

COMPARAndo cores em português e inglês

Susana Inácio
Diana Santos
Rosário Silva



Estrutura

- O COMPARA
- A cor
- Estudos em linguística com corpora
- Preparativos: “colorir” o COMPARA
- Caminhos a explorar
- Primeiros resultados

COMPARAndo cores em português e inglês 

COMPARA

www.linguateca.pt/COMPARA

- o maior corpus paralelo editado do mundo
 - 3 milhões de palavras, 100.000 “frases alinhadas”
- textos literários (originais e traduções) em **português e inglês**, 35 autores, 45 tradutores
- projecto no âmbito da **Linguateca**, iniciado 2000
- liderado por
 - Ana Frankenberg-Garcia e Diana Santos
- anotação morfossintáctica (parte portuguesa - PALAVRAS, parte inglesa - CLAWS)

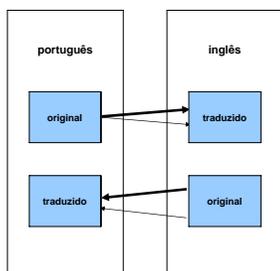
COMPARAndo cores em português e inglês 

Objectivos gerais do COMPARA

- Permitir estudos linguísticos das duas línguas e da tradução entre elas
- Ajudar no ensino e na aprendizagem da tradução e das línguas
- Ser uma ferramenta de consulta para tradutores e falantes das duas línguas
- Produzir material útil para investigação em processamento de linguagem natural e tradução

COMPARAndo cores em português e inglês 

A estrutura do COMPARA



(inspirada pelo ENPC-
English-Norwegian Parallel Corpus)

COMPARAndo cores em português e inglês 

Conteúdo do COMPARA 9.0.0

- 74 pares original-tradução
 - publicação do original: 1837 a 2000
 - publicação da tradução: 1862 a 2002
- alinhamento automático das frases revisto
 - reordenação, adição marcadas,...
- marcação/anotação
 - notas de tradução, ênfase, títulos,...
 - anotação morfossintáctica revista (em progresso)
- documentação sobre o corpus e seu uso

COMPARAndo cores em português e inglês 

A cor

- Forma de representação do universo, dada por diferentes culturas
- Análise
 - tradução da cor/cultura
 - aumento significativo da criação das palavras designando cor, para representar o universo

Dois tipos de estudos em linguística com corpora

- exploratório: observar os dados, definir categorias, detectar correlações
- experimental: dada uma hipótese, definir experiências que a possam corroborar

Trabalho preparatório

- lista de cores no COMPARA
 - palavras em português – 386 (209 lemas)
 - palavras em inglês – 335
- codificação no corpus [sem= "cor"]
- revisão manual
- primeiros resultados quantitativos: 4179 (PT) e 4060 (EN) ocorrências marcadas como COR

Exemplos de cores em português

- *rosa-shocking, amarelo-pálido, doirado, cinza-escuro, anilado, pérola*, grená, cinzento-claro, esverdinhado, verde-claro, bege, loura, castanho-escuro, morena, bronzeado, oiro*, lilás*, róseo, encarnado*, creme*, fulvo, azul-escuro, azulado, acastanhado, branco*

* desambiguado

Opções na marcação de "cor"

- palavras simples e compostas por justaposição, de qualquer categoria gramatical
- todas as palavras tendo um sentido colorido
 - denotando raça
 - outros domínios restritos: cabelo, pele, vinho,...
- não foram marcadas ausência de cor, ou cor não especificada (*descolorar, colorir, cor, corar*)

Distribuição da cor no COMPARA

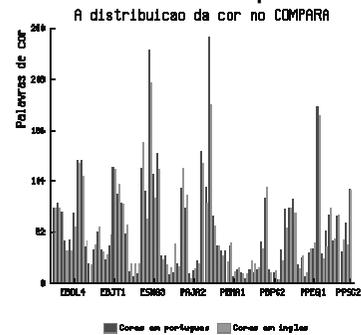
- por língua
 - português
 - inglês
- por variante
 - português (Portugal, Moçambique, Brasil e Angola)
 - inglês (Reino Unido, EUA e África do Sul)
- por texto
- por texto traduzido ou original
- por autor

Distribuição da cor por língua e variante

Tipo de texto	Textos	Ocorrências
Total (português)	74 (75)	4179
Original em português	40	1979
Traduzido para português	34	2200
Por variante do português		
Brasil	27 (28)	1430
Portugal	43	2683
África (Moçambique e Angola)	4	66
Por variante do inglês		
Reino Unido	39	2490
Estados Unidos	32	1117
África do Sul	4	572

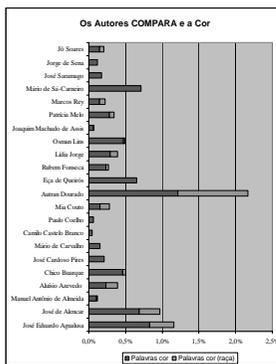
COMPARando cores em português e inglês

Palavras de cor por texto



COMPARando cores em português e inglês

Distribuição da cor por autor



COMPARando cores em português e inglês

Questões lexicais (PT)

- palavras simples (92 lemas, 3802 palavras)
- palavras em expressão *cor de X* (53 X, 99 casos)
- palavras compostas por justaposição (87 lemas, 200 palavras)

COMPARando cores em português e inglês

Exemplos mais frequentes

Palavras simples*		Palavras em expressão cor de X		Palavras compostas por justaposição*	
Cor (lema)	f	Cor (lema)	f	Cor (lema)	f
branco	786	cor de laranja	12	azul-marinho	12
negro	531	cor de areia	7	castanho-escuro	7
preto	383	cor de azeitona	4	azul-claro	5
vermelho	278	cor de cobre	4	azul-pálido	5
azul	256	cor de prata	4	verde-claro	5
verde	212	cor de âmbar	3	vermelho-vivo	5
amarelo	151	cor de ameixa	3	cinzento-claro	4
dourado	114	cor de bronze	3	azul-celeste	3
castanho	95	cor de cinza	3	cinzento-escuro	3
cinzento	90	cor de creme	3	cinzento-esverdeado	3

* marcadas com "cor"

COMPARando cores em português e inglês

Palavras compostas por justaposição (PT)

- 87 lemas
 - 9 dicionarizados (Houaiss, Academia, Aurélio)
 - 78 **não** dicionarizados
- Exemplos do primeiro caso: *azul-celeste, verde-azeitona, verde-garrafa,...*
- Exemplos do segundo caso: *amarelo-açafrão, azul-metálico, branco-manhã, louro-palha, verde-ramagem, verde-amarelo, vermelho-lacre, vermelho-framboesa,...*

COMPARando cores em português e inglês

Questões sintáticas

- categoria gramatical
- o género da cor
- construções gramaticais

Categoria gramatical

- Adjectivo: 3088 (*carro verde, fato azul*)
- Nome: 978
- Verbo: 74
- **ADJn: 27**
 - -- No entanto gosto do **azul** -- prosseguiu, a cliente -- e o **verde** é muito bonito.
- **N_ADJ: 9**
 - Mas Isidoro, mesmo **preto**, não era capaz de uma coisa dessas.
- ADV: 1 (*purpureamente*)

O género da cor

- M: 2515
- F: 1502
- Cores mais associadas ao masculino:
 - *ocre, róseo, violeta, marrom, pardo, grisalho, castanho, carmesim, ruivo, rosa,...*
- Cores mais associadas ao feminino:
 - *lilás, amarelado, roxo, acastanhado, encarnado, dourado,...*

Construções gramaticais

- nome + cor (2223):
 - *cabelo castanho, olhos azuis, seda preta*
- cor + conjunção e + cor (134):
 - *castanho e vermelho, rosa-shocking e branco*
- cor + pronome relativo que (5):
 - *cabelo **louro-claro** que já vai ficando branco*
 - *de um **azul** que só há nas porcelanas*

Questões semânticas

- O que é mais colorido
- Cores não padrão/criatividade
- Raça
- Outros domínios específicos
 - cabelo
 - céu
 - vinho

O que é mais colorido

- *cabelo* 117
- *olho* 78
- *seda* 35
- *papel* 32
- *vestido* 29
- *luz* 29
- *homem* 26

Cores não padrão/Criatividade (Mistura de cores)

- *amarelo-acinzentado, azul-esverdeado, castanho-azulado, castanho-dourado, cinzento-esverdeado, rosa-acinzentado, verde-amarelado, verde-cinza, vermelho-grená, vermelho-arroxeadado, cinza prateado, cinzento alilado, castanho dourado*

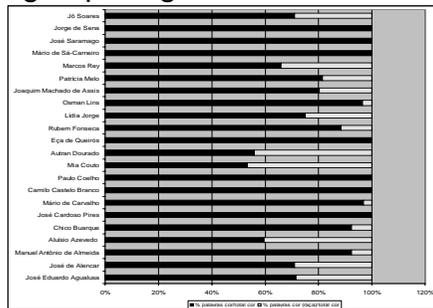
COMPARando cores em português e inglês

Cores de domínios específicos

- Cor de pele :
 - *moreno(a), negro(a), branco(a),...*
- Cor de cabelo:
 - *louro(s), loira(s), moreno(a), grisalha, ruivo,...*
- Céu:
 - *azul, anil, negro, vermelho,...*
- Vinho:
 - *branco, tinto, verde*

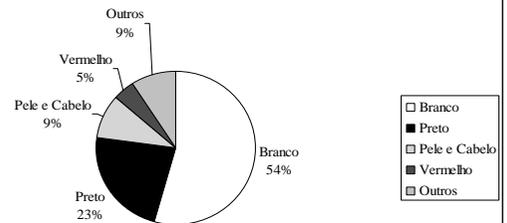
COMPARando cores em português e inglês

Cor indicando raça, por autor, em língua portuguesa



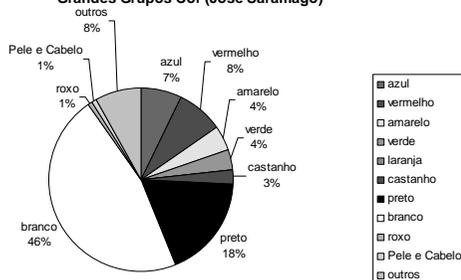
COMPARando cores em português e inglês

Cores Preferidas dos Autores de Língua Portuguesa



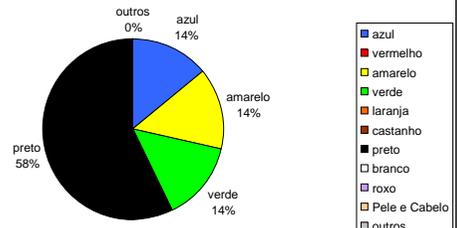
COMPARando cores em português e inglês

Grandes Grupos Cor (José Saramago)



COMPARando cores em português e inglês

Grandes Grupos Cor (Camilo Castelo Branco)



COMPARando cores em português e inglês

Questões discursivas

- em que tipos de frases aparece a cor
- a cor contribui para a avaliação ou para a descrição
- a cor em discurso directo
- que cores são positivas ou negativas

COMPARAndo cores em português e inglês



Questões de tradução

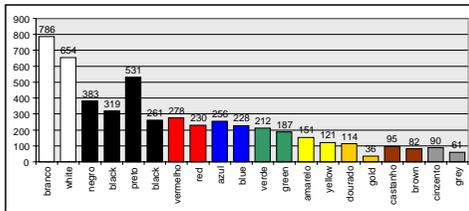
- há cores que se perdem?
- há cores mais difíceis de traduzir?
- existem estratégias específicas para a tradução da cor?
- há cores mais dependentes das culturas do que outras?

COMPARAndo cores em português e inglês



A tradução da cor (1/2)

- ✓ Ocorrências de cores em português com correspondente de cor em inglês

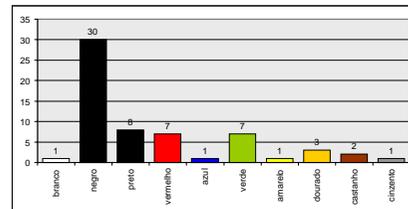


COMPARAndo cores em português e inglês



A tradução da cor (2/2)

- ✓ Ocorrências de cores em português sem correspondente de cor em inglês



COMPARAndo cores em português e inglês



Cores com mais correspondentes de cor em inglês

- **branco (48)**
- white, blank, pale, gray, blonde, silver,...
- **negro (48)**
- black, dark, gloomy, negro, charcoal,...
- **preto (27)**
- black, dark, negro, nigger, black-painted,...
- **vermelho (22)**
- red, scarlet, purple, vermilion-dyed, crimson,...
- **azul (22)**
- blue, navy, ice-blue, bluish, violet,...

COMPARAndo cores em português e inglês



Em conclusão

- Um manual de dados públicos para serem explorados
- Convite à participação da comunidade linguística nesta exploração da cor
- www.linguateca.pt/COMPARA/

COMPARAndo cores em português e inglês

