. Saunders (2005), numa perspectiva filosófica, denuncia a base biológica implícita de Berlin & Kay (1969). Nesse contexto, trabalhos com base em corpos podem oferecer um bom complemento no que se refere ao comportamento das cores nas línguas. Com relação ao português, além da exploração que vimos desenvolvendo no âmbito do AC/DC (Santos et al., 2011), do CorTrad (Teixeira et al., 2012) e do COMPARA (Frankenberg-Garcia & Santos, 2002) relativa à cor (Inácio et al., 2008, Silva et al., 2008, Santos et al., 2008, Santos et al., 2011), referimos também Biderman et al. (2007), que parte igualmente de corpos para explorar e descrever as diferenças na expressão da cor no Brasil e em Portugal. Ainda para o português, remetemos o leitor para os trabalhos de Bacelar do Nascimento et al. (1996), Correia (2006) e Farias & Marcuschi (2006), que se debruçam sobre diferentes questões relacionadas com a cor, e Jorge et al. (2003) e

Zavaglia (2006, 2007), que tratam o fenômeno de uma perspectiva contrastiva e/ou lexicográfica – veja-se também Philip (2003) para o contraste entre o inglês e o italiano. Finalmente, numa ótica diacrônica, Sletsjøe (1962) e Swearingen (1990) dissertam sobre, respetivamente, o conceito de verde aplicado a olhos e o da diferenciação dos termos vermelho e roxo em português.

2. Contexto

Nossa exploração de corpos toma por base o projeto AC/DC (Acesso a Corpos / Disponibilização de Corpos), que disponibiliza corpos do português, nas variantes do Brasil e de Portugal, na Internet, e contém atualmente cerca de 374 milhões de palavras (Costa et al, 2009, Santos, no prelo). Todo o material foi anotado automaticamente pelo analisador sintático PALAVRAS (Bick, 2000), tendo algumas partes passado por revisão humana. Além da anotação do PALAVRAS, os corpos do AC/DC também vêm recebendo anotação relativa à informação semântica no campo das cores (Mota & Santos, 2009) como mencionado na secção anterior. A secção seguinte apresenta brevemente as opções linguísticas subjacentes à anotação das cores, bem como problematiza os casos mais difíceis. Para a documentação detalhada do trabalho de anotação linguística das cores, consulte-se o Arco-Íris (Silva e Santos, 2012).

2.1. A anotação semântica

Como mencionado, todos os corpos do AC/DC possuem informação semântica relativa à cor, presente no atributo “sema”. Atualmente, são 7 as classes de cor (veja secção 2.2 sobre o processo de anotação):

SEMA	Explicação	Exemplo
[sema="cor"]	Cor pura, representação de atributos visuais	sapatos vermelhos ; nuvens cinzentas
[sema="cor:humana"]	cores que correspondem a atributos naturais humanos	cabelos ruivos , ou louros ; corar de vergonha
[sema="cor:vinho"]	Cores associadas aos tipos de vinho	vinho branco , tinto do Douro; espumante rosé
[sema="cor:raça"]	Cores usadas para representar raça	os pele-vermelha ; brancos , negros e pardos
[sema="cor:política"]	Cores associadas a partidos ou ideias políticas por metonímia	deputados verdes ; socialismo vermelho
[sema="cor:equipa"]	Cores associadas a equipas (geralmente de futebol) por metonímia	torcida alvinegra ; defesa encarnada
[sema="cor:original"]	Usos não convencionais; não representativos de atributos visuais (e que não se enquadram nos anteriores)	período negro ; concordata branca ; vida cor de rosa

Quadro 1: Classes de cores no AC/DC

A classe original é atribuída quando estamos diante de usos não relacionados a propriedades visuais – quando o vínculo com a propriedade colorida se perdeu, total ou parcialmente. Naturalmente se pode argumentar que nos usos de [sema="cor:política"] e [sema="cor:equipa"] o vínculo com a cor também pode estar distante, mas o uso da cor nesses domínios é sistemático o suficiente para nos

permitir agrupá-los sob um rótulo. No caso de [sema="cor:original"] (que admitimos talvez não ser o melhor nome, mas que provém do seguinte raciocínio “teve origem em cor”), agrupamos tanto (i) usos metafóricos de uma cor, separados ou englobados em expressões fixas (ficar *no vermelho*; *dar branco*); (ii) termos que originalmente contêm uma palavra de cor, mas que são considerados unidades, e que não se referem sobretudo a cor (*buraco negro*; *elefante branco*); (iii) termos que, sendo ainda possivelmente descritivos, são usados metaforicamente (*sinal verde*; *levou um cartão vermelho* (\neq *encontrei um cartão vermelho na mesa com um poema* (que é considerado [sema="cor"])). Além disso, (iv) usamos [sema="cor:original"] quando estamos diante de usos não previstos, que podemos apelidar de criativos, das palavras de cor, como ilustram os trechos (1-4) a seguir:

- (1) Se definem como socialistas (ou «socialismo **moreno**», como diz Brizola) , mas tem forte apelo populista .
- (2) Ainda que tingido do `socialismo **moreno**» do autor de "O Povo Brasileiro" , o livro ...
- (3) Impressionado com o bronzeado dos colegas, o baiano Jaques Wagner (PT-BA) , desbotado pelas viagens à Alemanha, Coréia do Sul e Inglaterra, disse: ` O socialismo **moreno** do Brizola não deu certo, mas pelo visto o neoliberalismo **mulatinho** do Fernando Henrique vai bem» .
- (4) Sem a ajuda soviética e vítima do bloqueio americano, o sonho do socialismo **moreno**, que tanto encantou a esquerda brasileira, parecia estar fadado a um final infeliz.

Se *moreno* nos exemplos (1-4) é cor tanto quanto o seriam "socialismo vermelho" ou "socialismo verde", por exemplo, então poderíamos atribuir-lhe a etiqueta [sema="cor:política"]. No entanto, não temos a cor *morena* sistematicamente associada a nenhum campo da política em português, e portanto anotamos como [sema="cor:original"], o mesmo se aplicando a "mulatinho" na frase (3). (E notamos que a etiqueta [sema="cor:humana"] jamais seria atribuída nesses casos.)

Além das classes semânticas, as cores (quando sema="cor") também são classificadas quanto ao grupo:

Grupos de cor: BRANCO, PRETO, AZUL, AMARELO, VERMELHO, LARANJA, VERDE, ROXO, CASTANHO, CREME, CINZENTO, ROSA, PRATEADO, DOURADO e também OUTRAS.*, MÚLTIPLA.*, AUSÊNCIA, NÃOESPECIFICADA e DESCONHECIDA.

O grupo Outras engloba expressões cuja cor pareceu impossível identificar ou definir, tais como *cor de apoplexia*; *cor de crime e traição*; *cor de ferro velho*, *cor de morte* etc. Além disso, usamos também o grupo Outras quando não é possível decidir sobre (e portanto anotar) a inclusão de uma cor em um determinado grupo. *Pardo*, por exemplo, é castanho/marrom ou cinzento? E *cor de telha*?

É importante esclarecer que a relação entre os grupos e as classes de cor não é de complementação, e não há dupla categorização. A classificação de grupo de cor só existe se a ocorrência de cor for do tipo [sema="cor"], isto é, se estivermos diante de usos descritivos, coloridos. Dessa forma, em *sorriso amarelo*, o amarelo não pertence ao grupo Amarelo, visto ser [sema="cor:original"], e o mesmo para *vinho verde*, classificado como [sema="cor:vinho"]. Estamos conscientes, no entanto, de que algumas das classes também são descritivas, como [sema="cor:humana"] (em *olhos verdes*, não há como negar a propriedade colorida de “verde”), e uma possibilidade para o futuro é adicionar grupos de cores específicas para a classe humana, por exemplo.

As motivações para a existência do grupo atual de cores (com 19 elementos) são, sobretudo, empíricas. Tomamos como critério o uso variado e frequente de um termo para uma cor/grupo de cor, pois isto significa que, de alguma maneira, esse grupo/cor é relevante na língua. No grupo *Creme*, por exemplo, são 18 as instâncias de cores, como bege; bege-areia; creme; marfim; perolado; pérola etc., o que o torna merecedor de um grupo próprio. Além disso, a busca por um “grupo” é mais prática que uma busca por lema (vide os 18 lemas de *Creme*). Vale lembrar, no entanto, que o grupo não foi fixo *a priori*, e que a exploração das cores nos corpos pode levar à criação de novas categorias/grupos de cores.

2.2 A ferramenta de anotação (corte-e-costura)

O corte-e-costura (Mota & Santos, 2011, Santos & Mota, 2012) é uma ferramenta desenhada para ajudar a anotação dos corpos, a partir do desenvolvimento de regras sucessivamente mais específicas, para obter cobertura total, pondo o anotador no ciclo: ou seja, desenvolvendo um conjunto inicial de regras, aplicando-as, revendo o resultado, e continuando num processo iterativo até considerar toda a anotação correta, e abrangendo todos os casos.

É baseado num léxico, quer de palavras simples quer de expressões com várias palavras, e na aplicação de cinco tipos de regras, que em seguida exemplificamos para a cor: 1) regras positivas adicionando casos de cor que não estão no léxico, 2) regras negativas removendo casos que não deviam ser considerados cor; 3) regras de especialização que transformam casos de cor noutros mais específicos; 4) regras que retiram a marcação de casos específicos; e 5) regras recursivas (cuja aplicação só pára se não puderem continuar a ser aplicadas). Além disso, para cada tipo de regras existe a distinção entre regras gerais e regras específicas associadas a cada corpo, que fica ao critério do anotador, conhecendo o conjunto dos corpos do AC/DC, decidir.

Abaixo temos um exemplo de uma regra de cada tipo: 1) a palavra *celeste* é considerada cor se antecedida pelas palavras *bicicleta*, *divã* ou *vestimenta*; 2) a palavra *louro* não é palavra de cor quando na expressão *louros da vitória*; 3) a palavra *branco* refere-se a vinho se estiver seguida por *tinto*, e passa pois de simples cor a cor:vinho; 4) a palavra *tricolor* que tinha sido atribuída à classe equipa volta a ser apenas “cor” (no caso específico de um dado corpo); e 5) se as palavras *açó*, *banana*, *borgonha* e *café* no singular estiverem seguidas da conjunção *e* e de uma palavra de cor, são consideradas cor.

```
a:[lema="bicicleta|divã|vestimenta"] b:[lema="celeste"] >> b:[sema="cor"]

a:[word="louros"] b:[word="da|do"] c:[word="triunfo|vitoria"] >> a:[sema="0"]

a:[lema="branco"] b:[word=","] c:[lema="tinto"] >> a:[sema="cor:vinho"]

[sema="cor:equipa" & lema="tricolor"] >> [sema="cor"]

a:[lema="açó|banana|borgonha|café" & pessnum="S"] b:[lema=","|e|ou"] c:[sema="cor"] >>
a:[sema="cor"]
```

4. Exploração das cores nos corpos

No presente trabalho, tomamos por base principalmente os corpos CONDIVport (Silva, 2008) e CHAVE (Rocha & Santos, 2007).

O CONDIVport, criado com o objetivo de permitir o estudo da convergência e divergência do português entre as variantes do Brasil e de Portugal, contém cerca de 5

milhões de palavras distribuídas em jornais desportivos e revistas de saúde e de moda – áreas que tendem a empregar cores de maneira recorrente. Além disso, é importante mencionar que toda a anotação das cores no CONDIV passou por revisão humana.

O CHAVE contém cerca de 99 milhões de palavras distribuídas em textos jornalísticos da Folha de São Paulo (Brasil) e do jornal Público (Portugal). Por ser um corpo maior, e de conteúdo mais geral, oferece dados complementares aos obtidos no CONDIVport, mas ainda não foi objeto de revisão humana.

4.1 Categorias gramaticais

A tabela 1 apresenta a distribuição das cores – sem distinção de tipo – por categoria gramatical nos corpos CHAVE e CONDIV. Como é possível observar, no CHAVE, não há qualquer diferença na distribuição das cores por categoria gramatical entre as variantes BR e PT, com uma ocorrência maior de adjetivos coloridos se comparados com os verbos. No CONDIV, os dados relacionados aos verbos coloridos também são idênticos entre as variantes, e por sua vez são menos frequentes que os verbos no CHAVE. Com relação aos adjetivos de cor, a situação no CONDIV é um pouco diferente, havendo mais adjetivos coloridos na variante PT do que na variante BR. Tal fato, no entanto, conforme explicado em Santos et al., (2011), pouco tem a ver com a distribuição de adjetivos, ou com o uso das cores, mas antes se deve à composição do corpo CONDIV: a porção BR do corpo contém muitos moldes de cor – e menos texto, em um sentido estrito – o que explica a diferença nos números.

	CHAVE		CONDIV	
	ADJ cor	Verbos cor	ADJ cor	Verbos cor
BR	0,6%	0,2%	2,2%	0,06%
PT	0,6%	0,2%	3,4%	0,06%

Tabela 1: Distribuição das cores por categoria gramatical e variante

4.2 Exploração dos verbos coloridos

Com relação aos verbos coloridos, nosso principal interesse nesse trabalho, constatamos que a imensa maioria está na forma de particípio, o que não surpreende, dado que o particípio frequentemente assume o papel de modificador. No entanto, como as formas participiais ainda não foram totalmente revistas no CHAVE (Douro (rio e região), e dourado (peixe), por exemplo, ainda estão como particípios passados do verbo *dourar*) optamos por não considerá-las neste trabalho.

A tabela 2 contém todos os verbos coloridos, excluindo o particípio, contidos no AC/DC⁴, em ambas as variantes. Os lemas marcados com ** ocorrem no CHAVE e no CONDIV, e portanto serão considerados no presente estudo. Também na tabela 2 indicamos a distribuição dos lemas por variante, independente de corpo, e notamos que, em termos gerais, há um uso mais frequente de verbos de cor na variante de Portugal. Embora não esteja na tabela, mas seja de fácil aferição na página do AC/DC, percebemos que os verbos presentes apenas na variante PT vêm, em grande

⁴ À data do presente trabalho. Como a anotação e sua revisão ainda estão em progresso, poderemos ainda encontrar mais casos no futuro, assim como modificar os dados quantitativos.

parte, do corpo Vercial, que contém obras literárias de autores portugueses, cujas datas de publicação variam desde 1500 a 1933⁵.

Grupo de cor	Lema e distribuição dos verbos por corpo e variante	TOTAL por variante (de todos os corpos)					
		CHAVE		CONDIV		PT	BR
		BR	PT	BR	PT	PT	BR
Grupo Vermelho	avermelhar**; vermelhejar; vermelhar, purpurar, revermelhar, purpurejar, enrubescer	4	4	2	10	50	14
Grupo Laranja	alaranjar**	1	6	---	---	17	2
Grupo Amarelo	amarelar**, amarelecer**, amarelejar	7 2	1 14	1 ---	1 1	14 67	16 3
Grupo Roxo	arroxear**, roxear	1	1	---	---	2	---
Grupo Preto	enegreecer**, negrejar	2	13	---	4	120	6
Grupo Verde	esverdear**; reverdecer**; verdejar**; esverdear; esverdinhar; verdecer	--- 1 ---	8 1 1	--- 1 ---	1 --- ---	25 69 86	3 6 2
Grupo Cinzento	acinzentar**	4	6	----	---	22	5
Grupo Rosa	rosar**, rosear	---	2	---	---	---	2
Grupo Dourado	dourar**/doirar**, sobredourar, sobre- dourar; sobre- doirar	51	50	4	3	857	64
Grupo Branco	branquear**; embranquecer**; branquejar; esbranquiçar; alvidecer	4 7	99 2	5 2	9 5	431 46	12 24
Grupos Castanho, Creme, Prateado e Azul	sem verbos nas formas finitas no CHAVE e CONDIV; pratear; azular; acastanhar, amarronzar	---	---	---	---		

Tabela 2: Distribuição dos verbos coloridos nos corpos CHAVE e CONDIV, sem particípio⁶.

⁵ O corpo Vercial contém obras digitalizadas pelo projeto Vercial, <http://alfarrabio.di.uminho.pt/vercial/>, e de fato podemos considerar que, exceto nos séculos XIX e XX, não faz sentido distinguir os autores “portugueses” dos “brasileiros”. Um estudo mais fino deverá ser feito, reduzindo os escritores portugueses ao período moderno, e não entrando em conta com os outros, ou considerando alguns autores antigos como pertencendo às duas ou mesmo à variante brasileira, como poderia ser o caso do Padre António Vieira ou do Judeu.

⁶ Para realizar a busca por lema no AC/DC, após a escolha do corpo desejado para a pesquisa, a expressão de busca é (o exemplo refere-se à busca pelos verbos em forma finita pertencentes ao grupo Vermelho): [sema="cor.*" & grupo="Vermelho" & temcagr!="PCP" & pos="V.*"]. Para a distribuição por corpo, uma vez selecionado o corpo TODOS, a expressão de busca é (o exemplo refere-se à busca da distribuição, pelos corpos, do lema “branquear”, na variante PT):

Uma conclusão possível, e que pode ser explorada futuramente com a análise detalhada dos verbos de cores no Vercial, é a perda gradual da propriedade descritiva dos verbos de cores – como se fossem tomando a forma de adjetivos/participios e convencionalizando-a em direções mais concretas.

4.3 Os usos das cores

Nesta seção relatamos os dados obtidos na exploração dos usos dos verbos de cor. Especificamente, buscamos os argumentos associados aos verbos coloridos, nas posições de sujeito e de objeto.

Com relação ao grupo Dourado, o primeiro dado a chamar a atenção é que quase metade das ocorrências de *dourar/doirar*, em ambas as variantes, não é colorida, mas se refere à expressão “dourar/doirar a pílula”, em um uso claramente original.

A tabela 3 apresenta detalhadamente os usos de *dourar/doirar*. Como se pode perceber, são (quase) todos não descritivos: doura-se a *imagem*, a *perspectiva*, o *perdão*, o *brasão*. A única ocorrência que talvez possa ser considerada descritiva, isto é, colorida, diz respeito a “dourar a sala”, em “.. o sol doira a sala...”. No entanto, consideramos esse (mais) um exemplo de difícil decisão, que pode ser visto tanto do ponto de vista da cor (*o sol deixar a sala da cor dourada*) como um uso não colorido, (*o sol ilumina a sala*). Não há como ter certeza da leitura, e consideramos, portanto, a vagueza uma propriedade inerente e importante da linguagem, que não deve ser eliminada ou vista como imperfeição. Como a língua não nos obriga a escolher uma única interpretação, não o faremos. Especificamente com relação à anotação, estamos diante de um exemplo em que o verbo recebe as duas informações semânticas, simultaneamente: será considerado tanto [sema="cor"] (para a interpretação colorida), quanto [sema="cor:original"] (para a interpretação não colorida).

Na variante BR, a situação é idêntica, em que se doura, além da pílula, a *imagem*, a *ignorância* e a *personalidade*, entre outros. Assim como na variante PT, o *sol* doura coisas como *paisagem*, *cidade* e *gente*. E, diferente da variante PT, no Brasil também se doura *alho*, *bacon*, *cebola*, etc, isto é, temos o uso associado ao domínio culinário, e, portanto, um uso original – o que não acontece na variante PT, em que o verbo associado a tais ações é *alourar*⁷.

Com relação aos sujeitos de *dourar*, são pouco frequentes, dada a alta ocorrência da forma infinitiva “dourar a pílula”. Ainda assim, na parte portuguesa do CHAVE apenas nomes próprios (de pessoas e instituições) exercem a função de sujeito, o que reforça a ideia de um uso original. Na variante BR, além de nome próprio, são *mestres*, *estatutos*, *virtudes*, *retórica* e *jornal* aqueles que *douram*, mais uma vez em um uso original. Ainda na variante BR, *discos de massa*, *carne* e *grill douram*, repetindo-se o uso original vinculado à culinária, e encontramos apenas uma ocorrência de um uso possivelmente colorido, em que a ação de dourar é feita pelo *sol*, como já mencionado⁸.

[sema="cor.*" & lema="branquear" & temcagr!="PCP" & pos="V.*" & variante="PT"], e nos resultados deve-se pedir “Distribuição de corpo”.

⁷ Por outro lado, não podemos desconsiderar o fato de haver muito mais notícias sobre comida/culinária no jornal Folha de S. Paulo que no Público – o que não invalida nossa explicação. São apenas 6 os casos de “alourar” na parte PT do CHAVE, e 1 na parte BR (ou 2, se considerarmos também “alourar”).

⁸ Para realizar a busca por lema dos sujeitos de *dourar* ou *doirar* no AC/DC, após a escolha do corpo desejado para a pesquisa, a expressão de busca é (o exemplo refere-se à busca apenas na

Lema dos objetos de <i>dourar/doirar</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
pílula* **	12	-	9	1	original
imagem* ** ; perspectiva** ; brasão** ; asa** ; sala ** ; chamado** ; perdão** ; ignorância* ; fisiologia* ; imagem* ; década* ; personalidade* ; jogada* ; triunfo*	5	2	9	--	original
alho* ; bacon* ; assado* ; cebola* ; manteiga* ; frango*	7	--	---	--	original (culinária)
paisagem* ; cidade* ; luz* ; gente*	3	2	---	--	original colorido
TOTAL	27	4	18	1	

Tabela 3: Objetos de *dourar/doirar* nos corpos CHAVE e CONDIV⁹

A análise dos verbos *branquear* e *embranquecer*, pertencentes ao grupo Branco, também revela fatos curiosos. Se voltarmos à tabela 2, vemos que, no CHAVE, tomando-se o lema *branquear*, é enorme a diferença entre a variante BR (4) e a variante PT (99). Quando exploramos essa forma verbal, percebemos que, das 99 ocorrências, 68 referem-se a um uso original da cor, em que o principal objeto de *branquear* é *dinheiro*, e os demais lemas são variados (tabela 4). O segundo uso mais comum de *branquear*, e que também comparece de maneira mais frequente na variante PT, diz respeito ao campo da cor humana. Branqueiam-se *dente*, *perna*, *pele* *mão*. Ainda na tabela 4, vemos que, na variante BR, não há uso original de *branquear*: o uso colorido é o mais comum, mas ainda assim são poucas as ocorrências no corpos (apenas 3). Com relação ao uso associado ao ser humano, a variante BR conta com apenas uma ocorrência (*branquear o pescoço*). Não encontramos nem uma ocorrência de *branquear* usado de maneira original. Com relação à expressão *branquear dinheiro*, muito frequente em PT, o seu equivalente em BR é *lavar dinheiro* (22 ocorrências no CHAVE), ou, mais especificamente, a expressão *lavagem de dinheiro* (96 ocorrências no CHAVE).

Lema dos objetos de <i>branquear</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
dinheiro**	---	--	14	1	original
imagem** ; PIDE** ; comportamento** ; situação** ; fascismo** ; derrota** ; tortura** etc.	---	---	68	--	original
roupa* ** ; rolha* ; celulose* ; marfim**	3	---	1	1	colorido
dente** ; pescoço* ; perna** , pele** , mão**	---	1	1	6	cor:humana
TOTAL	3	1	84	8	

Tabela 4: Objetos de *branquear* nos corpos CHAVE e CONDIV¹⁰

variante BR) [func="SUB]>.*"] [pos!="V.*"]* [pos="V.*" & lema="do[ui]rar" & sema="cor.*" & temcagr!="PCP" & variante="BR"] within s

⁹ Para realizar a busca por lema dos objetos de *dourar/doirar* no AC/DC, após a escolha do corpo desejado para a pesquisa, a expressão de busca é (o exemplo refere-se à busca apenas na variante PT) [sema="cor.*" & grupo="Dourado" & temcagr!="PCP" & pos="V.*" & variante="PT"] [pos!="V.*"]* @[func="<ACC"] within s

Já a exploração de *embranquecer* aponta para um quadro diferente. Voltando à tabela 2, não apenas a frequência de uso é menor do que em *branquear*, como também os dados são mais equilibrados entre as variantes. A análise qualitativa de *embranquecer* revela um uso associado à cor humana, independente da variante, como pode ser observado na tabela 5.

Lema dos objetos de <i>embranquecer</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
tudo e todos	1	---	---	---	cor:raça original
Pele* **; mão**; cabelo* **	2	---	1	2	cor:humana
TOTAL	3	---	1	2	

Tabela 5: Objetos de *embranquecer* nos corpos CHAVE e CONDIV

A exceção é para uma ocorrência de *embranquecer*, na variante BR, que pode ser associada a um uso original ou, simultaneamente, ao uso associado à utilização da cor como raça, como pode ser percebido na frase:

FSP951120-009: É que foram antes massacrados por uma mídia que busca embranquecer **tudo e todos**.

Por fim, se desconsideramos, na busca, a presença de objeto, e analisamos formas intransitivas de *embranquecer*, aparecem mais duas ocorrências, ambas na variante BR, nenhuma colorida, e ambas associadas ao campo semântico da raça:

FSP950319-128: Ao adotar como visão de si mesmo a ideologia de seus dominadores, o mestiço opta pelo recalque e pela traição de tudo que nele não for espelho da Europa: o filho de índia e europeu, identificando-se com o pai, tornou-se perseguidor do gentio materno; o mulato, buscando ascender socialmente, trata desesperadamente de **embranquecer**, reforçando e legitimando o preconceito com o negro.

FSP950623-092: Suassuna acredita que os espetáculos, se bem produzidos, seriam capazes de atrair ` 300 vezes mais gente do que o Michael Jackson» (que ele define como ` traidor da raça negra», por sua suposta vontade de ` **embranquecer**»).

Com relação ao grupo Preto, nossa análise considerou apenas o verbo *enegrecer*. Pela tabela 2, constatamos que *enegrecer* é muito mais frequente na variante PT (e com um uso muitíssimo mais alto no corpo Vercial (63), que não exploramos aqui), e é usado principalmente de maneira original. A tabela 6 oferece uma visão qualitativa de *enegrecer*, em ambas as variantes.

Lema dos objetos de <i>enegrecer</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
costa**, céu**, prata**	---	---	2	1	cor
Abril**; coração**; imagem**; verso**; cenário* **; desporto**	1	---	6	1	cor:original
dente**	---	---	1	---	cor:humana
TOTAL	1	---	9	2	

Tabela 6: Objetos de *enegrecer* nos corpos CHAVE e CONDIV

¹⁰ Expressão de busca utilizada, considerando a variante PT: [sema="cor.*" & lema="branquear" & temcagr!="PCP" & pos="V.*" & variante="PT"] [pos!="V.*"]* @[func="<ACC"] within s Distribuição de lema

Quando buscamos pelos responsáveis pelo enegrecimento, encontramos, nos usos coloridos, *petróleo (enegrece a costa)*, *fumo (enegrece o céu)* e *enxofre (enegrece a prata)* e, nos usos originais, *ódio (enegrece o coração)*, ou incertezas (*enegrecem o cenário*), sendo apenas o último exemplo relativo à variante BR. Se eliminamos a restrição referente à presença de objeto, temos mais uma ocorrência de *enegrecer*, na variante BR do CHAVE, vaga entre [sema="cor:raça"] ou [sema="cor:humana"]:

FSP951016-045: No começo do século, sociólogos acreditavam fortemente nas variáveis raciais e, otimistas, diziam que o Brasil não tinha um problema com os negros, pois iríamos branqueando gradativamente com os casamentos inter-raciais, enquanto os americanos se preocupavam, pois acreditavam que iriam **enegrecendo** por meio dos mesmos cruzamentos entre negros e brancos.

Os verbos do grupo Amarelo – *amarelar* e *amarelecer* – também apresentam comportamentos distintos conforme a variante.

Com relação a *amarelar*, em termos quantitativos, quase não há diferença entre as variantes PT e BR (tabela 2). No entanto, a análise qualitativa a partir dos corpos (tabela 7) revela que, considerando apenas o corpo CHAVE-BR encontramos (poucas) ocorrências transitivas¹¹ de *amarelar*, que é usado sobretudo relacionado a atributos humanos, mas encontramos também um uso original (*o remorso amarelado os olhos*) e um uso colorido.

Lema dos objetos de <i>amarelar</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
Pratos*	1	---	---	---	cor
Pessoas*; unhas*	2	---	---	---	cor:humana
olhos*	1				cor: original
TOTAL	4	---	---	---	

Tabela 7: Objetos de *amarelar* nos corpos CHAVE e CONDIV

No entanto, se consideramos os usos intransitivos ou transitivos indiretos de *amarelar*, o quadro se altera ligeiramente. Encontramos mais três ocorrências (duas no CHAVE BR, uma no CONDIV PT), todas em um uso claramente original, mas com diferenças de sentido conforme a variante. Na variante BR, *amarelar* equivale a *ter medo de*; *se acovardar* (sentido inexistente na variante PT). Na variante PT (cf. o terceiro exemplo), *amarelar* equivale a *distribuir cartões amarelos*.

FSP951030-086: Folha -- Você já «**amarelou**» para ondas grandes em alguma competição? (CHAVE-BR)
 par=fut48437: Mas não é menos que, depois que descobriram que a capital boliviana era nos Andes, os jogadores brasileiros literalmente **amarelaram**. (CONDIV- BR)
 par=fut18426: O jogo foi muito árduo e houve necessidade de «**amarelar**». (CONDIV- PT)

Com relação a *amarelecer*, a tabela 2 já indica o uso mais frequente na variante PT, com apenas 3 ocorrências na variante BR. A tabela 8 contém a descrição qualitativa do comportamento de *amarelecer*.

¹¹ [sema="cor.*" & lema="amarelar" & temcagr!="PCP" & pos="V.*" & variante="PT"] [pos!="V.*"]* @[func="<ACC"] within s

Lema dos objetos de <i>amarelecer</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
espetáculo*; mármore**	1	---	2	--	cor
pele*	1	---	--	--	cor:humana
TOTAL	2	---	-	--	

Tabela 8: Objetos de amarelecer nos corpos CHAVE e CONDIV

No entanto, como *amarelecer* é usado principalmente de maneira intransitiva, apresentamos os sujeitos de *amarelecer*, na variante PT: *página, tinta, relva, colectâneas, relvado* [sema="cor"]; *fruta, sorriso* [sema="cor:original"]. Notamos ainda uma ocorrência de *amarelecer de inveja*, referente a [sema="cor:original"].

Por fim, detalhamos os usos dos verbos do grupo Verde – *esverdear, reverdecer e verdejar*. A tabela 2 mostra que, para os três lemas, há muito mais ocorrências na variante PT. No entanto, se restringimos a análise aos corpos CHAVE e BR, as diferenças somem, com no máximo duas ocorrências por variante. A única diferença está no lema *esverdear*, na variante PT do CHAVE, com 8 ocorrências, que detalhamos a seguir. Todas as ocorrências de *esverdear* são originais, como vemos na tabela 9:

Lema dos objetos de <i>esverdear</i>	Distribuição por variante e corpo				Campo semântico
	BR*		PT**		
	CHAVE	CONDIV	CHAVE	CONDIV	
indústria**; pílula**; sistema**; projetos**; atividade econômica**	---	---	7	1	Cor:original

Tabela 9: Objetos de esverdear nos corpos CHAVE e CONDIV

Com relação aos demais verbos de cor – *rosar, acinzentar, arroxear; avermelhar; alaranjar* – a baixa frequência de ocorrência nos corpos CHAVE e CONDIV e o limite de páginas nos obrigam a postergar a descrição para outra ocasião. Lembramos, no entanto, que todo o material encontra-se publicamente disponível na página do projeto AC/DC para aqueles que se interessarem pelo campo das cores.

Considerações finais

Neste trabalho, exploramos dois corpos jornalísticos, com textos das variantes de Portugal e do Brasil, quanto a classes gramaticais que se referem às cores, com especial atenção aos verbos.

De maneira sumarizada, os resultados encontrados indicam que:

- A distribuição das categorias gramaticais é a mesma entre as variantes portuguesa e brasileira;
- A distribuição dos verbos de cor e dos adjetivos de cor também é a mesma entre as variantes portuguesa e brasileira;
- Como esperado, verbos se prestam pouco a representar o processo de “colorização” das coisas. Assim, não é surpresa que a grande maioria dos verbos de cores esteja no participípio. E, quando desconsideramos essa forma, independente de variante ou gênero, certos grupos de cores, como *Castanho* e *Crema* desaparecem. Por outro lado, grupos, como *Verde, Branco, Preto, Amarelo* e *Dourado*, são mais produtivos na formação de verbos, dando origem a dois ou mais lemas distintos.

Especificamente quanto às formas verbais que se referem às cores, acreditamos ter descoberto pontos que merecem análise mais detalhada. De uma perspectiva do uso, é interessante perceber o que fazem os verbos de cor. *Branquear* e

embranquecer, por exemplo, têm comportamentos diferentes não apenas quanto à frequência. *Branquear*, mais comum, é usado principalmente na variante de Portugal, e de maneira metafórica: muito mais que os *dentes*, branqueiam-se *dinheiro* e *imagem*. *Embranquecer* tem um uso mais literal – embranquecemos os *cabelos*, a *pele*. No entanto, o que mais nos chama a atenção é a frequência dos usos originais associados aos verbos de cor – excluída a forma participial –, o que parece sugerir uma perda gradativa da propriedade descritiva desses verbos. A alta ocorrência, em muitos casos, dos verbos de cor apenas no corpo VERCIAL, composto por textos literários antigos, reforça essa hipótese.

Alguns pontos permanecem para análise futura: nossa busca excluiu locuções verbais. Usamos mais “ficar vermelho” do que “avermelhar”? E o que nos diria a análise das cores no participípio?

Por fim, reforçamos a ideia de que para comparar termos de cor entre duas ou mais línguas é importante uma caracterização detalhada das cores em cada uma das línguas contrastadas. Em nosso estudo, buscamos conjugar as dimensões de uso e de forma, oferecendo um quadro das cores em português, considerando também distinções entre as variantes do Brasil e de Portugal.

Agradecimentos

O presente trabalho foi desenvolvido no âmbito da Linguatca, financiada pelo governo português, UMIC, FCCN e União Europeia (FEDER e FSE), por meio de POSC/339/1.3/C/NAC, e pela FCT.

REFERÊNCIAS BIBLIOGRÁFICAS

BACELAR DO NASCIMENTO, Maria Fernanda & Anabela Carvalho. "Preto e branco ou branco e preto? (Como se combinam os nomes de cores)". In *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística* (Lisboa, 2-4 de Outubro de 1995), 1996, Lisboa: APL/Colibri, pp. 367-380.

BERLIN, Brent & Paul Kay. *Basic Colour Terms: their Universality and Evolution*. Stanford: CSLI, 1991 [1.a edição: 1969]

BIDERMAN, Maria Tereza Camargo, Maria Fernanda Bacelar do Nascimento & Luisa Alice Santos Pereira. “Uso das cores no português brasileiro e no português europeu.” In: Aparecida Negri Isquierdo & Ieda Maria Alves (eds.), *As ciências do léxico: Lexicologia, lexicografia, terminologia*, vol. III, Editora UFMS, Associação editorial Humanitas, 2007, 105-124.

CORREIA, Margarita. "Towards a General Description of the Semantic Field of 'Colour' in European Portuguese". In C.P. Biggam & Christian J. Kay (eds.), *Progress in Colour Studies*, 1: Language and Culture. Amsterdam/Filadélfia: John Benjamins, 2006, pp. 111-25.

COSTA, Luís, Diana Santos & Paulo Alexandre Rocha. "Estudando o português tal como é usado: o serviço AC/DC". In *The 7th Brazilian Symposium in Information and Human Language Technology (STIL 2009)*, 2009, São Carlos, Brasil.

DEUTSCHER, Guy. *Through the language glass: Why the World Looks Different in Other Languages*. Metropolitan: Henry Holt, 2010.

FARIAS, Emília Maria Peixoto & Luís Antônio Marcuschi. "A metáfora das cores na linguagem e no pensamento". In Abuêndia Padilha Pinto (org.), *Tópicos em cognição e linguagem*. 1ed. Recife: Ed. Universitária da UFPE, 2006, v. 1, p. 19-55.

FRANKENBERG-GARCIA, Ana & Diana Santos. "COMPARA, um corpus paralelo de português e de inglês na Web". *Cadernos de Tradução* IX.1, 2002, pp. 61-79. Universidade de Santa Catarina. ISSN: 1676-7047.

INÁCIO, Susana, Diana Santos & Rosário Silva. "COMPARAndo cores em português e inglês". In Sónia Frota & Ana Lúcia Santos (orgs.), *Artigos selecionados do XXIII Encontro da Associação Portuguesa de Linguística*, APL, 2008, pp. 271-86.

JORGE, Guilhermina (coord.) et al. "As cores preto no branco: uma análise comparativa", *Polifonia* 6, Revista do Grupo Universitário de Investigação em Línguas Vivas da Universidade de Lisboa, Edições Colibri, Lisboa, 2003, pp. 119-133.

LUCY, John A. "Linguistic relativity". *Annual Review of Anthropology* 26. Palo Alto, CA: Annual Reviews Inc, 1997, pp 291-312.

MOTA, Cristina & Diana Santos. "Corte e costura no AC/DC: auxiliando a melhoria da anotação nos corpos". Setembro de 2009. <http://www.linguateca.pt/acesso/corte-e-costura.pdf>

PHILIP, Gillian Susan. "Collocation and connotation: A Corpus-Based Investigation Of Colour Words In English And Italian", PhD thesis, University of Birmingham, Março de 2003.

PHILIP, Gill. *Colouring Meaning: Collocation and connotation in figurative language*. John Benjamins Publishing Company, 2011.

ROCHA, Paulo & Diana Santos. "CLEF: Abrindo a porta à participação internacional em avaliação de RI do português". In Diana Santos (ed.), *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. Lisboa, Portugal : IST Press, 2007, pp. 143-158.

ROSCH, Eleanor. "Universals and cultural specifics in human categorization". In Brislin, Richard, Stephen Bochner & Walter Lonner (eds.), *Cross-cultural perspectives on learning*. Wiley, 1975, pp. 177-206.

SANTOS, Diana. "Corpora at Linguateca: vision and roads taken". In Tony Berber Sardinha & Telma São Bento Ferreira (eds.), *Working with Portuguese corpora*. No prelo.

SANTOS, Diana & Cristina Mota. "Experiments in human-computer cooperation for the semantic annotation of Portuguese corpora". In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2010)* (Valletta, Malta, 17-23 de Maio de 2010), European Language Resources Association, pp. 1437-1444.

SANTOS, Diana, Rosário Silva & Susana Inácio. "What's in a colour? Studying and contrasting colours with COMPARA". *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008)* (Marraqueche, 26 de Maio -1 de Junho de 2008), ELDA.

SANTOS, Diana, Rosário Silva & Cláudia Freitas. "Pluralidades na cor: contrastando a língua do Brasil e de Portugal". In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.), *Línguas Pluricêntricas: Variação Linguística e Dimensões Sociocognitivas. Pluricentric Languages: Linguistic Variation and Sociocognitive Dimensions*. Braga: Aletheia, Publicações da Faculdade de Filosofia da Universidade Católica Portuguesa, 2011, pp. 555-572.

SANTOS, Diana, Stella E.O. Tagnin & Elisa Duarte Teixeira. "Colours, clothing and food in CorTrad: why corpus-based translation studies are revealing". Apresentação no ICAME 2011 (Oslo, 1-5 de junho de 2011).

SAUNDERS, Barbara. "Revisiting basic colour terms", *Science as Culture, The Human Nature Review*, 2005, <http://human-nature.com/science-as-culture/saunders.html>.

SILVA, Augusto Soares da. "O corpus CONDIV e o estudo da convergência e divergência entre variedades do português". In Luís Costa, Diana Santos & Nuno Cardoso (eds.), *Perspectivas sobre a Linguateca / Actas do encontro Linguateca : 10 anos*. Linguateca, 2008, pp. 25-28.

SILVA, Rosário & Diana Santos. "Arco-íris: notas sobre a anotação do campo semântico da cor em português". Primeira edição: 25 de junho de 2009. Versão atual: 24 de janeiro de 2012. <http://www.linguateca.pt/aceso/ArcoIris.pdf>

SILVA, Rosário, Susana Inácio & Diana Santos. "Documentação da anotação relativa à cor no COMPARA". Última versão: 31 de Dezembro de 2008. <http://www.linguateca.pt/COMPARA/DocAnotacaoCorCOMPARA.pdf>

SLETSJØE, Leif. "Sobre el topico de los ojos verdes", *Strenae: Estudios de filologia e historia dedicados al profesor Manuel Garcia Blanco*, Salamanca, 1962, pp. 445-459.

SWEARINGEN, Andrew. "Seeing red in Roxo: The Evolution of Portuguese colour terms", MA thesis, Univ. of Copenhagen, 2000.

TEIXEIRA, Elisa D., Diana Santos & Stella E. O. Tagnin. "CorTrad: um novo corpus paralelo multiversão para o par de línguas português-inglês". Em Tania M.G. Shepherd, Tony Berber Sardinha & Marcia Veirano Pinto (eds.), *Caminhos na Linguística de Corpus*, Mercado de Letras, 2012, pp. 151-176.

WIERZBICKA, Anna. The meaning of color terms: semantics, culture, and cognition. *Cognitive Linguistics*, 1-1, 1990, 99-150.

ZAVAGLIA, Cláudia. "Dicionário e cores". *Alfa* 50, 2, São Paulo, 2006, pp. 25-41.

ZAVAGLIA, Claudia. "A prática lexicográfica multilingue: Questões concernentes ao campo das cores". In Aparecida Negri Isquierdo & Ieda Maria Alves (eds.), *As*

ciências do léxico: Lexicologia, lexicografia, terminologia, vol. III, Editora UFMS, Associação editorial Humanitas, 2007, pp. 209-222.