

# Compilação e uso de corpora paralelos

Ana Frankenberg-Garcia



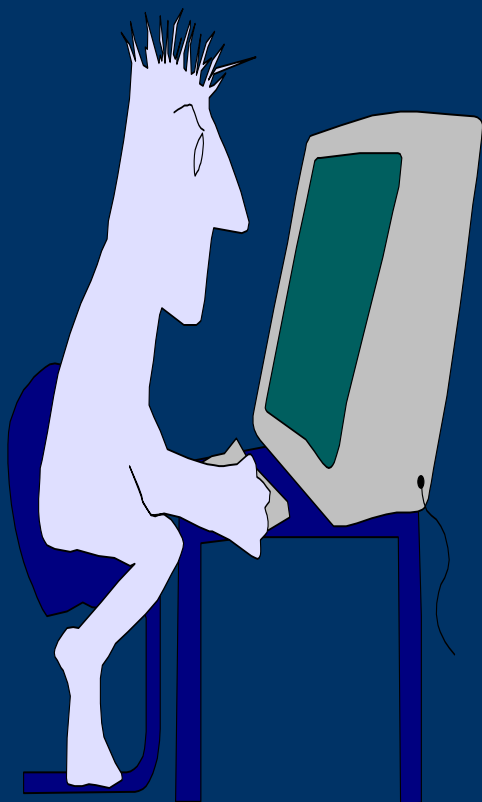
# O que é um corpus?

software específico

grande

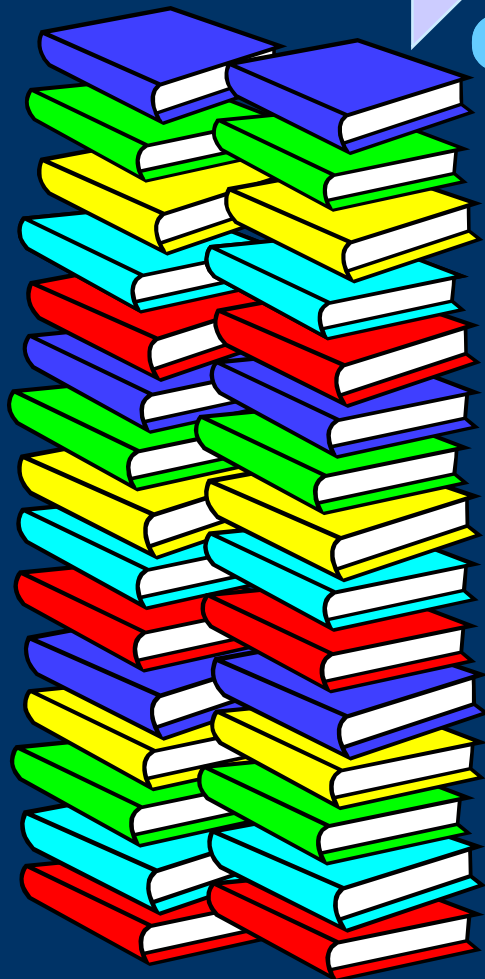
critérios definidos

em formato digital



unidirecional

# O que é um corpus paralelo?

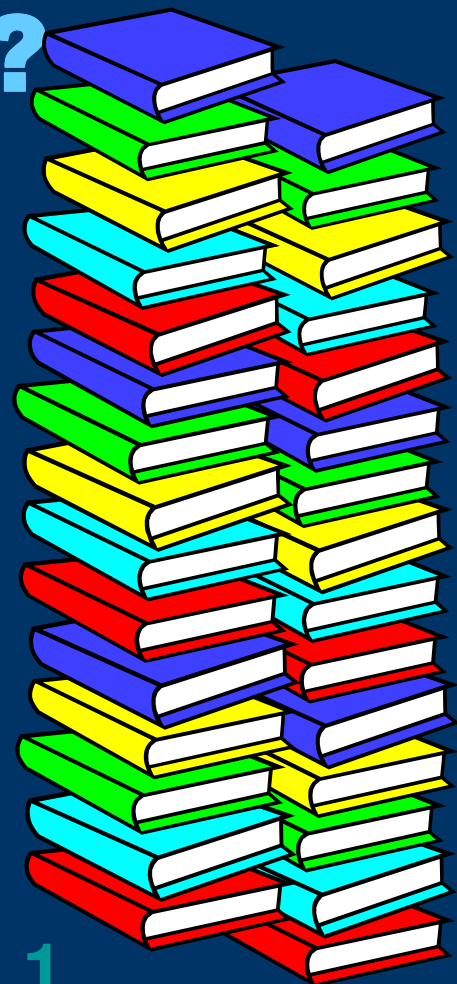


Traduções L2

Originais L1



bidirecional



Traduções L1

Originais L2

# Vantagens de corpora paralelos

- Permitem fazer tudo que um corpus monolíngüe faz
- Permitem ainda fazer análises impossíveis em corpora monolíngües

# Limitações de corpora paralelos

- Muito menos textos (e tipos de textos) traduzidos no mundo
  - Corpora paralelos em geral muito menores
- Originais e traduções tem de ser alinhados
  - Corpora paralelos dão muito mais trabalho

# Utilizações de corpora paralelos

## Diferenças entre corpora paralelos

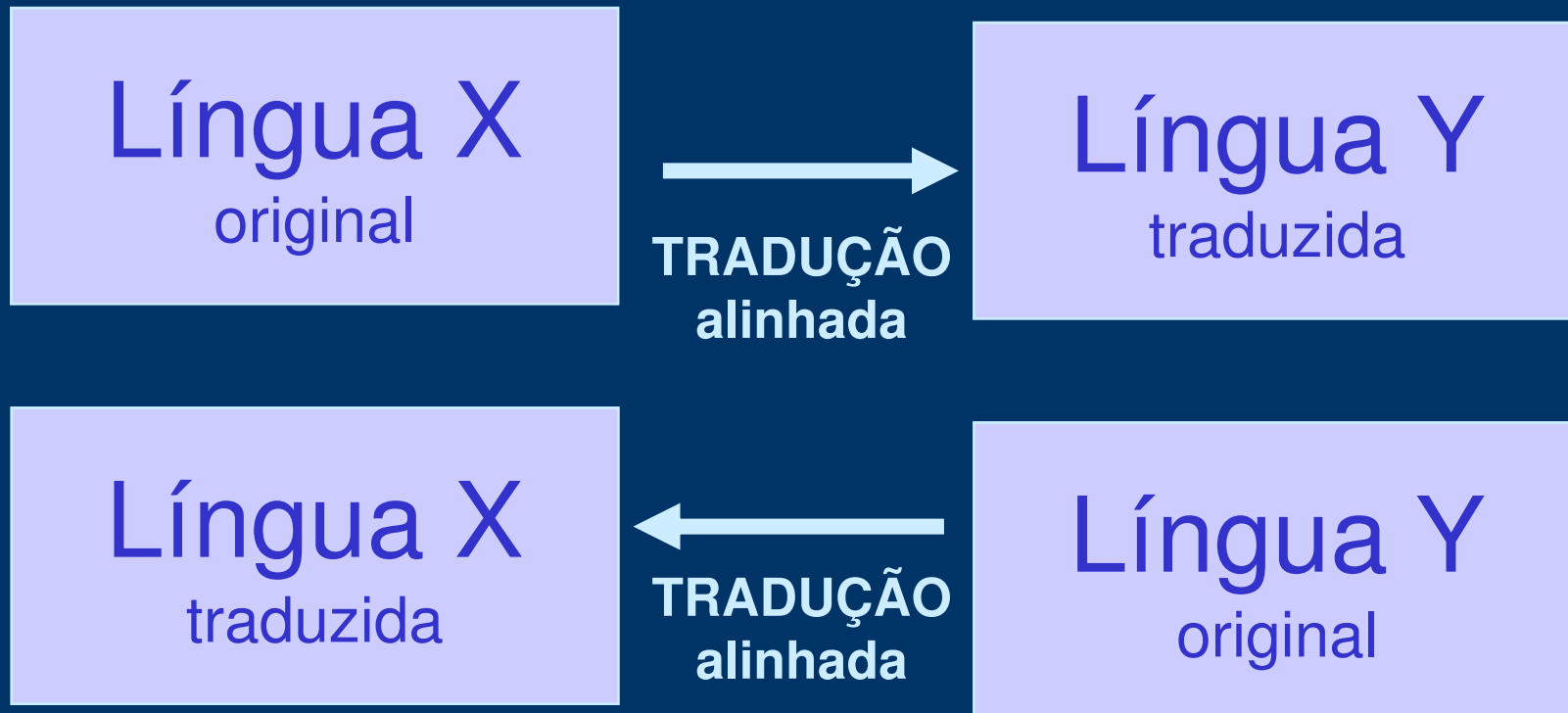


# 1. corpora paralelos unidirecionais



- dicionários, gramáticas bilíngües
- tradução automática

## 2. corpora paralelos bidirecionais

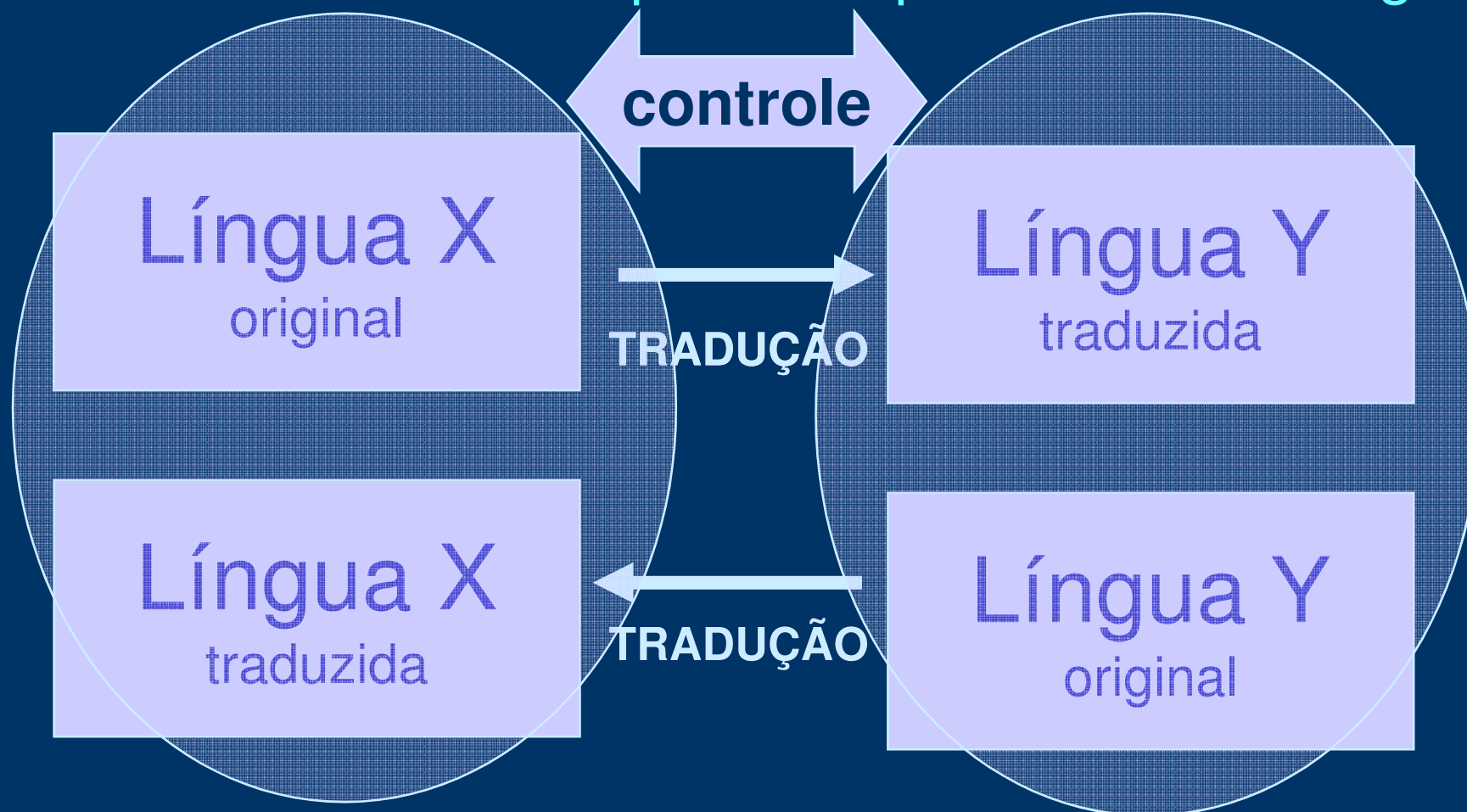


- dicionários, gramáticas bilíngües (dois sentidos)
- tradução automática (dois sentidos)



### 3. corpora paralelos bidirecionais

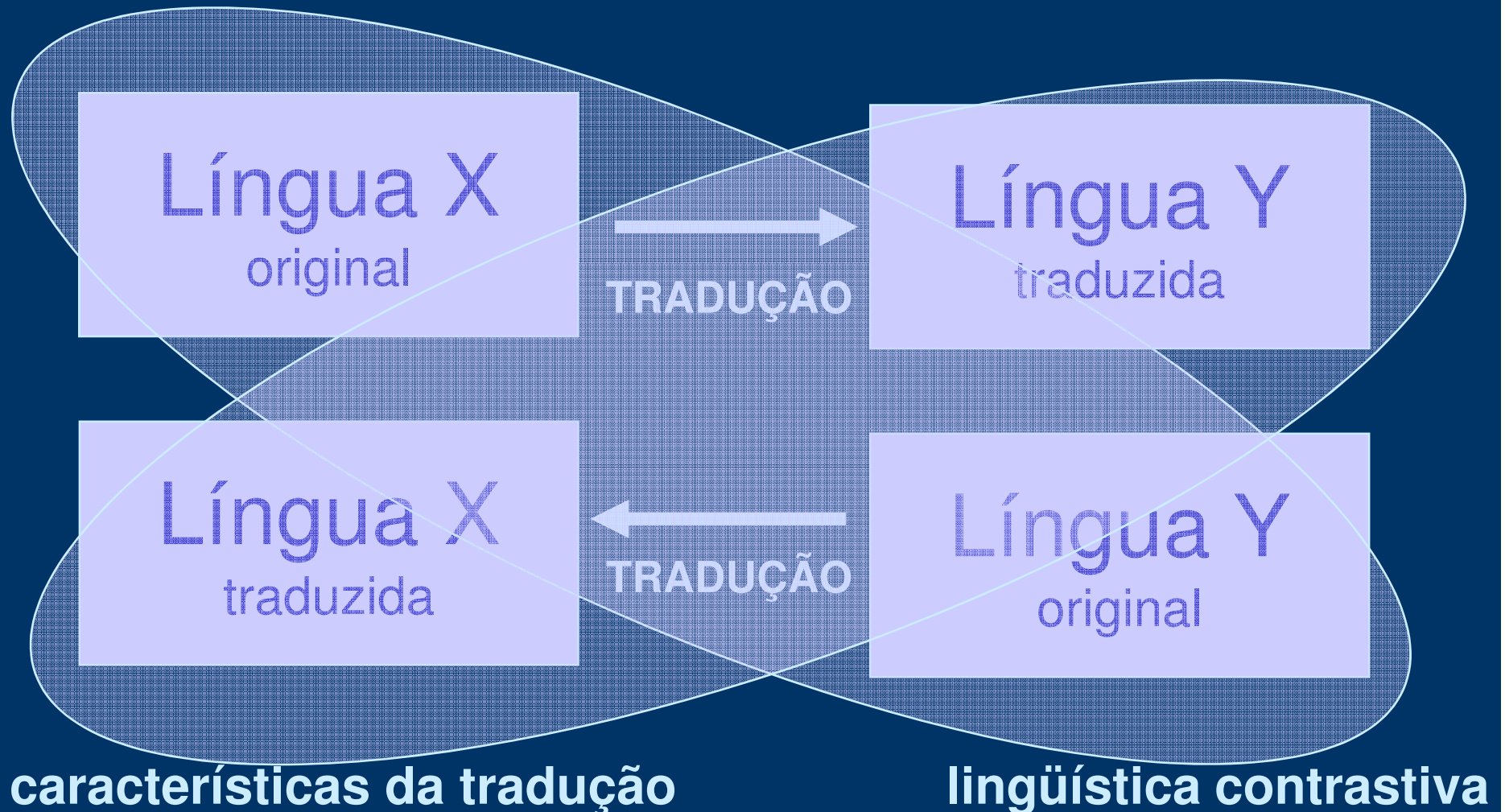
Também servem de corpora comparáveis monolíngües



- Diferenças entre língua original e língua traduzida
- Tradutês
- Estudos teóricos sobre universais da tradução

## 4. corpora paralelos bidirecionais

Também servem de corpora comparáveis bilíngües



# Compilação seleção de textos

- Gênero (científico, literário, técnico, etc.)
- Modo (textos orais? escritos?)
- Estilo (norma culta? popular?)
- Época (textos contemporâneos? antigos?)

**Há traduções?**

- Línguas (quais? só duas ou mais?)
- Traduções (profissionais? falantes nativos? tradutores diferentes? )
- Unidirecional ou bidirecional?

# Compilação seleção de textos

PT-EN

**unidirecional**



**bidirecional**

oral, popular  
científico

turismo

literário

político (PE)

# Compilação direitos de autor

Uso pessoal

Uso partilhado

Menos trabalho Autorizações

Aproveitamento

Estudos replicáveis

# **Compilação** direitos de autor

- Autorizações duplas, duplo trabalho
- Muitos editores, autores e tradutores não sabem o que é um corpus
- Proteger textos
- Publicitar textos

# Compilação alinhamento

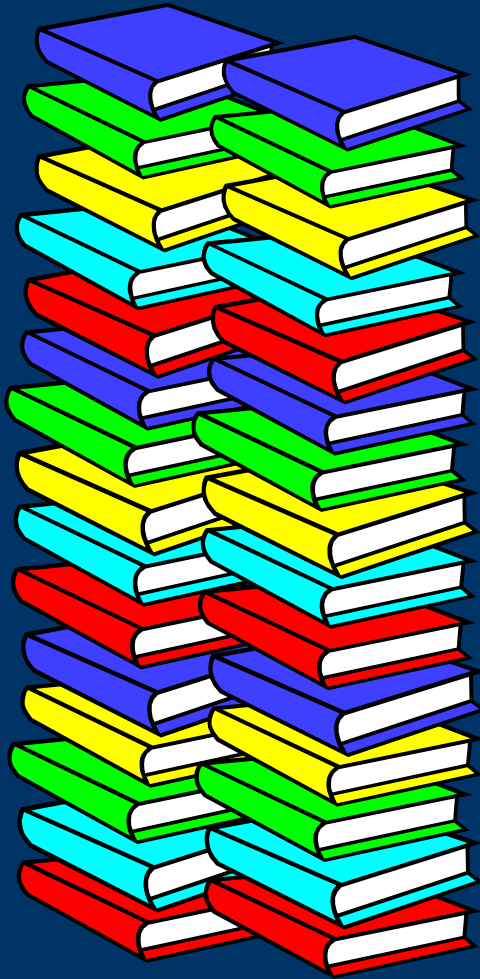
Que partes das traduções correspondem a que partes dos originais?

Texto?

Parágrafo?

Frase?

Palavra?



# Compilação etiquetagem e anotação

Etiquetas de alinhamento

Outras?

O mais leve possível

```
<id=EBDL3T2:294> -- Posso pedir um  
conselho de etiqueta?
```

Anotação gramatical detalhada

```
<id=EBDL3T2:294> --/PU Posso/V_fmc  
pedir/V um/DETarti conselho/N de/PRP  
etiqueta/N ?/PU
```

O que queremos automatizar? Quanto tempo temos?



# Alguns corpora paralelos acessíveis em rede

- **COMPARA**
  - Português-Inglês
- **Multisemcor**
  - Italiano-Inglês, Romeno-Inglês
- **OPUS, EUROPARL**
  - 11 línguas da UE
- **HUNGLISH CORPUS**
  - Inglês-Húngaro
- **CORPUS PARALELO CLUVI**
  - Vários mini-corpora com diferentes combinações de línguas, incluindo Galego, Espanhol, Catalão, Basco, Português, Inglês

# **As opções do corpus COMPARA**

<http://www.linguateca.pt/COMPARA/>

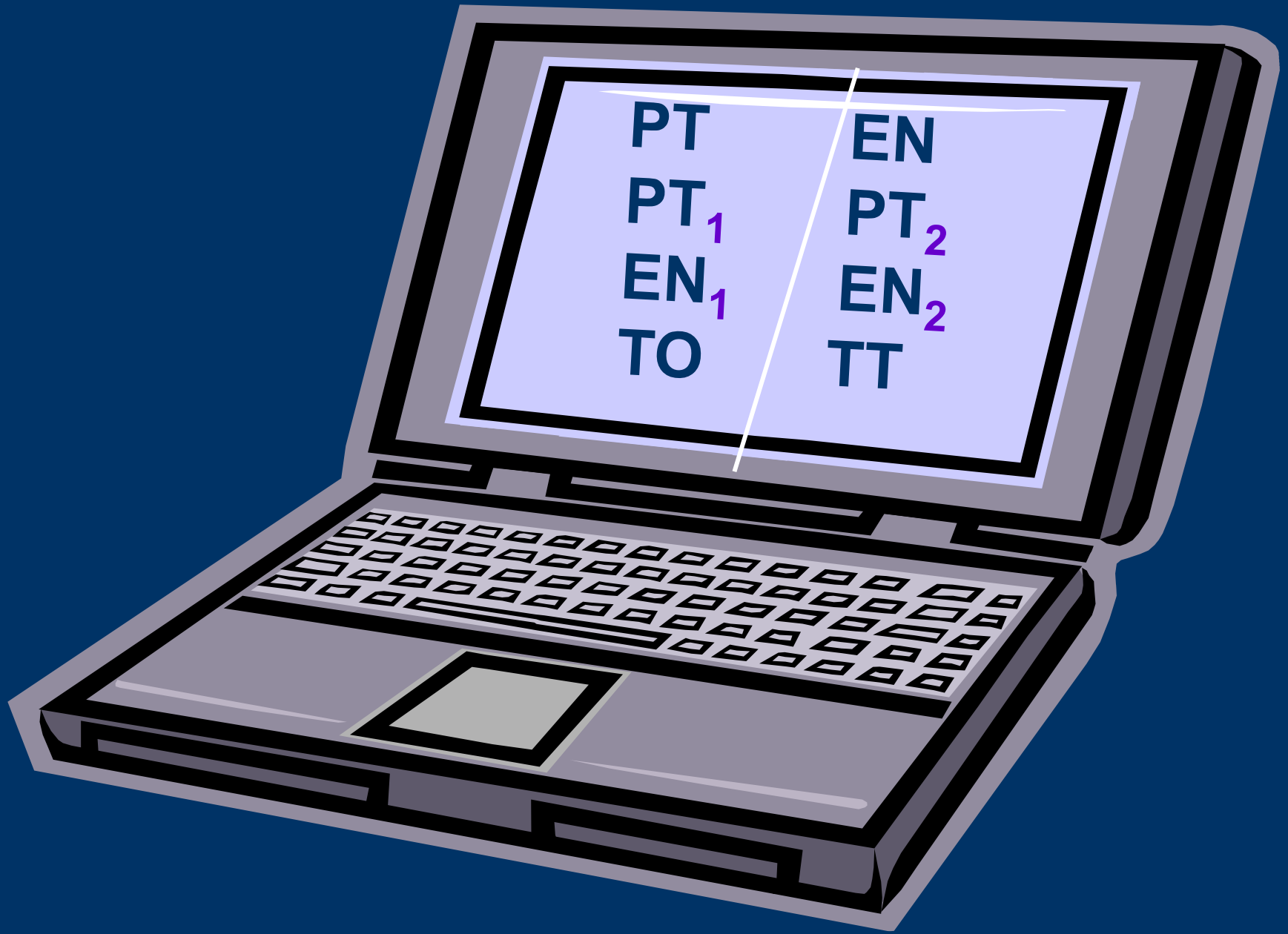
# COMPARA estrutura



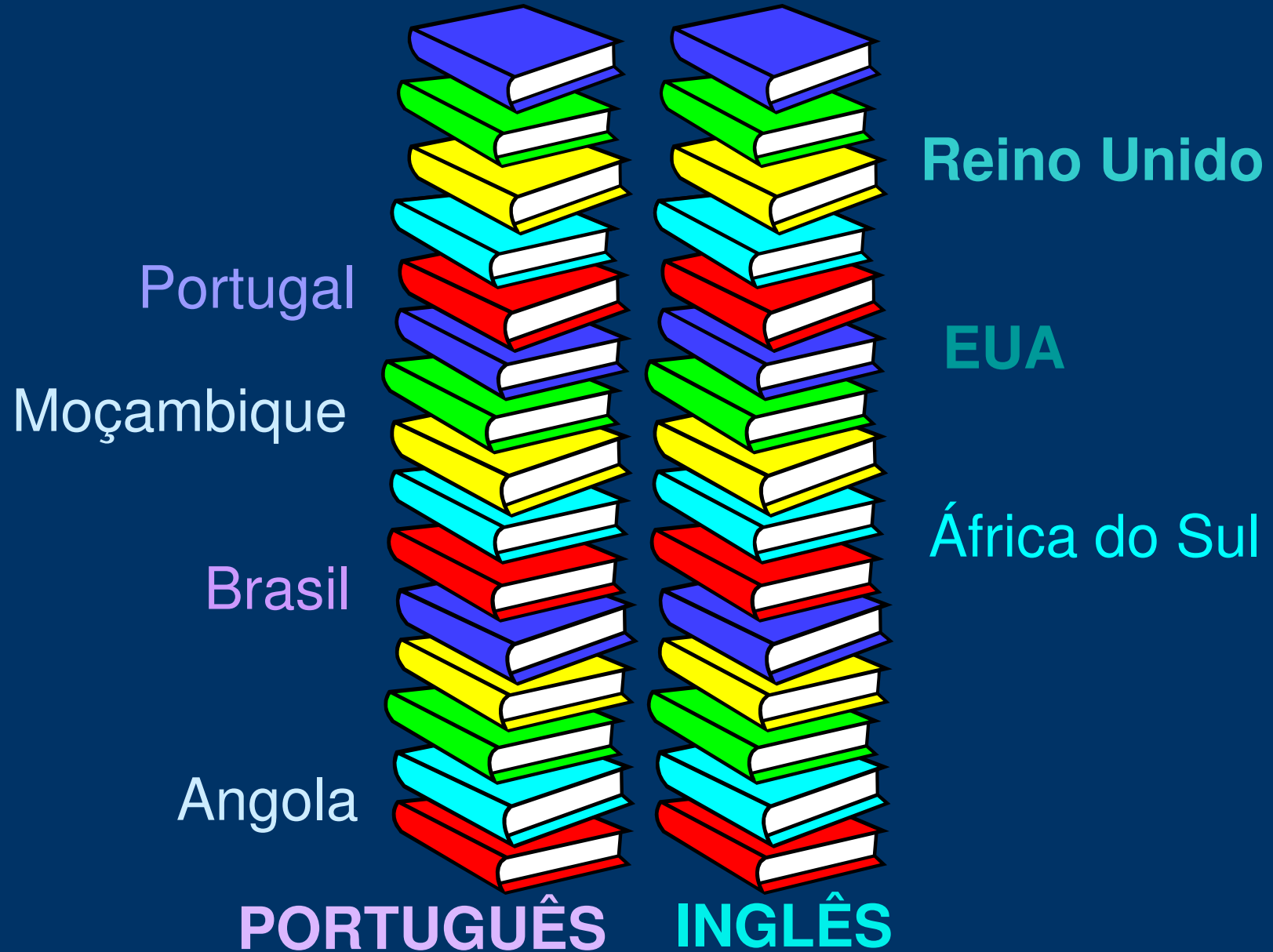
Traduções EN  
Originals PT

Traduções PT

Originals EN

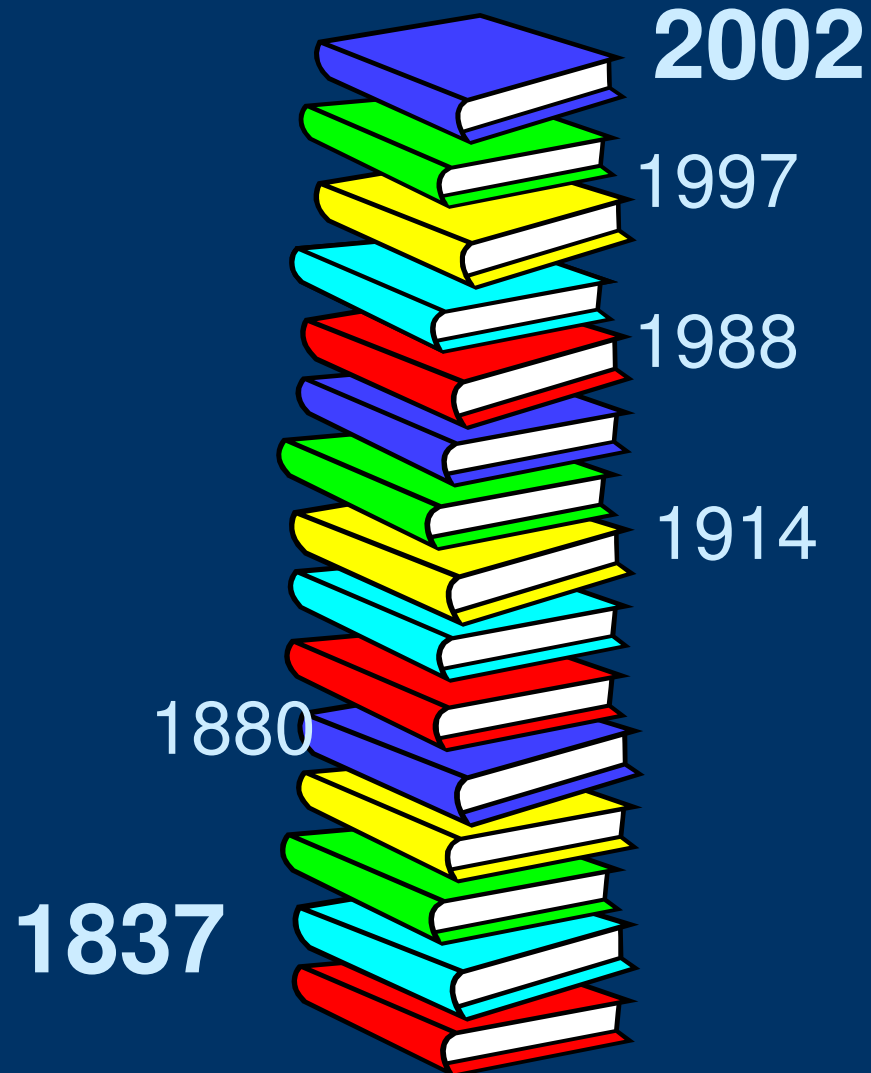


# COMPARA 8.2 variantes



# COMPARA 8.2

Datas de publicação



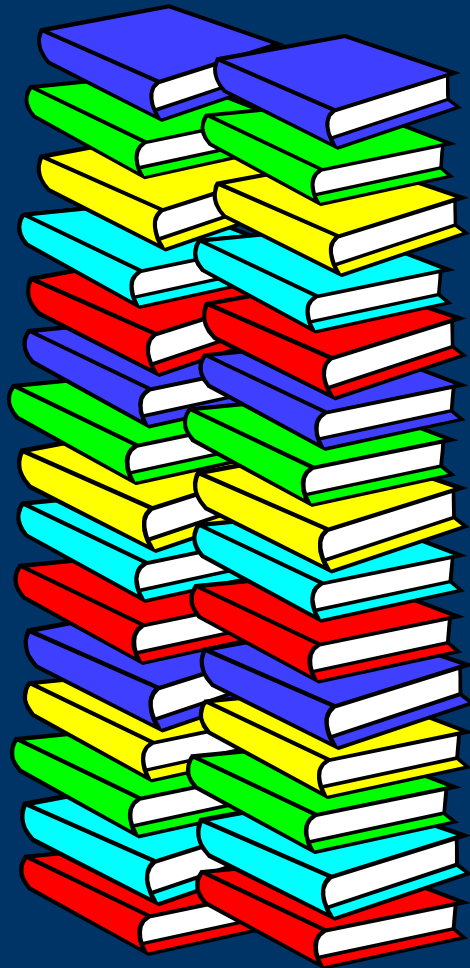
# COMPARA 8.2 gênero

Literatura  
publicada

extensível



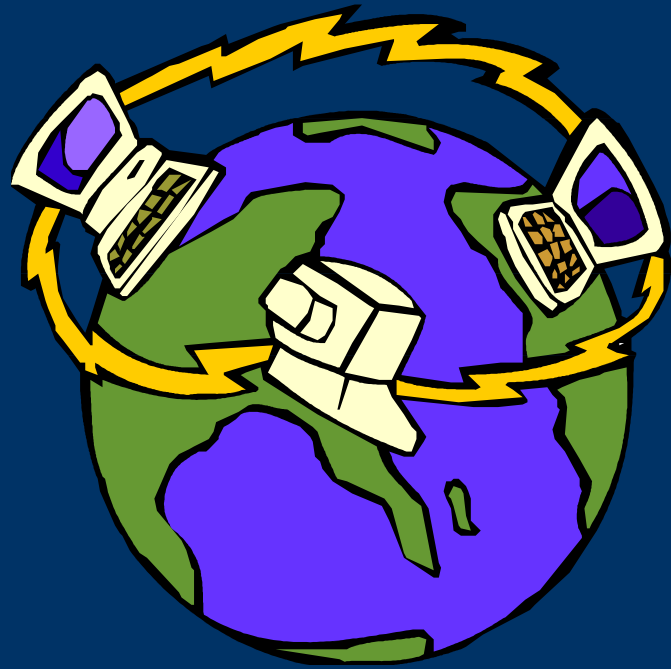
# Outros critérios de seleção de textos



- originais e traduções publicados
- inglês traduzido diretamente do português e português traduzido diretamente do inglês
- traduções feitas por seres humanos!
- Possível alinhar um original com mais de uma tradução



# COMPARA disponibilidade



Uso partilhado

acesso gratuito, em rede

para investigação e  
fins educacionais

# COMPARA Interface

- Pesquisa simples
- Pesquisa avançada
- Serviço em português
- Serviço em inglês

# COMPARA etiquetas semânticas

## EBJB1.pt

ele revelou-me o seu interesse por Gosse  
<tnote> Edmund William Gosse (1849–  
1928), crítico inglês </tnote> e pela  
sociedade literária inglesa dos finais  
do século passado.

## EBDL2T1.en

When we sat on the sofa together to  
watch <title>News at Ten</title>

# COMPARA etiquetas semânticas

**EBDL1T1.pt**

passou-me uma receita de **<named>**  
Valium **</named>**

**EBJB1.en**

the white bear, **<foreign>** thalassarctos  
maritimus **</foreign>**, is the aristocrat  
of bears...

**EBDL1T1.pt**

acaba por se esquecer de ter medo, até  
que acaba por verificar que não há  
**<emph>** de que **</emph>** ter medo.

## Atributos suplementares

### [Ajuda](#)

- Notas de tradução
- Títulos
- Palavras ou expressões estrangeiras
- Palavras ou expressões com ênfase
- Entidades mencionadas
- Adição de frases na tradução
- Supressão de frases na tradução
- Reordenamento de frases na tradução
- Junção de frases na tradução
- Separação de frases na tradução
- Todas as mudanças frásicas acima

## 3. Pretende utilizar todos os textos do corpus?

**Sim** - vá directamente para o [passo seguinte](#)

**Não** - utilize os parâmetros abaixo para seleccionar as suas preferências

Procurar

Limpar formulário

### 3.1. Restringir variantes:

#### [Ajuda](#)

#### Português

- Angola
- Brasil
- Moçambique
- Portugal

#### Inglês

- África do Sul
- Reino Unido
- Estados Unidos

### 3.2 Restringir datas de publicação:

#### [Ajuda](#)

Originais com 1ª edição

Traduções com 1ª edição

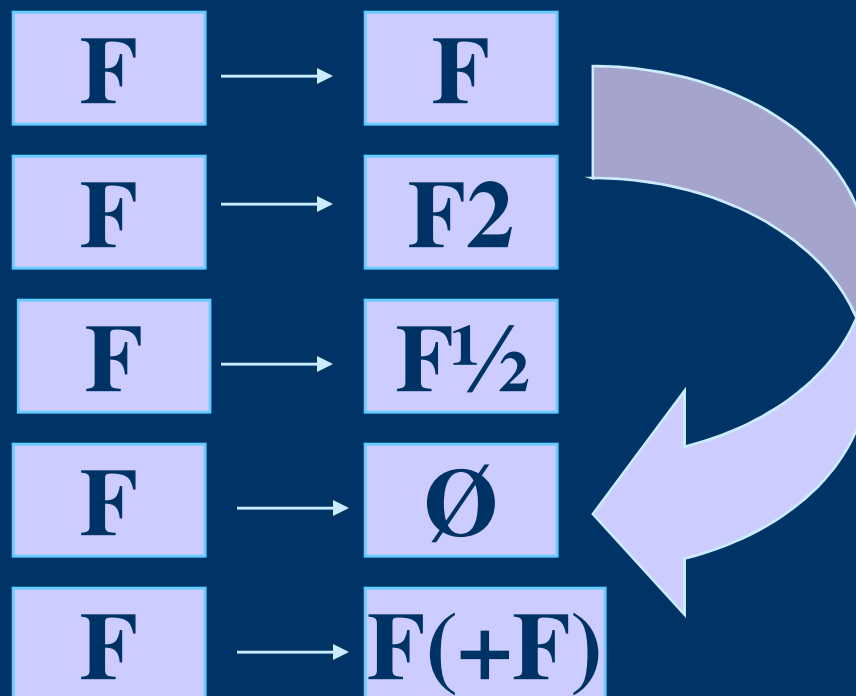
### 3.3 Conduzir pesquisas apenas de:

#### [Ajuda](#)

# COMPARA Alinhamento

1 unidade de alinhamento = 1 frase do texto original

Original Tradução



Etiquetas de alinhamento

## Atributos suplementares

### [Ajuda](#)

- Notas de tradução
- Títulos
- Palavras ou expressões estrangeiras
- Palavras ou expressões com ênfase
- Entidades mencionadas

- Adição de frases na tradução
- Supressão de frases na tradução
- Reordenamento de frases na tradução
- Junção de frases na tradução
- Separação de frases na tradução
- Todas as mudanças frásicas acima

## 3. Pretende utilizar todos os textos do corpus?

**Sim** - vá directamente para o [passo seguinte](#)

**Não** - utilize os parâmetros abaixo para seleccionar as suas preferências

Procurar

Limpar formulário

### 3.1. Restringir variantes:

#### [Ajuda](#)

#### Português

- Angola
- Brasil
- Moçambique
- Portugal

#### Inglês

- África do Sul
- Reino Unido
- Estados Unidos

### 3.2 Restringir datas de publicação:

#### [Ajuda](#)

Originais com 1ª edição

Traduções com 1ª edição

### 3.3 Conduzir pesquisas apenas de:

#### [Ajuda](#)

# **COMPARA** Anotação Gramatical

1. Português – PALAVRAS

2. Inglês - CLAWS



## Concordância

Procura: "grande".

EBDL1T1 (55):	Com/PRP os/DETDartd rendimentos/N que/ADVinterr_quant actualmente/ADV tenho/V_fmc ./PU podia/V_fmc dar/V_10/NUMcard 0/NUMcard libras/N por/PRP ano/N sem/PRP <b>grande/ADJ</b> esforço/N ./PU	I/0 could/0 afford/0 / _UNDEF_ ten/0 / _UNDEF_ thousand/0 a/0 year/0 from/0 my/0 present/0 income/0 ./0 without/0 much/0 pain/0 ./0
EBDL1T1 (436):	Respondi/V_fmc que/KS há/Vprp anos/N não/ADV escrevia/V um/DETDarti trecho/N em/PRP prosa/N tão/ADVquant_dem <b>grande/ADJ</b> ./PU	I/0 said/0 it/0 was/0 the/0 longest/0 piece/0 of/0 continuous/0 prose/0 I'd/0 written/0 in/0 years/0 ./0
EBDL1T1 (541):	o/DETDartd Joe/PROP ./PU que/SPECrel tem/V problemas/N graves/ADJ na/PRP+DETDartd coluna/N ./PU anda/V_fmc sempre/ADV com/PRP um/DETDarti colete/N e/KC_co-fmc_co-vfin mal/ADV consegue/V_fmc servir/V ./PU o/DETDartd Rupe/PROP que/SPECrel teve/V um/DETDarti <b>grande/ADJ</b> acidente/N de/PRP carro/N há/Vprp uns/DETDartd anos/N e/KC_co-vfin coxeia/V_fmc de/PRP ambas/DETDartd as/DETDartd pernas/N se/KS é/V que/KS é/V possível/ADJ ./PU e/KC_co-vfin o/DETDartd Humphrey/PROP ./PU que/KS tem/V artrite/N nos/PRP+DETDartd pés/N e/KC uma/DETDarti articulação/N da/PRP+DETDartd anca/N de/PRP plástico/N ./PU	There's/0 Joe/0 ./0 he's/0 got/0 serious/0 back/0 trouble/0 ./0 wears/0 a/0 corset/0 all/0 the/0 time/0 and/0 can/0 barely/0 manage/0 to/0 serve/0 overarm/0 ./0 Rupert/0 ./0 who/0 was/0 in/0 a/0 bad/0 car/0 crash/0 a/0 few/0 years/0 ago/0 and/0 limps/0 with/0 both/0 legs/0 ./0 if/0 that's/0 possible/0 ./0 and/0 Humphrey/0 ./0 who/0 has/0 arthritis/0 in/0 his/0 feet/0 and/0 a/0 plastic/0 hip-joint/0 ./0
EBDL1T1 (852):	Não/ADV podia/V_fmc ser/V Universidade/PROP de/PROP Rummidge/PROP ./PU porque/KS já/ADV há/V uma/NUMcard ./PU e/KC_co-fmc_co-vfin então/ADVkc puseram-lhe/V_fmc+PERS o/DETDartd nome/N de/PRP Universidade/PROP James/PROP Watt/PROP ./PU um/DETDarti <b>grande/ADJ</b> inventor/N local/ADJ ./PU	They/0 couldn't/0 call/0 it/0 the/0 University/0 of/0 Rummidge/0 because/0 there/0 was/0 one/0 already/0 ./0 so/0 they/0 called/0 it/0 James/0 Watt/0 University/0 ./0 after/0 the/0 great/0 local/0 inventor/0 ./0
EBDL1T1 (875):	Tirando/V isto/SPECdem ./PU significou/V_fmc um/DETDarti <b>grande/ADJ</b> melhoramento/N na/PRP+DETDartd estação/N central/ADJ da/PRP+DETDartd cidade/N ./PU e/KC_co-fmc_co-vfin estou/V_fmc a/PRP menos/DETDartd doze/NUMcard minutos/N de/PRP Londres/PROP ./PU	Apart/0 from/0 that/0 ./0 it's/0 a/0 great/0 improvement/0 on/0 the/0 City/0 Centre/0 station/0 ./0 and/0 is/0 twelve/0 minutes/0 nearer/0 London/0 for/0 me/0 ./0
EBDL1T1	E/KC é/ADVfoc verdade/N que/ADVfoc ela/PERS anda/V_fmc	And/0 it's/0 true/0 that/0 she/0 carries/0 an/0 exceptionally/0

[\[This page in English\]](#)

COM

# Anotação gramatical: palavras flexionadas

As pesquisas simples permitem-lhe consultar a totalidade do COMPARA na direcção português-inglês **ou** na direcção inglês-português.

## De português para inglês

[lema="ajudar"]

Escreva uma palavra ou sequência de palavras *em português*. Use aspas em torno de cada palavra (ex. "Até" "logo")

Pesquisar de português para inglês

Apagar

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

## De inglês para português

Escreva uma palavra ou sequência de palavras *em inglês*. Use aspas em torno de cada palavra (ex. "See" "you")

Pesquisar de inglês para português

Apagar

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore connosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

<a href="#">EBDL1T1</a> (334):	E, de facto, a Alexandra tem-me <u>ajudado</u> .	And Alexandra has certainly been a help.
<a href="#">EBDL1T1</a> (506):	Mas o melhor de tudo é que <u>ajuda</u> a dormir.	But best of all, it helps you sleep.
<a href="#">EBDL1T1</a> (608):	Conheci a Amy há seis anos, quando a contrataram para <u>ajudar</u> a seleccionar os actores para a primeira série de <i>Os Vizinhos do Lado</i> .	I met Amy six years ago when she was hired to help cast the first series of <i>The People Next Door</i> .
<a href="#">EBDL1T1</a> (748):	Comprei todas as revistas de automóveis, na esperança de encontrar uma crítica qualquer ao carro que me <u>ajudasse</u> a decidir não o comprar.	I bought all the motoring magazines hoping that I would find some damning criticism of the car that would enable me to decide against it.
<a href="#">EBDL1T1</a> (1101):	Levou-me a galerias de arte e a teatros independentes e, com o argumento de que isso me <u>ajudaria</u> a perceber a televisão americana, a locais mais plebeus de lazer: Burger Kings e Donuts Delites com <i>drive-in</i> , recintos de <i>bowling</i> e um desafio de basebol.	She took me to art galleries and little theatres and, on the grounds that it would help me understand American television, to more plebeian places of resort: drive-in Burger Kings and Donut Delites, ten-pin bowling alleys, and on one occasion a baseball game.
<a href="#">EBDL1T1</a> (1341):	Embora tivesse uma certa lógica, este argumento não me <u>ajudou</u> muito, porque não só não tenho qualquer ideia para um programa novo, como, no estado de espírito em que me encontro, é altamente improvável que venha a tê-la.	Which was quite logical, but not particularly helpful, since I don't have an idea for a new show, and am unlikely to get one in my present state.
<a href="#">EBDL1T2</a> (334):	E Alexandra tem <b>ajudado</b> .	And Alexandra has certainly been a help.
<a href="#">EBDL1T2</a> (339):	De certa maneira, eu já sei disso, mas me <b>ajuda</b> ter Alexandra para esclarecer.	In a way I know that already, but it helps to have Alexandra spell it out.
<a href="#">EBDL1T2</a> (506):	Mas, o melhor de tudo é que <b>ajuda</b> você a dormir.	But best of all, it helps you sleep.
<a href="#">EBDL1T2</a> (1101):	Levou-me para galerias de arte e pequenos teatros e, com o pretexto de me <b>ajudar</b> a entender a tevê americana, para lugares mais plebeus: Drive-in Burger Kings e Donut Delites, pistas de boliche e, numa ocasião, para um jogo de beisebol.	She took me to art galleries and little theatres and, on the grounds that it would help me understand American television, to more plebeian places of resort: drive-in Burger Kings and Donut Delites, ten-pin bowling alleys, and on one occasion a baseball game.
<a href="#">EBDL1T2</a> (1341):	O que era muito lógico, mas não <b>ajudava</b> muito, já que eu não tenho qualquer idéia para um novo <i>show</i> e será muito difícil que venha a ter	Which was quite logical, but not particularly helpful, since I don't have an idea for a new show, and am unlikely to get one in

[This page is English]

# COMPARA Anotação gramatical: palavras ambíguas

As pesquisas simples permitem-lhe consultar a totalidade do COMPARA na direcção português-inglês ou na direcção inglês-português.

## De português para inglês

[word="gosto" & pos="N"]

Escreva uma palavra ou sequência de palavras *em português*. Use aspas em torno de cada palavra (ex. "Até" "logo")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

## De inglês para português

Escreva uma palavra ou sequência de palavras *em inglês*. Use aspas em torno de cada palavra (ex. "See" "you")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore connosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

BDL1T1 408):	Claro que actualmente tenho posses para mandar fazer camisas por medida, mas o ar <i>snob</i> dos camiseiros de Picadilly dissuade-me de lá entrar e as popelines às riscas expostas nas montras são demasiado afectadas para o meu <b>gosto</b> .	Of course, I could afford to have my shirts made to measure nowadays, but the snobby-looking shops around Picadilly where they do it put me off and the striped poplins in the windows are too prim for my taste.
BDL1T1 1642):	Por outras palavras, são feitas como os filmes, sendo cada cena filmada várias vezes de diversos ângulos e a distâncias focais diferentes, <i>take</i> após <i>take</i> , de preferência num local escolhido para o efeito, e não num estúdio, e depois montada a <b>gosto</b> do realizador.	In other words, they're made like movies, every scene being shot several times from various angles and focal distances, in take after take, on location rather than in a studio, and then edited at the director's leisure.
BDL1T1 2020):	É um jogador encantador, dá <b>gosto</b> vê-lo a defender, como se tivesse a bola atada aos pés, desviando os adversários como se fossem cordeirinhos.	He's a lovely player, thrilling to watch when he's running at a defence with the ball apparently tied to his feet, scattering them like sheep.
BDL1T1 2110):	Fica numa correnteza de salas por cima de uma loja de <i>t-shirts</i> e recordações de mau <b>gosto</b> em Carnaby Street.	It's a small set of rooms over a tatty tee-shirt and souvenir shop in Carnaby Street.
BDL1T2 408):	Claro que agora tenho condições de mandar fazer minhas camisas sob medida, mas essas lojas com ar de esnobe que a gente encontra em Piccadilly me fazem perder a vontade e, além disso, os tecidos de popelina listrados que eles põem nas vitrinas são muito certinhos para o meu <b>gosto</b> .	Of course, I could afford to have my shirts made to measure nowadays, but the snobby-looking shops around Picadilly where they do it put me off and the striped poplins in the windows are too prim for my taste.
BDL1T2 2110):	O lugar é um conjunto de salas em cima de uma loja de mau <b>gosto</b> de camisetas e uma de <i>souvenirs</i> em Carnaby Street.	It's a small set of rooms over a tatty tee-shirt and souvenir shop in Carnaby Street.
BDL2 346):	O <b>gosto</b> de Vic em música é limitado, mas exigente.	Vic's taste in music is narrow but keen.
BDL2 577):	Olha para cima, para as nuvens cinzentas que deslizam pelo céu, e para baixo, para o panorama de estreitos jardins de traseiras, alguns deles limpos e arranjados,	She looks up at the grey clouds scudding across the sky, down at a vista of narrow back gardens, some neat and trim with goldfish ponds and brightly

[This page is English]

# COMPARA Anotação gramatical: palavras ambíguas

As pesquisas simples permitem-lhe consultar a totalidade do COMPARA na direcção português-inglês ou na direcção inglês-português.

## De português para inglês [word="gosto" & pos="V"]

Escreva uma palavra ou sequência de palavras *em português*. Use aspas em torno de cada palavra (ex. "Até" "logo")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

## De inglês para português

Escreva uma palavra ou sequência de palavras *em inglês*. Use aspas em torno de cada palavra (ex. "See" "you")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore connosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

	silenciosa que ele não entendera.	sorrow that he hadn't understood at all.
<a href="#">EBJT1</a> (1029):	Claro que <b>gosto</b> .	Course I do.'
<a href="#">EBJT1</a> (2066):	-- <b>Gosto</b> delas -- disse Judy.	'I like them,' Judy said.
<a href="#">EBJT2</a> (649):	-- <b>Gosto</b> disto.	'I like this.
<a href="#">EBJT2</a> (833):	-- Deteve-se e a seguir prosseguiu: -- <b>Gosto</b> dele, percebes?	She paused and then she said, 'I just love him, you see.
<a href="#">EBJT2</a> (1827):	-- Não preciso de dizer que <b>gosto</b> de ti! -- gritara-lhe ele havia pouco tempo.	'I don't have to <i>say</i> I love you!' he'd shouted at her not long ago.
<a href="#">EBJT2</a> (2301):	-- Porque eu <b>gosto</b> , e também porque é fácil de dizer em inglês.	'Because I like it, because it's easy to say in English too.
<a href="#">EBJT2</a> (2342):	Não sei se <b>gosto</b> dela ou não, mas ela parece gostar de mim e claro que estou grata por isso.	I don't know if I really like her or not, but she seems to like me which of course I'm grateful for.
<a href="#">EBJT3</a> (195):	-- E <b>gosto</b> -- respondeu, deixando-os cair precipitadamente.	'I am,' she said, dropping them hurriedly.
<a href="#">EBJT3</a> (624):	-- Acho que <b>gosto</b> mais dele desde que deixou de ser Leslie.	'I think I prefer him since he stopped being Leslie.
<a href="#">EBKI1</a> (4288):	-- Mas confesso que não <b>gosto</b> do que está a passar-se mais do que você.	'But I must say, I don't like this any more than you do.
<a href="#">EBLC1</a> (236):	« <b>Gosto</b> dos que falam», disse Alice.	'I like them when they can talk,' Alice said.
<a href="#">EBOW1</a> (553):	<b>Gosto</b> , é claro, do Henry.	Of course, I am very fond of Harry.
<a href="#">EBOW1</a> (1317):	<b>Gosto</b> dos escândalos que atingem os outros, mas os escândalos que me dizem respeito não me interessam	I love scandals about other people, but scandals about myself don't interest me.

[\[This page in English\]](#)

# COMPARA Anotação gramatical: colocações

As pesquisas simples permitem-lhe consultar a totalidade do COMPARA na direcção português-inglês **ou** na direcção inglês-português.

## De português para inglês

[pos="V.\*"] "silêncio"

Escreva uma palavra ou sequência de palavras *em português*. Use aspas em torno de cada palavra (ex. "Até" "logo")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

## De inglês para português

Escreva uma palavra ou sequência de palavras *em inglês*. Use aspas em torno de cada palavra (ex. "See" "you")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore connosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)



- [PBFC2](#)
- [PBRF2](#)
- [PPEQ1](#)
- [PPJSA2](#)
- [PBFC3](#)
- [PMMC1](#)
- [PPEQ2](#)
- [PPLJ1](#)
- [PBPM1](#)
- [PMMC2](#)
- [PPEQ3](#)
- [PPMC1](#)
- [PBPM2](#)
- [PPCC1](#)
- [PPJS1](#)
- [PPSC1](#)
- [PBRF1](#)
- [PPCP1](#)
- [PPJSA1](#)
- [PPSC2](#)

## 4. Escolha os resultados

[Ajuda](#)

**Concordâncias**

( mostrar propriedades de alinhamento  esconder notas de tradução  mostrar categoria gramatical - disponível só para português)

**Distribuição das formas**

- Distribuição da categoria gramatical (disponível só para português)
- Distribuição dos lemas (disponível só para português)
- Distribuição das fontes
- Distribuição em texto original e traduzido
- Distribuição em termos da variante do português
- Distribuição em termos da variante do inglês
- Distribuição combinada das expressões de busca em português e inglês

Procurar

Limpar formulário

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore conosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

Esta é a interface DISPARA do COMPARA.

Última actualização desta página: 7 de Maio de 2007.

Última actualização do corpus: 18 de Junho de 2007.

[Enviar questões, comentários e sugestões](#)

# Resultados da pesquisa

Os resultados das buscas efectuadas no COMPARA podem ser usados para fins educacionais e investigação, desde que se mencione a fonte. Para citar textos específicos do corpus, seleccione o código azul ao lado de cada concordância de modo a obter a sua referência completa. Para se referir ao corpus como um todo, cite: **COMPARA 8.2.2** <http://www.linguateca.pt/COMPARA/> [31-Julho-2007]

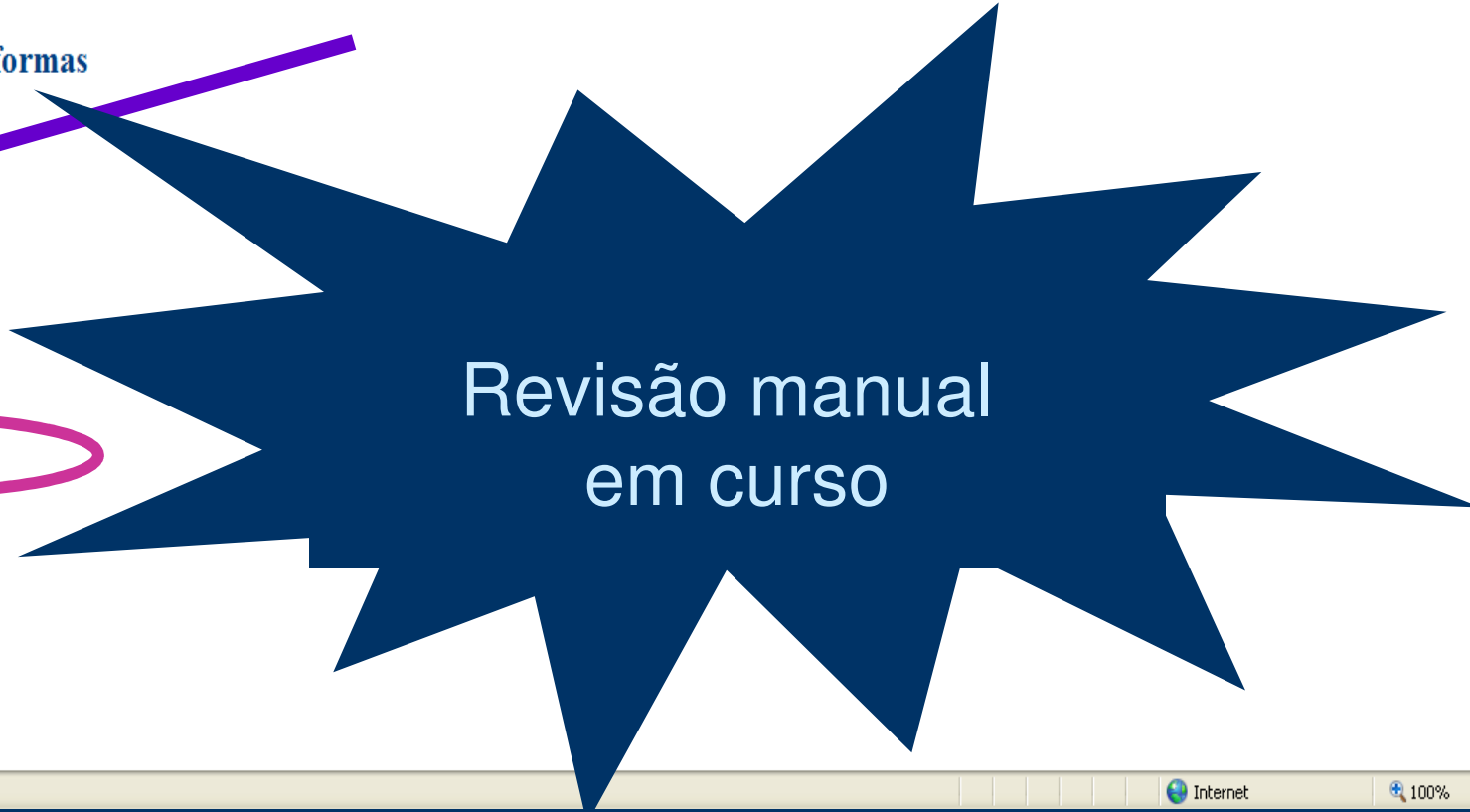
## Breve descrição do corpus usado nesta procura:

Palavras portuguesas	Palavras inglesas	Unidades de alinhamento
1436493	1549551	97714

18 ocorrências.

## Distribuição das formas

fez-se 2  
era 2  
impôs-lhe 1  
guardasse 1  
Fez-se 1  
guardou 1  
solicitando 1  
sagrado 1  
impos 1  
pedindo 1  
pedir 1  
pediu 1  
fazia 1  
querer 1  
seria 1



Revisão manual  
em curso

# **COMPARA usos e utilizadores**

## **Lingüistas e engenheiros / PLN**

tradução automática e outras aplicações

## **Lexicógrafos**

dicionários bilíngües

## **Teóricos da tradução**

estudos empíricos sobre tradução

## **Professores de tradução**

exercícios e problemas de tradução

# **COMPARA usos e utilizadores**

## **Professores de línguas**

exercícios e testes para os alunos

## **Tradutores e estudantes de tradução**

equivalências lingüísticas

## **Estudantes de línguas e qualquer pessoa que utilize o inglês na sua profissão**

Dicionário/gramática bilíngüe com muitos extras

# Alguns exemplos de utilização do corpus

## 1. Lexicografia bilíngüe

- Traduções mais prováveis para uma palavra polisêmica

## 2. Estudos de tradução

- Palavras com distribuição muito diferente em originais e traduções

[\[This page in English\]](#)

# 1. Lexicografia bilíngüe

## COMPARA - pesquisa simples

As pesquisas simples permitem-lhe consultar a totalidade do COMPARA na direcção português-inglês **ou** na direcção inglês-português.

### De português para inglês

Escreva uma palavra ou sequência de palavras *em português*. Use aspas em torno de cada palavra (ex. "Até" "logo")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

### De inglês para português

Escreva uma palavra ou sequência de palavras *em inglês*. Use aspas em torno de cada palavra (ex. "See" "you")

- Não fazer distinção entre maiúsculas e minúsculas
- Juntar acentos e cedilhas automaticamente

[Ajuda](#)

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore conosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

Esta é a interface DISPARA do COMPARA.

Última actualização desta página: 4 de Dezembro de 2005.

<a href="#">PPJS1</a> (3113):	Um <b>grande</b> estrondo fez as mulheres gritarem.	a <u>loud</u> noise made the women scream.
<a href="#">PPJS1</a> (3123):	Nós, parados em grupo, só víamos uma <b>grande</b> agitação de corpos e de tiras da camisa rasgada.	so that we, hanging back in a group, could only see a <u>confused</u> movement of bodies and strips of torn shirt.
<a href="#">PPJS1</a> (3149):	E uma <b>grande</b> aclamação saudou a erecção momentânea que ele conseguiu.	Then a <u>great</u> cry of triumph greeted the momentary erection that he finally managed.
<a href="#">PPJSA1</a> (23):	Fica aqui muito perto, seria um <b>grande</b> favor que me faziam.	It's close by and you could do me no <u>greater</u> favour.
<a href="#">PPJSA1</a> (211):	Ela procurava afastar do pensamento o roubo do carro, apertava carinhosamente as mãos do marido entre as suas, enquanto ele, com a cabeça baixa para que o motorista não pudesse ver-lhe os olhos pelo retrovisor, não parava de perguntar-se como era possível que tão <b>grande</b> desgraça lhe estivesse a acontecer a ele,	She tried not to think about the stolen car, squeezed her husband's hand affectionately, while he, his head lowered so that the driver would not see his eyes through the rear-view mirror, could not stop asking himself how it was possible that such a <u>terrible</u> tragedy should have befallen him,
<a href="#">PPJSA1</a> (304):	No fim das contas, estas ou as outras, não é assim tão <b>grande</b> a diferença entre ajudar um cego para depois o roubar e cuidar de uma velhice caduca e tatebitate com o olho posto na herança.	When all is said and done, there is not all that <u>much</u> difference between helping a blind man only to <u>rob</u> him afterwards and looking after some tottering and stammering old person with one eye on the inheritance.
<a href="#">PPJSA1</a> (587):	Tem porém um inconveniente, ser demasiado <b>grande</b> , tornaria difícil e dispendiosa a vigilância dos internados,	There is, however, one drawback, the <u>size</u> of the place is likely to make it both difficult and costly to keep an eye on those interned,
<a href="#">PPJSA1</a> (623):	Havia mais camaratas, corredores longos e estreitos, gabinetes que deviam ter sido de médicos, sentinas encardidas, uma cozinha que ainda não perdera o cheiro de má comida, um <b>grande</b> refeitório com mesas de tampo forrados de zinco, três celas acolchoadas até à altura de dois metros e forradas de cortiça daí para cima.	There were more wards, long and narrow corridors, rooms that must have been the doctors' rooms, dingy latrines, a kitchen that still reeked of bad cooking, a <u>vast</u> refectory with zinc-topped tables, three padded cells in which the bottom six feet of the walls had padding and the rest was lined with cork.
<a href="#">PPJSA1</a> (745):	Guiou as mãos do marido, ela própria tomou à sua conta o cego que encontrou mais a jeito, e com <b>grande</b> esforço conseguiram separá-los.	She guided her husband's hands, she herself took charge of the blind man whom she found more manageable, and with <u>much</u> effort, they managed to separate them.
<a href="#">PPJSA1</a> (991):	E ela, por causa da <b>grande</b> tristeza de estar cega e da grande alegria de ter recuperado o marido, a alegria e a	And she, because of her deep distress at being blind and her great joy on regaining <u>her</u> husband, joy and sorrow can

[\[This page in English\]](#)

# COMPARA - pesquisa avançada

As pesquisas avançadas dão-lhe a oportunidade de efectuar consultas mais sofisticadas e escolher as partes do COMPARA que pretende utilizar. Também poderá seleccionar diferentes tipos de resultados. Basta seguir os passos 1 a 4 abaixo.

## 1. Seleccione a direcção

[Ajuda](#)

- de português para inglês
- de inglês para português

Procurar

Limpar formulário

## 2. Preencha com o tipo de pesquisa pretendido

### 2.1 Pesquisa por palavra ou expressão

Digite uma palavra ou expressão: <a href="#">Ajuda</a>	Restrição de alinhamento (opcional): <a href="#">Ajuda</a>
<input type="text" value="grande"/>	<input type="text" value="great"/>
<input type="checkbox"/> Não fazer distinção entre maiúsculas e minúsculas	<input type="checkbox"/> Não fazer distinção entre maiúsculas e minúsculas
<input type="checkbox"/> Juntar acentos e cedilhas automaticamente	<input type="checkbox"/> Juntar acentos e cedilhas automaticamente

### 2.2 Mais opções de pesquisa

[Ajuda](#)

- (a) Procurar a palavra ou expressão inserida acima apenas em:
- (b) Ou não inserir nada acima e mostrar todos os resultados com:

- Notas de tradução
- Títulos
- Palavras ou expressões estrangeiras
- Adição de frases na tradução
- Supressão de frases na tradução
- Reordenamento de frases na tradução



- [PPEQ1](#)       [PPEQ2](#)       [PPEQ3](#)       [PPJS1](#)       [PPJSA1](#)  
 [PPJSA2](#)       [PPLJ1](#)       [PPMC1](#)       [PPSC1](#)       [PPSC2](#)

## 4. Escolha os resultados

[Ajuda](#)

**Concordâncias**

(  mostrar propriedades de alinhamento  esconder notas de tradução  mostrar categoria gramatical - disponível só para português)

**Distribuição das formas**

**Distribuição da categoria gramatical** (disponível só para português)

**Distribuição dos lemas** (disponível só para português)

**Distribuição das fontes**

**Distribuição em texto original e traduzido**

**Distribuição em termos da variante do português**

**Distribuição em termos da variante do inglês**

**Distribuição combinada das expressões de busca em português e inglês**

Procurar

Limpar formulário

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore connosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

Esta é a interface DISPARA do COMPARA.

Última actualização desta página: 7 de Maio de 2007.

Última actualização do corpus: 18 de Junho de 2007.

[Enviar questões, comentários e sugestões](#)

Procura inicial: "grande" %c. Pedido de : **concordância em contexto, distribuição combinada das expressões de busca em português e inglês.** Corpus: **De português para inglês .**

## Resultados da pesquisa

Os resultados das buscas efectuadas no COMPARA podem ser usados para fins educacionais e investigação, desde que se mencione a fonte. Para citar textos específicos do corpus, seleccione o código azul ao lado de cada concordância de modo a obter a sua referência completa. Para se referir ao corpus como um todo, cite: **COMPARA 8.2.2**  
<http://www.linguateca.pt/COMPARA/> [3-Agosto-2007]

### Breve descrição do corpus usado nesta procura:

Palavras portuguesas	Palavras inglesas	Unidades de alinhamento
1436493	1549551	97714

446 ocorrências.

### Distribuição combinada das expressões de busca em português e inglês

1259	"grande" %c
446	"grande" %c :COMPARA_ING_ANOT "great" %c
963	"great" %c

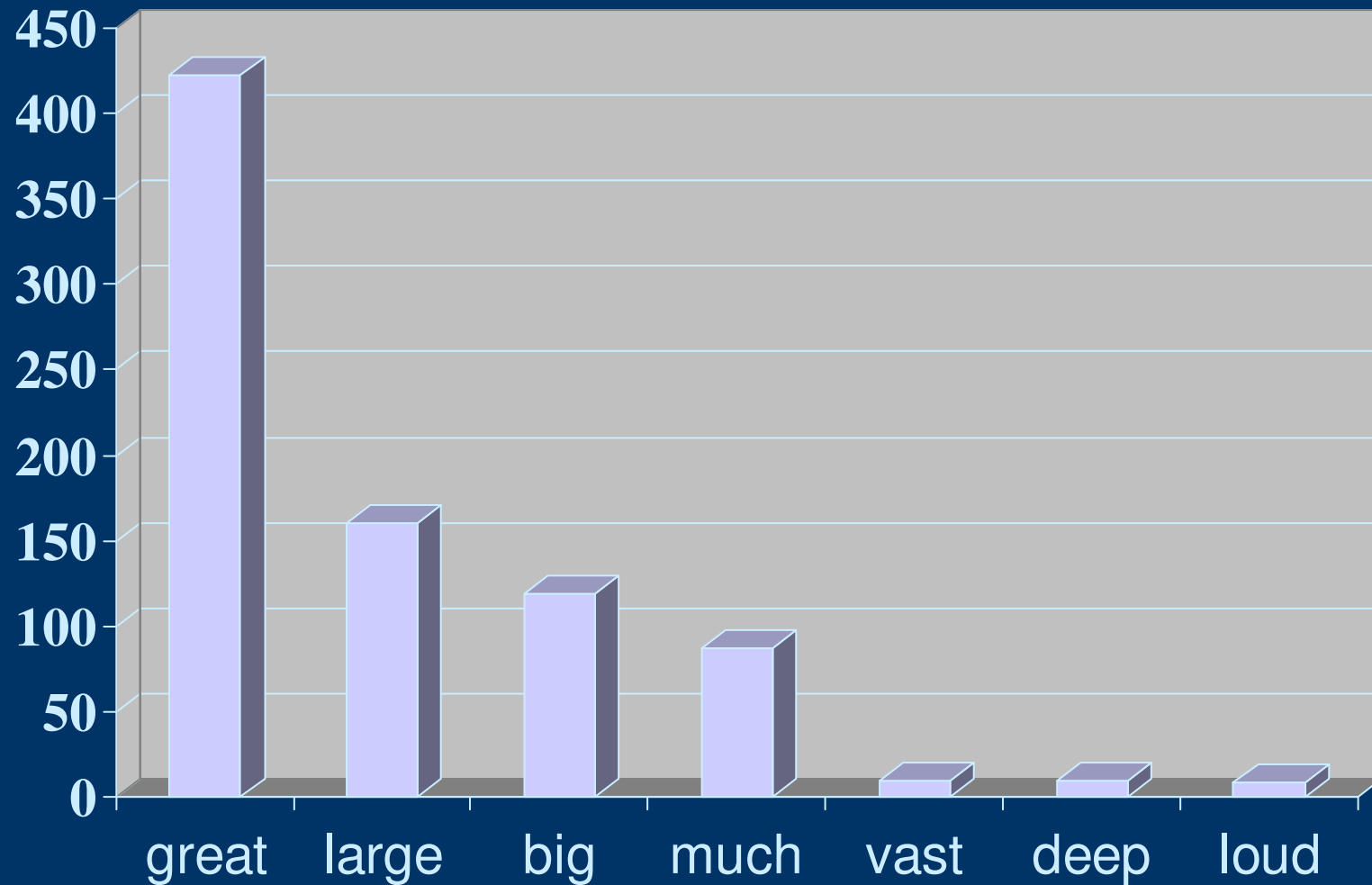
### Concordância

Procura: "grande" %c :COMPARA\_ING\_ANOT "great" %c.

<a href="#">EBDL1T1</a> (852):	Não podia ser Universidade de Rummidge, porque já há uma, e então puseram-lhe o nome de Universidade James Watt, um grande inventor local	They couldn't call it the University of Rummidge because there was one already, so they called it James Watt University, after the great local inventor
-----------------------------------	---	---

<a href="#">EBIM2</a> (734):	O <b>grande</b> F. S. Dibben.	The <b>great</b> F. S. Dibben.
<a href="#">EBIM3</a> (20):	Com <b>grande</b> alívio, ouvi a porta do quarto a abrir-se e passos no chão.	With <b>great</b> relief I heard the door of my bedroom open, and footsteps across the floor.
<a href="#">EBIM3</a> (353):	Desta vez, surpreendentemente, abriu-se uma enorme fissura e um pedaço <b>grande</b> , satisfatório, de betão soltou-se.	This time, surprisingly, a <b>great</b> fissure opened up and a large, satisfying piece of <b>concrete</b> came away.
<a href="#">EBIM3</a> (539):	A mãe leu aquela passagem várias vezes, com <b>grande</b> satisfação.	Mother read this line over a couple of times with great satisfaction.
<a href="#">EBIM3</a> (879):	Trabalhava a <b>grande</b> velocidade, mexendo e remexendo o enorme monte até ele ficar numa lama espessa, cinzenta.	She worked at great speed, turning and folding the huge pile in on itself till it was a stiff, grey sludge.
<a href="#">EBJB1</a> (293):	Ela trata-o com <b>grande</b> bondade e afecto; mantêm-se em contacto nos quarenta anos seguintes.	She treats him with great kindness and affection; they remain in touch for the next forty years.
<a href="#">EBJB1</a> (321):	Sainte-Beuve escreve a Matthew Arnold: <i>Salammbô</i> é o nosso <b>grande</b> acontecimento!>	Sainte-Beuve writes to Matthew Arnold: ' <i>Salammbô</i> is our great event!'
<a href="#">EBJB1</a> (473):	1846 - Sou apenas um lagarto literário que passa o dia a aquecer-se ao <b>grande</b> sol da Beleza.	1846 -- I am only a literary lizard basking the day away beneath the great sun of Beauty.
<a href="#">EBJB1</a> (531):	Nunca foi tão manifesto o ódio a tudo o que é <b>grande</b> -- desdém pela Beleza, desprezo pela literatura.	Never has hatred for everything great been so manifest -- disdain for Beauty, execration of literature.
<a href="#">EBJB1</a> (592):	Juliet Herbert é um <b>grande</b> buraco ligado com fio.	Juliet Herbert is a great hole tied together with string.
<a href="#">EBJB1</a> (891):	«Com um salto deixava-me cair na <b>grande</b> pele branca de urso, que adorava, e cobria-lhe de beijos a cabeça enorme».	'With one bound I would throw myself on the large white bearskin, which I adored, and cover its great head with kisses.'
<a href="#">EBJB2</a> (81):	Havia, por exemplo, um preguiça especialmente preguiçoso -- uma criatura de gostos muito requintados, nosso afiançar-lhes pessoalmente -- que	There was a particularly relaxed sloth, for instance -- an exquisite creature, I can vouch for it personally -- which had scarcely got down to the foot of its tree

# Traduções de “grande”



[\[This page in English\]](#)

## 2. Estudos de tradução

# COMPARA - pesquisa avançada

As pesquisas avançadas dão-lhe a oportunidade de efectuar consultas mais sofisticadas e escolher as partes do COMPARA que pretende utilizar. Também poderá seleccionar diferentes tipos de resultados. Basta seguir os passos 1 a 4 abaixo.

### 1. Selecciona a direcção

Procurar

Limpar formulário

[Ajuda](#)

- de português para inglês
- de inglês para português

### 2. Preencha com o tipo de pesquisa pretendido

#### 2.1 Pesquisa por palavra ou expressão

Digite uma palavra ou expressão: <a href="#">Ajuda</a>	Restrição de alinhamento (opcional): <a href="#">Ajuda</a>
<input type="text" value="“nodded”"/>	<input type="text"/>
<input type="checkbox"/> Não fazer distinção entre maiúsculas e minúsculas	<input type="checkbox"/> Não fazer distinção entre maiúsculas e minúsculas
<input type="checkbox"/> Juntar acentos e cedilhas automaticamente	<input type="checkbox"/> Juntar acentos e cedilhas automaticamente

#### 2.2 Mais opções de pesquisa

[Ajuda](#)

(a) Procurar a palavra ou expressão inserida acima apenas em:

(b) Ou não inserir nada acima e mostrar todos os resultados com:

- Notas de tradução
- Títulos
- Palavras ou expressões estrangeiras
- Adição de frases na tradução
- Supressão de frases na tradução
- Reordenamento de frases na tradução

- [TLEQ1](#)     [TLEQ2](#)     [TLEQ3](#)     [TJSA1](#)     [TJSA2](#)  
 [PPJSA2](#)     [PPLJ1](#)     [PPMC1](#)     [PPSC1](#)     [PPSC2](#)

## 4. Escolha os resultados

[Ajuda](#)

**Concordâncias**

( mostrar propriedades de alinhamento  esconder notas de tradução  mostrar categoria gramatical - disponível só para português)

**Distribuição das formas**

**Distribuição da categoria gramatical** (disponível só para português)

**Distribuição dos lemas** (disponível só para português)

**Distribuição das fontes**

**Distribuição em texto original e traduzido**

**Distribuição em termos da variante do português**

**Distribuição em termos da variante do inglês**

**Distribuição combinada das expressões de busca em português e inglês**

Procurar

Limpar formulário

[PESQUISA SIMPLES](#)

[PESQUISA AVANÇADA](#)

[Conteúdo: textos e resumo quantitativo do corpus](#)

[Agradecimentos](#)

[Ajuda](#)

[Aula Prática](#)

[Breve descrição do COMPARA](#)

[Colabore connosco](#)

[Construção do COMPARA](#)

[Dúvidas dos utilizadores](#)

[O sistema DISPARA](#)

[Quem somos](#)

[Trabalhos publicados](#)

Esta é a interface DISPARA do COMPARA.

Última actualização desta página: 7 de Maio de 2007.

Última actualização do corpus: 18 de Junho de 2007.

[Enviar questões, comentários e sugestões](#)

Breve descrição do corpus usado nesta procura:

Palavras portuguesas	Palavras inglesas	Unidades de alinhamento
1436493	1549551	97714

135 ocorr

Denominador comum : 100 000

Distribuição de texto original ou traduzido

Texto original	109 [ 822160 ]
Texto traduzido	26 [ 727391 ]

Concordância

Procura: "noddod".

$$822\ 160 \div 8,22160 = 100\ 000$$

$$727\ 391 \div 7,27160 = 100\ 000$$

<a href="#">EBDL1T1</a> (1571):	Sally <b>noddod</b> and v	
<a href="#">EBDL1T1</a> (1720):	Hal <b>noddod</b> gravely.	O Hal <u>acenoou</u> com a cabeça em sinal de <u>assentimento</u> , com um ar pesaroso.
<a href="#">EBDL1T1</a> (1782):	«The best audience in the world is going to be stymied if they turn find that it's all about the prime of life said his agreement.	-- Até o melhor público do mundo ficaria <u>era de uma comédia e</u> <u>fez</u> <u>mostrar</u>
<a href="#">EBDL1T1</a> (1908):	I <b>noddod</b> .	-- <u>Confirmei</u> com um gesto de cabeça.
<a href="#">EBDL1T2</a> (1571):	Sally <b>noddod</b> and went back to her reading.	Sally <u>acquiesceu</u> e voltou a ler.
<a href="#">EBDL1T2</a>	Hal <b>noddod</b> gravely.	<u>acenando</u> com a cabeça gravemente.

$$109 \div 8,22160 = 13,2$$

$$26 \div 7,27160 = 3,6$$

em 100 000

# Distribuição de “nodded” em texto original e traduzido

Inglês original

13,2 / 100 mil palavras

Inglês traduzido do português

3,6 / 100 mil palavras

**3,7 vezes mais frequente em inglês original!**



# Tarefas

## 1. Lexicografia bilíngüe

Traduções mais prováveis de  
tempo  
sombra  
sensação

# Tarefas

## 2. Estudos de tradução

Distribuição em português original e português traduzido de:

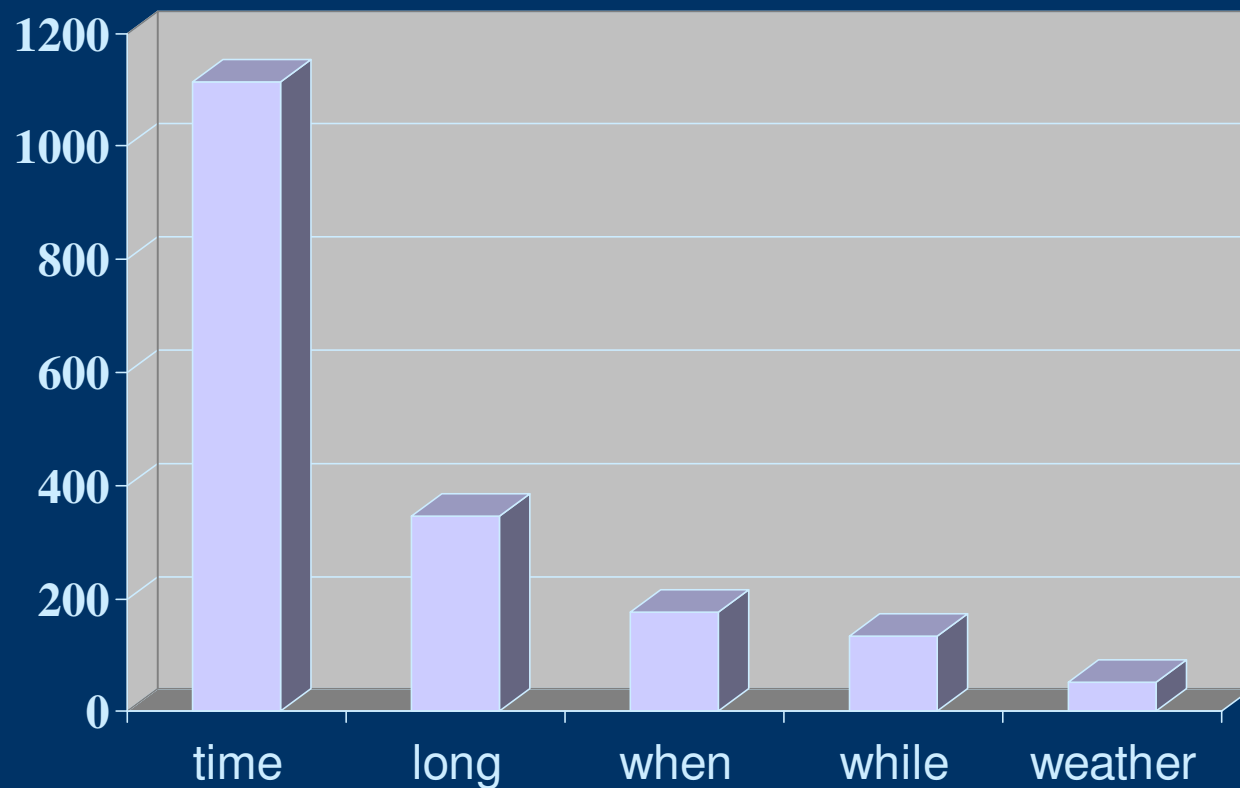
Diferente(s)

Simplesmente

Lema = “rezar”

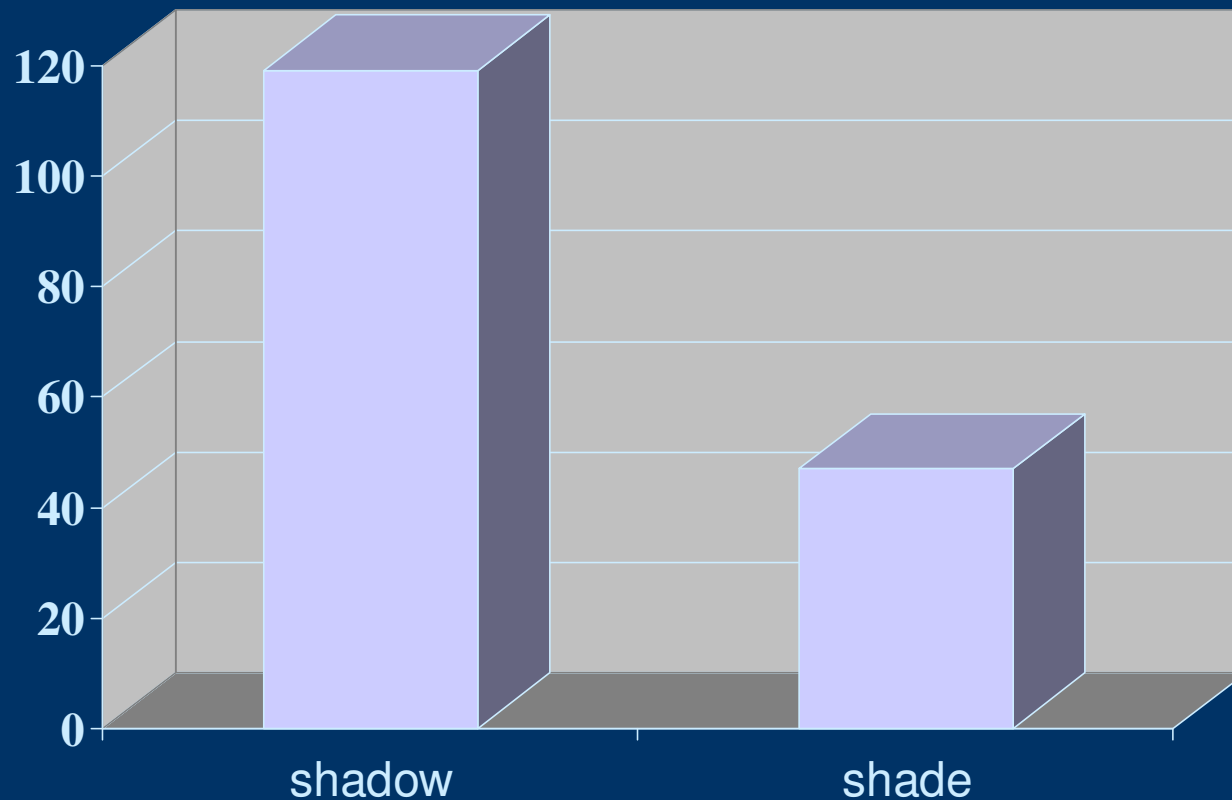
	descansar sobre vosso leito.	you will be told to sleep in thy bed.»
<a href="#">PBPC3</a> (432):	« <b>Obrigada</b> por levar-me com você», disse ela.	«Thank you for taking me with you,» she said.
<a href="#">PBPC3</a> (1111):	« <b>Obrigado</b> por acolher-nos», disse Elias.	«Thank you for receiving us,» Elijah said.
<a href="#">PBPM1</a> (393):	<b>Obrigada</b> , pessoal, disse minha mãe, amanhã eu pego o da pamonha.	'Thank you, folks. Tomorrow it's the ice-cream man,' my mother said.
<a href="#">PBPM2</a> (408):	Tinha um bilhete também, com letra de criança: <b>Obrigado</b> , Máiquel.	There was also a note, in a child's handwriting: Thank you, Máiquel.
<a href="#">PBPM2</a> (1206):	<b>Obrigado</b> , eu falei.	Thanks, I said.
<a href="#">PBPM2</a> (1638):	Não, <b>obrigado</b> .	No, thanks.
<a href="#">PBPM2</a> (1643):	Muito <b>obrigado</b> , mas eu não vou matar ninguém.	Thank you very much, but I'm not going to kill anybody.
<a href="#">PBRF1</a> (61):	«Muito <b>obrigada</b> », arquejou afinal.	'Thank you very much,' she gasped finally.
<a href="#">PBRF1</a> (86):	« <b>Obrigada</b> , obrigada!»	'Thank you, thank you!'
<a href="#">PBRF1</a> (86):	«Obrigada, <b>obrigada</b> !»	'Thank you, thank you!'
<a href="#">PBRF1</a> (118):	«Meu amigo», a letra de Angélica era miúda e difícil de ler, «muito <b>obrigada</b> por ter salvo a minha vida.	'My friend,' Angélica's tiny writing was hard to read. 'Thank you very much for saving my life.
<a href="#">PBRF1</a> (576):	«Sim, <b>obrigado</b> .»	'Yes. Thanks.'
<a href="#">PBRF1</a> (1728):	<b>Obrigado</b> , querido.	Thank you, dear.
<a href="#">PBRF2</a> (1239):	Muito <b>obrigado</b> .»	Thanks.»

# Resultados das Tarefas



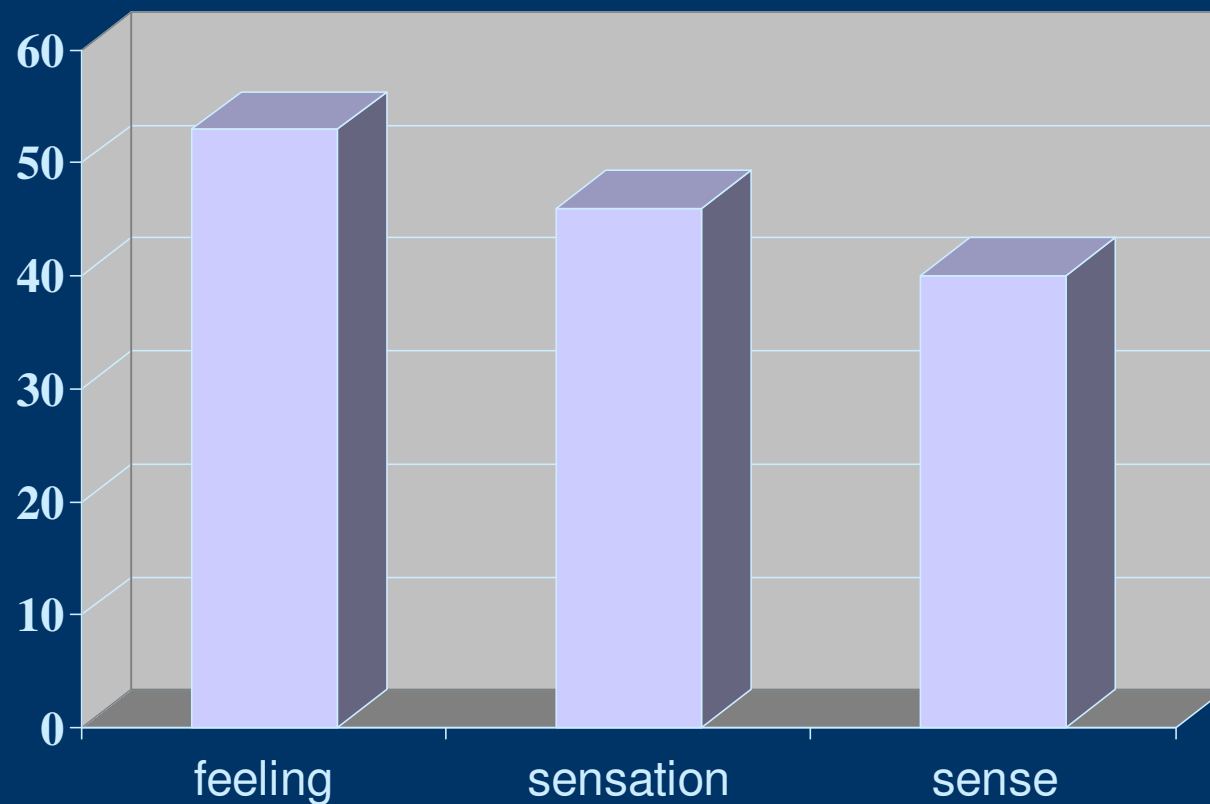
Traduções de “tempo” no COMPARA 8.2

# Resultados das Tarefas



Traduções de “sombra” no COMPARA 8.2

# Resultados das Tarefas



Traduções de “sensação” no COMPARA 8.2

# Resultados das Tarefas

	PT original	PT traduzido
Diferente(s)	16,4	31,1
Simplesmente	5,4	15,8
Lema="rezar"	11,9	5,5

ocorrências em cada 100 mil palavras no COMPARA 8.2