

**Stella E. O. Tagnin**  
**Estudos em Tradução –**  
**Teorias, Práticas e Tecnologias**  
**PUC-RS - 10-11/10/2014**  
1ª. Parte

# LINGUÍSTICA DE CORPUS E TRADUÇÃO

# Resumo do curso

## Parte 1:

- Fluência/Naturalidade e Convencionalidade
- Noções gerais sobre Linguística de Corpus

## Parte 2:

- Como a LC pode ajudar o tradutor
  - Fonte de referência – consulta direta aos *corpora on-line*
  - Metodologia para a construção de *corpora* personalizados

## Parte 3:

- A exploração de *corpora* personalizados – WST e AntConc
  - WordList
  - Keywords
  - Concord

## Parte 4:

- A construção de um glossário de especialidade
  - Extração de termos
  - Identificação de equivalentes
  - Seleção de exemplos

# PARTE 1:

Convencionalidade e Linguística de Corpus

# Fluência e Naturalidade

- É um dos objetivos do tradutor?
- De que consiste?
- Como alcançá-la?
- **Conhecendo as convenções da língua**

# Convencionalidade

- “que é de uso ou de praxe; consolidado pelo uso ou pela prática” (Houaiss)
- “que obedece a **padrões aceitos; não original**, comum” (Houaiss)
- “**Consagrado** ou aprovado pelo uso, pela experiência” (Aurélio)

# Frasesologias / Convenções

**Regência:** insist **on**, instructor **in**, dressed **in**, good **at**  
insistir **em**, instrutor **de**, vestido **de**, bom **em**

**Phrasal Verbs:** look up, get across, make up

## **Colocações Adjetivas:**

public television – televisão educativa

Japanese current – corrente do Japão

hard luck – má sorte

## **Colocações Nominais:**

labor pains – dores de parto

Seashore – beira-mar

cost of living – custo de vida

# Frasesologias / Convenções

## **Colocações Verbais:**

**pay** a compliment – fazer um elogio

**take** a step – dar um passo

**make** a date – marcar um compromisso

## **Colocações Adverbiais:**

**lavishly** illustrated – fartamente / ricamente ilustrado

lie **outright** – mentir descaradamente

## **Binômios e Multinômios:**

dead or alive – vivo ou morto

cats and dogs – cães e gatos, gatos e cachorros

red, white and blue – vermelho, azul e branco, vermelho branco e azul

# Outras fraseologias

**Expressões Idiomáticas:** kick the bucket; blow one's top

**Provérbios:** Not all that glitters is gold.

**Fórmulas Situacionais:** Act your age; It takes two to tango.

E mais outras:

Pelo presente contrato particular de prestação de serviços as partes abaixo qualificadas, de comum acordo, contratam entre si os serviços descritos...

○ **objetivo** deste **artigo** é discutir / analisar / apresentar / refletir (1.240.000 hits no Google)

Este **artigo** tem por **objetivo** apresentar/ discutir /analisar... (431.000 hits no Google)



Assim...

**há muito mais expressões fixas  
na língua do que se imagina!**

- Como fazer para reconhecê-las?

É onde entra a  
Linguística de  
Corpus

---

# O que é um corpus hoje?

Uma coletânea de textos em formato eletrônico, compilada segundo critérios específicos, considerada representativa de uma língua (ou da parte que se pretende estudar), destinada à pesquisa

# O que permite...

- **examinar uma grande quantidade de ocorrências**
- **de linguagem autêntica**
- **de uma só vez**
- **por meio de CONCORDÂNCIAS**

# E descobrir

- Padrões lexicais
- Padrões gramaticais

Padrões = **repetições**

# Concordância para "image\*"

## N Concordance

1 am na aba inferior do módulo Library. Com as **imagens** classificadas, você pode acionar os f  
2 ior. Quando for importar uma nova leva de **imagens**, não se assuste pelo fato das fotos i  
3 com as bandeirinhas, volte classificando as **imagens** com uso dos rótulos colondos e das e  
4 r do disco. A primeira opção apenas retira a **imagem** da biblioteca. A segunda apaga definit  
5 iblioteca. A segunda apaga definitivamente a **imagem**. Depois da primeira passada marcand  
6 do cabeçalho do menu. Para adicionar outras **imagens** a uma coleção já existente, basta sel  
7 do Collections também servem para agrupar as **imagens** exibidas na mesa de luz. No menu, Key  
8 a. Para criar uma nova coleção, selecione as **imagens** que deseja e clique no botão +, ao la  
9 s na última importação. Para controlar quais **imagens** você deseja exibir na biblioteca, use  
10 tão. O Lightroom permite cnar coleções de **imagens** que podem ser acessadas rapidamente p  
11 tl+Backspace. Sempre que tentar deletar uma **imagem** da biblioteca do Lightroom, o programa  
12 apenas serão gravadas na hora de exportar as **imagens**. O menu de importação traz informa  
13 ormações essenciais para o armazenamento das **imagens**. Se você estiver importando fotos que  
14 leitor do cartão). Depois de importadas, as **imagens** ficam disponíveis para você manipular  
15 ftwares. Assim, com ele é possível baixar as **imagens** do cartão, organizá-las, editá-las, f  
16 adicado em São Paulo, onde mantém o banco de **imagens** Argos ([www.argosfoto.com.br](http://www.argosfoto.com.br)) e minist  
17 para diminuir ou apagar as luzes em volta da **imagem** selecionada. Depois de importadas a  
18 ecíficos. No modo Grid, selecione o grupo de **imagens** que correspondem à mesma situação e m  
19 bre a seta correspondente. Para editar as **imagens** que você importou, classificá-las, ap  
20 tidas no próprio arquivo. Na era digital, as **imagens** ganham muita mobilidade e precisam se

# O tradutor

- é um **PRODUTOR DE TEXTOS**



**Linguística  
de Corpus**

- precisa de:
  - **EQUIVALENTE** – tradução
  - **CONTEXTO** – exemplo de uso
  - **CONVENCIONALIDADE** – padrões linguísticos e textuais
  - Informações sobre a **TIPOLOGIA TEXTUAL**
  - Especificidades da **CULTURA**

# Corpora para Tradução

Corpora *on-line*

- **Corpora bilíngues inglês-português**

- COMPARA: [www.linguateca.pt/COMPARA](http://www.linguateca.pt/COMPARA)

- CorTec:  
[http://www.fflch.usp.br/dlm/comet/consulta\\_cortec.html](http://www.fflch.usp.br/dlm/comet/consulta_cortec.html)

- CorTrad:  
[http://www.fflch.usp.br/dlm/comet/consulta\\_cortrad.html](http://www.fflch.usp.br/dlm/comet/consulta_cortrad.html)

- **Corpora personalizados**

- Compilados pelo tradutor



# CorTec 2005

- 5 corpora comparáveis:
  - Culinária – receitas
  - Meio ambiente - Ecoturismo
  - Informática - Geral
  - Cardiologia – Hipertensão
  - Direito – instrumentos contratuais
- inglês – português
- aproximadamente 200.000 palavras cada

# CorTec 2005

- Ferramentas
  - lista de frequência
  - concordanciador
    - igual a
    - começando com
    - terminando com
    - Contendo
  - n-gramas

# CorTec 2008

## 14 corpora (CNPq)

- Ecoturismo
- Hipertensão
- Instrumentos Contratuais
- Astronomia
- Insuficiência Renal
- Linguística
- Medidores Eletromagnéticos de Vazão
- Suplementos Nutricionais
- Computação
- Futebol
- Café
- Turismo Cultural
- Culinária 1 e 2

# Cortec 2013

- **Culinária**
- Ecoturismo**
- Hipertensão**
- Informática**
- Instrumentos Contratuais**
- Astronomia**
- Insuficiência Renal**
- Linguística**
- Medidores Eletromagnéticos de Vazão**
- Suplementos Nutricionais**
- Computação**
- Futebol atualizar**
- Café**
- Turismo Cultural**
- Culinária 2**
- Prostodontia**
- Fotografia**
- Autoclaves**
- Moda**
- Turismo – Hotelaria**



[Projeto](#)

[Equipe](#)

[CorTec](#)

[CorTrad](#)

[Publicações](#)

[Corpora Online](#)

# Bem-Vindos

## PROJETO **CO**rpus **M**ultilíngue para **E**nsino e **T**radução

Projeto CoMET - FFLCH/USP



[Projeto](#)

[Equipe](#)

[CorTec](#)

[CorTrad](#)

[Publicações](#)

[Corpora Online](#)

[Início](#)

## Publicações

[Artigos](#)

[Apresentações](#)

[Minicursos e Tutoriais](#)

[Livros Publicados](#)

[Revistas Organizadas](#)

[Capítulos de Livros](#)

---

### Artigos publicados

[TAGNIN, S.E.O.]

TAGNIN, S. E. O. . Corpus-driven terminology in Brazil. In: Armelle Le Bars Poupet; Claudia Xatara. (Org.). Cahiers de lexicologie - Dynamique de la recherche en lexicologie, lexicographie et terminologie au Brésil. 101ed.Paris: Classiques Garnier, 2012, v. 2012-2, p. 169-182.

DAYRELL, C. ; CANDIDO Jr., A. ; LIMA, G. ; MACHADO JR., D. ; COPESTAKE, A. ; FELTRIM, V. ; TAGNIN, S. E. O. ; ALUISIO, S. M. . Rhetorical Move Detection in English Abstracts: Multi-label Sentence Classifiers and their Annotated Corpora. In: LREC (Language Resources and Evaluation) 2012, 2012, Istanbul. LREC 2012 Proceedings, 2012.

TAGNIN, S. E. O. ; TEIXEIRA, E. D.. Translator-oriented, corpus-driven technical glossaries: the case of cooking terms. Corpora, v. 7, p. 51-67, 2012.



# CORTEC

English

Busca | O que é

## Passo 1/3: Escolhendo os *Corpora*.

Abaixo estão listados os *corpora* que compõem o CorTec.  
Selecione os *corpora* que deseja pesquisar:

### Lista de *Corpora*

- Culinária
- Ecoturismo
- Hipertensão
- Informática
- Instrumentos Contratuais
- Astronomia
- Insuficiência Renal
- Linguística
- Medidores Eletromagnéticos de Vazão
- Suplementos Nutricionais
- Computação
- Café
- Turismo Cultural
- Culinária 2
- Prostodontia
- Fotografia
- Autoclaves
- Moda
- Turismo - Hotelaria
- Futebol
- Futebol
- Futebol
- Futebol

### Língua

- Português
- Inglês

>> Próximo Passo



Corpus Multilíngue para Ensino e Tradução

USP

Sej

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

**CORT**

English

Busca O que é

### Passo 2/3: Escolhendo a Ferramenta.

Selecione abaixo a ferramenta a ser aplicada:

#### Lista de Ferramentas

- Concordanciador
- Gerador de Lista de Palavras
- Gerador de N-Gramas

>> Próximo Passo

CorTec - *Corpus* Técnico-Científico





Corpus Multilíngue para Ensino e Tradução

publicações no link Artigos

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

busca | O que e

### Passo 3/3: Aplicando a Ferramenta sobre o *Corpus*.

Configure a ferramenta usando as opções abaixo:

#### Concordanciador

Expressão ou Palavra:

Diferenciar maiúsculas:  Igual a  
 Começando com  
 Terminando com  
 Contendo

Tamanho do contexto reduzido (caracteres):

**>> Aplicar Concordanciador**

**Escolher outra Ferramenta <<**

**CorTec - *Corpus* Técnico-Científico**

Login

Internet

100%



USP

Corpus Multilíngue para Ensino e Tradução

Seja bem vindo ao s

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Foram encontradas 1993 ocorrências!! Fazer **DOWNLOAD** do resultado.

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 ulos esqueléticos e **cardiacos** 15, 18, 20 , f  
2 inistração de Mg em **cardiopatas** causa a redu  
3 ular esquelético e **cardíaco** que os valores  
4 rovidos de efeito "**cardioprotetor**". Entreta  
5 duções do dé- bito **cardíaco** e do volume int  
6 sobre a fisiologia **cardiovascular** e exigem  
7 a de insuficiência **cardíaca** congestiva, hem  
8 idade e mortalidade **cardiovas-** cular. Os ef  
9 ia de insuficiência **cardíaca** e re- nal. Qua  
10 é indicada para os **cardiopatas**, especialme  
11 e- dução do débito **cardíaco** e do fluxo sang  
12 ca e de frequência **cardíaca** induzidas pelo  
13 metabólica 21 . Os **cardioseletivos** causam  
14 prolol iedurilesl) (**cardioseletivo**), 200 mg  
15 seu possível efeito **cardioprotetor**. Podem se  
16 xiste insuficiência **cardíaca** congestiva ass



USP

Corpus Multilíngue para Ensino e Tradução

Projeto Comet!

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Ocorrência 2

cos, em

hipertensos, produz resistência ao tratamento exigindo maiores quantidades de drogas para controle de pressão 7 .

A administração de Mg em **cardiopatias** causa a redução da resistência vascular sistêmica 25, 26 , em coronariopatas com aumento secundário da contratilidade 27 , mas não produz

Ocorrência 3

inação (comunicação pessoal).

Apesar destes fatores de erro, o nível de Mg linfocitário correlaciona-se melhor com o "pool" muscular esquelético e **cardíaco** que os valores séricos ou eritrocitários 16-20 .

Há correlação entre os níveis sérico-plasmáticos e as outras variáveis que avaliam o Mg nos o

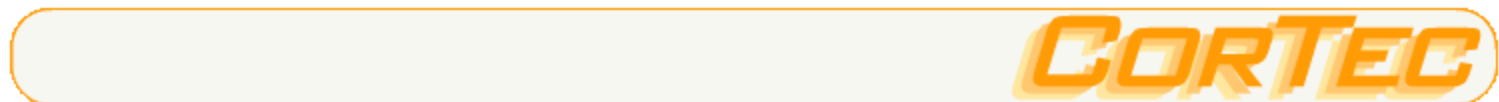


Corpus Multilíngue para Ensino e Tradução

US

indo ao site do Projeto Com

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nformativo
- Contato
- ite FFLCH
- Site USP



Busca O que é

### Passo 2/3: Escolhendo a Ferramenta.

Selecione abaixo a ferramenta a ser aplicada:

#### Lista de Ferramentas

- Concordanciador
- Gerador de Lista de Palavras
- Gerador de N-Gramas

>> Próximo Passo

CorTec - *Corpus* Técnico-Científico

Login

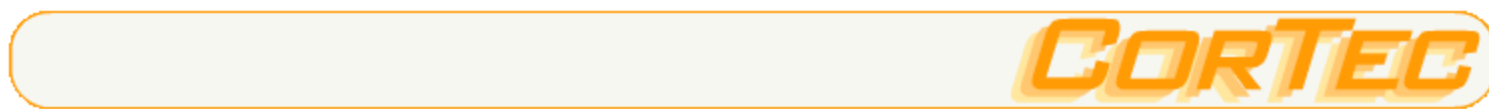


Corpus Multilíngue para Ensino e Tradução

US

Conheça

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP



Englis

Busca O que é

### Passo 3/3: Aplicando a Ferramenta sobre o *Corpus*.

Configure a ferramenta usando as opções abaixo:

#### Gerador de Lista de Palavras (Contador padrão)

Ordenado por:

- Ordem de frequência (decrecente)
- Ordem Alfabética

>> *Aplicar Gerador de Lista de Palavras*

#### Gerador de Lista de Palavras (Contador por palavra)

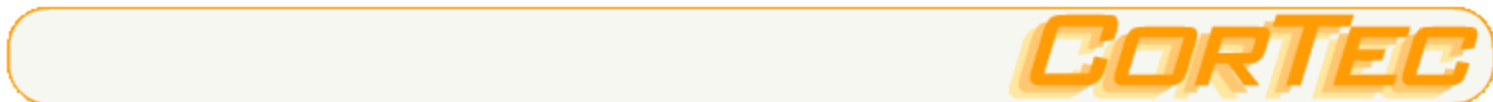
Número de palavras acima e abaixo da palavra escolhida:

Digite a palavra de busca:

>> *Aplicar Gerador de Lista de Palavras*



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nformativo
- Contato
- ite FFLCH
- Site USP



English

Busca O que é

Aguarde, executando ferramenta. Esta ação pode demorar alguns minutos.

**Hipertensão Ptg**

Fazer **DOWNLOAD** do Resultado.

<b>Ocorrências (tokens), considerando a Lexia Complexa de Nomes Próprios</b>
Total de Ocorrências: 442275
Total de Ocorrências que aparecem uma vez: 19133
Total de Ocorrências que aparecem mais de uma vez: 423142

<b>Palavras únicas/formas (types), considerando a Lexia Complexa de Nomes Próprios</b>
Total de Palavras: 35058
Total de Palavras que aparecem uma vez: 17259
Total de Palavras que aparecem mais de uma vez: 17799

<b>Índice Vocabular (token/type ratio)</b>
Token/Type: 0.5077



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nformativo
- Contato
- ite FFLCH
- Site USP

## TABELA DE FREQUÊNCIA

Pos.	Palavra	Freq.
1	de	18945
2	a	13959
3	e	10322
4	o	6447
5	da	5928
6	the ←	5721
7	em	5484
8	do	5331
9	que	5269
10	of ←	4654
11	com	4348
12	os	3563
13	and ←	3452
14	in	3358
15	na	2640
16	para	2620
17	no	2613
18	as	2507
19	to	2485
20	não	2232
21	dos	2186
22	por	1962
23	ou	1841
24	é	1820
25	pacientes	1663
26	como	1588
27	um	1562



Corpus Multilíngue para Ensino e Tradução

Conheça o CorTrad

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

25	pacientes	1663
26	como	1588
27	um	1563
28	is	1502
29	mais	1470
30	with	1459
31	à	1440
32	se	1371
33	ao	1305
34	foi	1301
35	arterial	1289
36	for	1275
37	uma	1226
38	ser	1153
39	nos	1151
40	das	1085
41	that	1037
42	entre	1036
43	foram	1035
44	pressure	978
45	blood	955
46	são	942
47	risco	840
48	or	827
49	hipertensão	826
50	estudo	803
51	tratamento	803
52	pressão	759
53	níveis	735





US

Corpus Multilíngue para Ensino e Tradução

/ inglês

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

35033	diabéti-cos	1
35034	editorial	1
35035	sub-distritos	1
35036	gestivos	1
35037	matúria	1
35038	descoberto	1
35039	analíticas	1
35040	labilidade	1
35041	restando	1
35042	ativamente	1
35043	rá-pida ←	1
35044	seguí-am	1
35045	atu-ante ←	1
35046	88	1
35047	comu ←	1
35048	popula-ções ←	1
35049	caminhe	1
35050	nomeamos	1
35051	azoxy	1
35052	orifícios	1
35053	desfibrilador	1
35054	acianótico	1
35055	tonicidade	1
35056	ierigorli	1
35057	lengthens	1
35058	stop-hypertension	1

>> Fazer Nova Busca

## Parte 2a:

Como a Linguística de Corpus pode ajudar o tradutor: Fonte de referência  
– consulta direta aos *corpora on-line*

# Tarefa 1

- ***Situação: Você tem uma tradução sobre hipertensão, mas não conhece o vocabulário, muito menos os equivalentes em português.***
- Faça uma lista de frequência para o corpus de Hipertensão em inglês.
- Quais as palavras lexicais mais frequentes?
- São correspondentes às em português?



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

Fazer **DOWNLOAD** do Resultado.

Ocorrências (tokens), considerando a Lexia Complexa de Nomes Próprios
Total de Ocorrências: 403764
Total de Ocorrências que aparecem uma vez: 12881
Total de Ocorrências que aparecem mais de uma vez: 390883

Palavras únicas/formas (types), considerando a Lexia Complexa de Nomes Próprios
Total de Palavras: 23420
Total de Palavras que aparecem uma vez: 11631
Total de Palavras que aparecem mais de uma vez: 11789

Índice Vocabular (token/type ratio)
Token/Type: 0.5034

**Hypertension  
Eng**

### TABELA DE FREQUÊNCIA

Pos.	Palavra	Freq.
1	the	22102
2	of	16798
3	and	13302
4	in	12932
5	to	7968
6	a	6243
7	with	5363
8	for	4189
9	was	4135



TABELA DE FREQUÊNCIA

Pos.	Palavra	Freq.
1	the	22102
2	of	16798
3	and	13302
4	in	12932
5	to	7968
6	a	6243
7	with	5363
8	for	4189
9	was	4135
10	were	3828
11	that	3293
12	is	3051
13	by	2982
14	as	2364
15	at	2347
16	or	2269
17	on	2142
18	from	1983
19	patients	1944
20	this	1760
21	be	1736
22	not	1707
23	pressure	1616
24	blood	1603
25	are	1518
26	an	1397
27	study	1328



Corpus Multilíngue para Ensino e Tradução

US

Trad - nosso novo corpus pa

Principal	19	patients	1944
O Projeto	20	this	1760
Equipe	21	be	1736
CorTec	22	not	1707
oMAprend	23	pressure	1616
CorTrad	24	blood	1603
rtigos, etc.	25	are	1518
Links	26	an	1397
Informativo	27	study	1328
Contato	28	hypertension	1270
Site FFLCH	29	bp	1222
Site USP	30	p	1221
	31	have	1171
	32	after	1164
	33	than	1121
	34	these	1086
	35	between	1056
	36	been	988
	37	no	937
	38	subjects	936
	39	may	881
	40	it	859
	41	has	845
	42	which	808
	43	hypertensive	807
	44	also	802
	45	we	800
	46	data	795
	47		770



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nformativo
- Contato
- ite FFLCH
- Site USP

35	between	1056
36	been	988
37	no	937
38	subjects	936
39	may	881
40	it	859
41	has	845
42	which	808
43	hypertensive	807
44	also	802
45	we	800
46	data	795
47	group	778
48	plasma	777
49	rats	746
50	levels	740
51	age	732
52	during	725
53	mm	702
54	ang ii	685
55	treatment	682
56	more	677
57	other	671
58	but	666
59	all	663
60	effect	655
61	increased	655
62	studies	651
63	effects	646

# Comparando as frequências

## Inglês:

- 19 – patients
- 23 – pressure
- 24 – blood
- 27 – study
- 28 – hipertension
- 29 – bp
- 30 – p
- 38 – subjects
- 43 – hypertensive
- 46 – data
- 47 – group
- 48 – plasma
- 49 – rats
- 50 - levels

## Português:

- 25 – pacientes
- 35 – arterial
- 44 – *pressure*
- 45 – *blood*
- 49 – hipertensão
- 50 – estudo
- 51 – tratamento
- 52 – pressão
- 53 – níveis
- 54 - maior



### Passo 3/3: Aplicando a Ferramenta sobre o *Corpus*.

Configure a ferramenta usando as opções abaixo:

#### Concordanciador

Expressão ou Palavra:

Diferenciar maiúsculas e minúsculas:

Tamanho do contexto reduzido (caracteres):

[>> Aplicar Conco](#)

[Escolher outra Ferra](#)

29 ncreased intramural **blood** volú me is unlikel  
 30 ve coronary venous **blood** would probably drai  
 31 enhance delivery of **blood** during SRP. Increa  
 32 this arterIALIZED **blood** to give up oxygen d  
 33 occlusion. Systolic **blood pressure** and LV en  
 34 pressures, systolic **blood pressure**, cardiac  
 35 ous differences of **blood** density gradients i  
 36 nutrient myocardial **blood** flow during SRP, m  
 37 egiional myocardial **blood** flow in experimenta  
 38 cooling of arterial **blood** while it is being  
 39 nation of increased **blood** supply by SRP with  
 40 atory, we have used **blood** infusate temperatu  
 41 enous pressure and **blood** drainage 35,36 . Th  
 42 content of venous **blood** (redistribution pha  
 43 lex coronary artery **blood** flow measured with  
 44 lex coronary artery **blood** flow (Cx Flow) meas  
 45 or antegrade cold **blood** cardioplegia. Menas  
 46 limited collateral **blood** flow, since therape  
 47 rdium with arterial **blood** during such interv  
 48 provide myocardial **blood** supply, hoping that  
 49 llimprove if enough **blood** reaches the acutely  
 50 aintain an adequate **blood pressure** and perip  
 51 e direct myocardial **blood** supply. Ho wever,  
 52 designed to supply **blood** flow to ischemic my  
 53 t rate or systolic **blood pressure** between th  
 54 thout the need for **blood** transfusion. Trans  
 55 heartrate, arterial **blood pressure**, right he  
 56 onal veins draining **blood** from the three ma  
 57 of cooled arterial **blood**, which can rapidly  
 58 tion in thearterial **blood pressure** after a 26  
 59 re, the lowering of **blood pressure** can occur  
 60 cal conversions of **blood** volume and velocity  
 61 ase in circulating **blood** volume. implies a d  
 62 t rate and systemic **blood pressure** are gener  
 63 al fall in systemic **blood** pres sure are like  
 64 by an in crease in **blood** flow 16 . Arterial  
 65 Increased systemic **blood** flow secondary to a

# Comparando as frequências

- **Português:**

- 25 – pacientes
- 35 – arterial
- 44 – *pressure*
- 45 – *blood*
- 49 – hipertensão
- 50 – estudo
- 51 – tratamento (54 – treatment)
- 52 – pressão
- 53 – níveis
- 54 - maior

- **Inglês:**

- 19 – patients
- 23 – pressure
- 24 – blood
- 27 – study
- 28 – hypertension
- 29 – bp
- 30 – p
- 38 – **subjects**
- 43 – hypertensive
- 46 – data
- 47 – group
- 48 – plasma
- 49 – rats
- 50 - levels

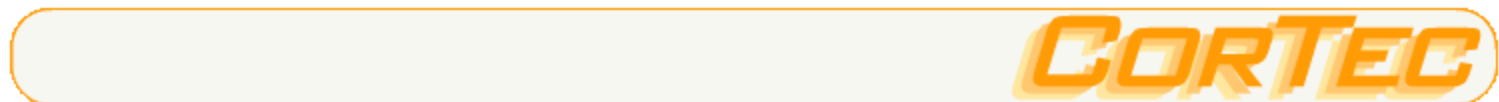


Corpus Multilíngue para Ensino e Tradução

USP

Bem-vindo ao site do Projeto Corpus

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP



Busca O que é

### Passo 3/3: Aplicando a Ferramenta sobre o *Corpus*.

Configure a ferramenta usando as opções abaixo:

#### Concordanciador

Expressão ou Palavra: Igual a

Diferenciar maiúsculas e minúsculas:

Tamanho do contexto reduzido (caracteres):

>> Aplicar Concordanciador

Escolher outra Ferramenta <<

CorTec - *Corpus* Técnico-Científico

Login



Principal

O Projeto

Equipe

CorTec

oMAprend

CorTrad

rtigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Busca O que é

Foram encontradas **976** ocorrências!! Fazer **DOWNLOAD** do resultado.

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 ction or serum enzymes among **subjects** submitted to retrograde as o  
2 Material and Methods: Twenty **subjects** divided in two groups, accor  
3 ts: In obese and hypertensive **subjects** not on medication we observe  
4 patient, there were 20 study **subjects** together with 10 individuals  
5 recorded by a printer in all **subjects**. In some representative indi  
6 ct, tissue perfusion of some **subjects** bearing unchanged CO through  
7 ion. Data obtained on normal **subjects** were used as a reference to  
8 ). In the obese hypertensive **subjects** not receiving medications (G  
9 universal improvement seen in **subjects** of G II, can be, at least pa  
10 ng-term study, lean diabetic **subjects** incorporated 14-26 g of guar  
11 lin requirements in diabetic **subjects**. Anderson and other investig  
12 Leitzmann 62 while studying **subjects** who ate oatmeal for bre akfa  
13 , and hyper cholesterolemic **subjects** 48 . Guar gum appears to sol  
14 11% in hypercholesterolem.c **subjects** taking 25 g of soy bean poly  
15 in weight managementof obese **subjects** by providing a greater volum  
16 in re requirements in diabetic **subjects**. Finally, dietary fiber also  
17 during angioplasty in human **subjects** has been controversial. Redu  
18 s feasible and safe in human **subjects** 67 . Although the protocol us  
19 increased several-fold over **subjects** with a negative result, the a  
20 G monitoring in asymptomatic **subjects** as for conventional exercise  
21 ing method for CAD in healthy **subjects** but its limitations in this

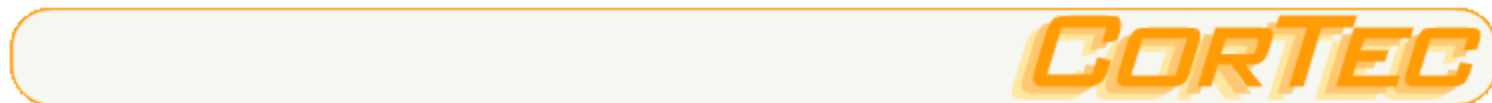


Corpus Multilíngue para Ensino e Tradução

US

publicações no link Artigos

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nformativo
- Contato
- Site FFLCH
- Site USP



Engl

Busca O que é

Foram encontradas **6** ocorrências!! Fazer **DOWNLOAD** do resultado.

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 ivíduos idosos não estão mais **sujeitos** a apresentar manifestações a  
 2 cúrio, mais precisos e menos **sujeitos** à descalibração 8 . Utilizara  
 3 a que os seres humanos estão **sujeitos**. Na prática clínica, e mesmo  
 4 índices de contratilidade não **sujeitos** a influências das condições d  
 5 dos a exercício, estarão mais **sujeitos** a apresentar as alterações el  
 6 ais não ficam artificialmente **sujeitos** a riscos; 2) pacientes podem

>> Fazer Nova Busca

CorTec - Corpus Técnico-Científico

# Comparando as frequências

- **Português:**

- 25 – pacientes
- 35 – arterial
- 44 – *pressure*
- 45 – *blood*
- 49 – hipertensão
- 50 – estudo
- 51 – tratamento (54 – treatment)
- 52 – pressão
- 53 – níveis
- 54 - maior

- **Inglês:**

- 19 – **patients**
- 23 – pressure
- 24 – blood
- 27 – study
- 28 – hypertension
- 29 – bp
- 30 – p
- 38 – subjects
- 43 – hypertensive
- 46 – data
- 47 – group
- 48 – plasma
- 49 – rats
- 50 - levels



## CORTEC

Busca O que é

Foram encontradas 1668 ocorrências!! Fazer **DOWNLOAD** do resultado.

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 ue a hipertensão arterial de **pacientes** com idades en- tre 60 e 80 a  
2 diuréticos ou empregados em **pacientes** com reninemia eleva- da. Exe  
3 ernentes espe- cificamente a **pacientes** idosos ainda não estão dis-  
4 agonistas, **envolven-** do 1900 **pacientes** com ICC de duração de um a 24  
5 ces de função car- diaca nos **pacientes** com ICC, em grau maior do que  
6 CC, em grau maior do que nos **pacientes** que estavam sob terapia com  
7 Power e col 5 observaram que **pacientes** com ICC apresentavam melhor  
8 de grupo controle, por- que **pacientes** submetidos a teste ergométric  
9 rpotassemia (especialmente em **pacientes** com insuficiência renal, ou  
10 ais são observados em 3% dos **pacientes** submetidos à terapia com lis  
11 a. Material e Métodos: **Vinte** **pacientes** submeti- dos durante seis se  
12 E MÉTODOS **Foram avaliados 20** **pacientes**, 11 (55%) do sexo feminino,  
13 melito insulino dependente e **pacientes** em uso de anticoncepcional or  
14 a- nas de uso de placebo, os **pacientes** eram designados a tomar drog  
15 tratados com placebo. Um dos **pacientes** deste últi- mo grupo foi ret





Principal

O Projeto

Equipe

CorTec

oMAprend

CorTrad

rtigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Foram encontradas 2001 ocorrências!! Fazer **DOWNLOAD** do resultado.

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 therapeutic option for many **patients** with severe coronary artery d  
2 he procedure, particularly in **patients** with highrisk anatomy. One s  
3 able to evaluate and manage **patients** with documented myocardial i  
4 ainly bring renewed hope for **patients** with coronary artery disease  
5 increases effort tolerance in **patients** with exercise in duced ST se  
6 he exercise capacity also in **patients** with intermittent Fig. 1 | P  
7 idamoleEchocardiography test **Patients** are instructed to fast for at  
8 Chronic bronchopneumopathic **patients** needing therapy with xanthin  
9 wer panel). state. In these **patients**, dipyridamole infusion can be  
10 lography 18 . In all studied **patients**, the quality of echocardiogra  
11 Safety. About two thirds of **patients** experienced some minor side  
12 we completed the test in all **patients**. The symptoms quickly resolve  
13 hylline administration. A few **patients** required nitrates as well to  
14 test even in unstable angina **patients** in whom exercise stress test  
15 accuracy. In a subset of 710 **patients** with history of chest pain an  
16 nd 89%, respectively) and in **patients** with normal resting function  
17 dial ischemia in over 50% of **patients** with history of chest pain,  
18 c (exercise tolerance in CAD **patients**) as well as prognostic standa  
19 tivessel disea se 15,16 . In **patients** with single vessel disease, m  
20 ischemia producing vessel in **patients** with multivessel disease who  
21 ifferent coronary arteries in **patients** with normal resting function  
22 this pattern only in the few **patients** in whom the multiple asynerg  
23 e (heterozonal) positivity in **patients** with a resting dyssynergy: t  
24 rn is relatively frequent in **patients** with a previous myocardial i



Corpus Multilíngue para Ensino e Tradução

US

Seja bem vindo ao site do Pr

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

19	patients	1944
20	this	1760
21	be	1736
22	not	1707
23	pressure	1616
24	blood	1603
25	are	1518
26	an	1397
27	study	1328
28	hypertension	1270
29	bp	1222
30	p	1221
31	have	1171
32	after	1164
33	than	1121
34	these	1086
35	between	1056
36	been	988
37	no	937
38	subjects	936
39	may	881
40	it	859
41	has	845
42	which	808
43	hypertensive	807
44	also	802
45	we	800
46	data	795
47		770



Principal

O Projeto

Equipe

CorTec

oMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Foram encontradas 1300 ocorrências!! Fazer **DOWNLOAD** do resultado.

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 vs. moderate blood pressure (BP) reduction -- would either pr  
2 to moderate (target diastolic BP, 80-89 mm Hg) or intensive (t  
3 r intensive (target diastolic BP, 75 mm Hg) antihypertensive t  
4 tudy medication did not bring BP to goal, add-on therapy with  
5 th the intensive and moderate BP treatment groups (25 fatal/no  
6 g goal systolic and diastolic BP. At randomization, BP was <14  
7 astolic BP. At randomization, BP was <140/90 mm Hg in only 20.  
8 sive medications whenever the BP was  $\geq 140/90$  mm Hg, 84.8% of  
9 he subjects attained the goal BP. During the 2 years of treatm  
10 ing the 2 years of treatment, BP control was maintained in 70%  
11 f the subjects for a systolic BP <140 mm Hg and in 90% of the  
12 the subjects for a diastolic BP <90 mm Hg.[15] These data, li  
13 If patients did not meet the BP goal with the maximum tolerat  
14 strated to effectively reduce BP and to do so comparably to th  
15 nd to do so comparably to the BP reduction observed with hydro  
16 beyond its ability to reduce BP, doxazosin would confer addit  
17 oxazosin treatment arm in the BP component of the trial. It wa  
18 p should be informed of their BP treatment assignment and that  
19 of continuing the rest of the BP and lipid-lowering trial.  
20 as necessary to achieve goal BP. In this regard, 4 years into  
21 routinely accepted equivalent BP reduction as a suitable surro  
22 become available. Equivalent BP reduction by different antihy  
23 such as cigarette smoking, a BP greater than 140/90 mm Hg, or  
24 and reducing blood pressure (BP) have been documented in nume  
25 s two key enzymes involved in BP homeostasis: neutral endopept



Corpus Multilíngue para Ensino e Tradução

US

as novas publicações no link

- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nformativo
- Contato
- Site FFLCH
- Site USP

Clique na palavra de busca para obter um contexto expandido com 150 caracteres

1 Fisiologia Clinical CNR Via P Savi, 8156100 Pisa, Italy. M  
2 is in logarithmic scale. (\*  $p < .05$  vs intravenous at endo r  
3 ravenous at endo regions, and  $p < .01$  at mid and epi regions  
4 TTAR, CARLOS E. SALAS, DONALD P. BERNSTEIN, DAVID LEHR, ROBER  
5 TTAR, CARLOS E. SALAS, DONALD P. BERNSTEIN, DAVID LEHR, ROBER  
6  $4 \pm 357$  dynes. sec.  $\text{cm}^{-5}/\text{m}^2$ ;  $p < 0.05$  (Data obtained in supi  
7 test for matched pairs, with  $p < 0.05$  considered significant  
8 G II (3.27 to 3.58  $\text{L}/\text{min}/\text{m}^2$ ;  $p < 0.05$ ). Despite similar bod  
9 12 week experimental period ( $p < 0.05$ ) 15 . In a 3 month pl  
10 ssigned to the control group ( $p = 0.06$ ) 21 . By statistical  
11 phosphorylation. Is opentenyl p yroph osph ate c an isomeriz  
12 ytic Drugs> P. Maggioni, Claudio Fresco Mil  
13 of Thrombolytic Drugs Aldo P. Maggioni, Claudio Fresco  
14 Milano Correspondência: Aldo P. Maggioni Istituto di Ricerc  
15 intervals from 1.09 to 1.84 ( $p < 0.008$ ). In addition, subcutan  
16 significant for altepla-se ( $p < 0.01$ ) but not for APSAC ( $p < 0.$   
17 ( $p < 0.01$ ) but not for APSAC ( $p < 0.08$ ). When the strokes were  
18 APSAC in hemorrhagic strokes ( $p < 0.00001$ ). The GUSTO trial, w  
19 and 0.49% and 0.54% with SK ( $p < 0.03$ ). The calculated excess  
20 reated patients (0.9% vs 0.7%  $p < 0.05$ ). Also in the GUSTO tri  
21 ith alteplase (11.8% vs 7.1%  $p < 0.0001$ ). In GISSI-2 5 the inc  
22 alteplase and 4.4% with SK ( $p < 0.001$ ). In GUSTO 9 sustained  
23 iac arrest, or reinfarction (2 $p < 0.0002$ ). Although this endpo  
24 initiation of therapy (25%, 2 $p < 0.003$ ) most likely due to a r  
25 uces mortality by about 13% ( $p < 0.02$ ; 95% CI, -2% to -25%), n  
26 n-fatal reinfarction by 19% ( $p < 0.01$ ; 95% CI, -5% to -33%) an  
27 -fatal cardiac arrest by 19% ( $p < 0.02$ ; 95% CI, -2% to -30%) i

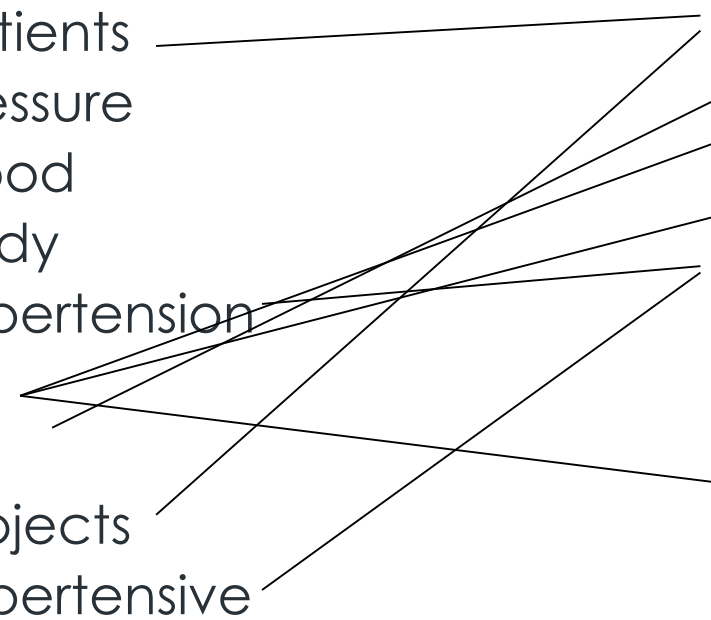
# Comparando as frequências

## ○ Inglês:

- 19 – patients
- 23 – pressure
- 24 – blood
- 27 – study
- 28 – hypertension
- 29 – bp
- 30 – p
- 38 – subjects
- 43 – hypertensive
- 46 – data
- 47 – group
- 48 – plasma
- 49 – rats
- 50 – levels

## ○ Português:

- 25 – pacientes
- 35 – arterial
- 44 – *pressure*
- 45 – *blood*
- 49 – hipertensão
- 50 – estudo
- 51 – tratamento (54 – treatment)
- 52 – pressão
- 53 – níveis
- 54 – maior



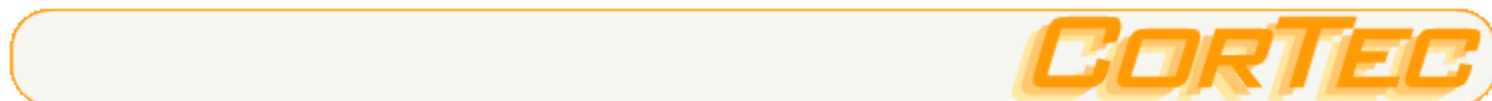
# Tarefa 2

- *Situação: Você precisa traduzir “medidores de vazão” para o inglês.*
- Use o corpus de “Medidores eletromagnéticos de vazão”
- Qual foi seu procedimento para estabelecer a equivalência?



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nfornativo
- Contato
- Site FFLCH
- Site USP

68 singularidades. Cada um dos **medidores** tem tomadas de pressão adequa  
69 tomação Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
70 tomação Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
71 tomação Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
72 tomação Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
73 ais parâmetros são medidos, **medidores** e transdutores utilizados  
74 das saídas digitais, esses **medidores** de nível e vazão possuem tamb  
75 tomação Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
76 o, é necessário recorrer à **medidores** de vazão do tipo eletromagnét  
77 tomação Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
78 egra TEXTO FONTE Título: **Medidores** magnéticos Subtítulo: n.c.  
79 ontrele Assunto específico: **Medidores** de Vazão DISTRIBUIÇÃO Tip  
80 Blücher Ltda. 1983. IV.2.4 **Medidores** magnéticos Os medidores magn  
81 .2.4 Medidores magnéticos Os **medidores** magnéticos foram desenvolvido  
82 básico O princípio básico dos **medidores** magnéticos é semelhante ao de  
83 rma que existem dois tipos de **medidores** magnéticos : de corrente cont  
84 de material termoplástico. Os **medidores** magnéticos podem ser executad  
85 ões de operações em que estes **medidores** têm aplicações, depende da re  
86 m 150°C. É possível encontrar **medidores** magnéticos construídos para s  
87 rodos (em m) ( na maioria dos **medidores** magnéticos) Velocidade do flu  
88 lação entre a relação e/B. Os **medidores** de vazão magnéticos de corren  
89 to secundário do medidor. Nos **medidores** magnéticos de corrente altern  
90 fluência da condutividade nos **medidores** de vazão deve ser entendida c  
91 Influência da viscosidade: os **medidores** são independentes da viscosid  
92 dutividade para a maioria dos **medidores** magnéticos. Com equipamentos  
93 o elétrica Cabos de sinal Nos **medidores** convencionais, o sinal do med  
94 bobinas de campo requer, nos **medidores** magnéticos convencionais, uma  
95 cantes para instalação destes **medidores**. Em geral, os seguintes crité  
96 r instalado na posição. 2) Os **medidores** de mais de 300 mm de diâmetro  
97 iformemente os flanges. 4) Os **medidores** magnéticos podem ser instalad  
98 tais ou inclinados. Quando os **medidores** são instalados na posição ver



Busca O que é

### Passo 3/3: Aplicando a Ferramenta sobre o *Corpus*.

Configure a ferramenta usando as opções abaixo:

#### Concordanciador

Expressão ou Palavra: Igual a

Diferenciar maiúsculas e minúsculas:

Tamanho do contexto reduzido (caracteres):

>> Aplicar Concordanciador

Escolher outra Ferramenta <<





- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nfornativo
- Contato
- Site FFLCH
- Site USP

1 ters and over 60 suppliers of **magnetic** flowmeters. These three techn  
2 rmed orifice plate, a damaged **magnetic** flowmeter liner and other pot  
3 Flood, product specialist, **magnetic** flowmeters, for Krohne, these  
4 ria to choose among Coriolis, **Magnetic**, Ultrasonic and Vortex Meters  
5 echnologies include Coriolis, **magnetic**, ultrasonic, and vortex flowm  
6 es to minimize meter wear. **Magnetic** Flowmeters Magnetic flowmete  
7 wear. Magnetic Flowmeters **Magnetic** flowmeters use Faraday's Law  
8 um when it passes through a **magnetic** field. This voltage is direct  
9 rtional to the density of the **magnetic** field, the length of the co  
10 magnitude of the voltage. **Magnetic** flowmeters use wire coils mou  
11 rough these coils generates a **magnetic** field inside the pipe section  
12 either side of the pipe. The **magnetic** flowmeter uses this voltage  
13 ltage to compute flowrate. **Magnetic** flowmeters are used to measur  
14 ious limitation on the use of **magnetic** flowmeters is that they onl  
15 In addition, the accuracy of **magnetic** flowmeters can be affected  
16 natural gas applications. And **magnetic** flowmeters do not work on h  
17 y, followed by ultrasonic and **magnetic** meters. Many users are now  
18 le vortex and multivariable **magnetic** flowmeters have also been dev  
19 Consumers' Guide to Full-Bore **Magnetic** Flowmeters Subtitulo: Guest  
20 Consumers' Guide to Full-Bore **Magnetic** Flowmeters Guest article by  
21 Spitzer PE and Walt Boyes **Magnetic** flowmeters utilize Faraday's  
22 where wires travel through a **magnetic** field and produce a voltage.  
23 eter can be moved through the **magnetic** field of a horseshoe magnet a  
24 ent to move higher. In **magnetic** flowmeters, a magnetic field  
25 In magnetic flowmeters, a **magnetic** field is generated and channe  
26 c to allow penetration of the **magnetic** field into the liquid. Locati  
27 ower necessary to deliver the **magnetic** field, as well as reduce the  
28 conductive liquid through the **magnetic** field will cause a voltage si  
29 will move (flow) through the **magnetic** field and generate a voltage



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- nfornativo
- Contato
- Site FFLCH
- Site USP

1 ters and over 60 suppliers of **magnetic** flowmeters. These three techn  
2 rmed orifice plate, a damaged **magnetic** flowmeter liner and other pot  
3 Flood, product specialist **magnetic** flowmeters for Krohne, these  
4 ria to choose among Coriolis, **magnetic**, ultrasonic and Vortex Meters  
5 echnologies include Coriolis, **magnetic**, ultrasonic, and vortex flowm  
6 es to minimize meter wear. **Magnetic** Flowmeters Magnetic flowmete  
7 wear. Magnetic Flowmeters **Magnetic** flowmeters use Faraday's Law  
8 um when it passes through a **magnetic** field. This voltage is direct  
9 rtional to the density of the **magnetic** field, the length of the co  
10 magnitude of the voltage. **Magnetic** flowmeters use wire coils mou  
11 rough these coils generates a **magnetic** field inside the pipe section  
12 either side of the pipe. The **magnetic** flowmeter uses this voltage  
13 ltage to compute flowrate. **Magnetic** flowmeters are used to measur  
14 ious limitation on the use of **magnetic** flowmeters is that they onl  
15 In addition, the accuracy of **magnetic** flowmeters can be affected  
16 natural gas applications. And **magnetic** flowmeters do not work on h  
17 y, followed by ultrasonic and **magnetic** meters. Many users are now  
18 le vortex and multivariable **magnetic** flowmeters have also been dev  
19 Consumers' Guide to Full-Bore **Magnetic** Flowmeters Subtitulo: Guest  
20 Consumers' Guide to Full-Bore **Magnetic** Flowmeters Guest article by  
21 Spitzer PE and Walt Boyes **Magnetic** flowmeters utilize Faraday's  
22 where wires travel through a **magnetic** field and produce a voltage.  
23 eter can be moved through the **magnetic** field of a horseshoe magnet a  
24 ent to move higher. In **magnetic** flowmeters, a magnetic field  
25 In magnetic flowmeters, a **magnetic** field is generated and channe  
26 c to allow penetration of the **magnetic** field into the liquid. Locati  
27 ower necessary to deliver the **magnetic** field, as well as reduce the  
28 conductive liquid through the **magnetic** field will cause a voltage si  
29 will move (flow) through the **magnetic** field and generate a voltage

# Tarefa 3

- *Situação: Alguém pediu para você traduzir uma receita do inglês para o português.*
- Veja quão frequentes são os advérbios *gently*, *lightly* e *finely* no corpus de Culinária.
- Como você os traduziria?
- Confirme a tradução no CorTec.

# Tarefa 3

- *Gently* – 648 ocorrências
- *Lightly* – 445 ocorrências
- *Finely* – 538 ocorrências
  
- Tradução:
- Gentilmente?
- Levemente?
- Finamente?

# Tarefa 3

- ...to a floured board and knead gently until smooth.
- Turn onto a lightly floured surface.
- Mix lightly with a wooden spoon or fork
- ...a floured surface and knead lightly.
- Walnuts - 50g (2 oz), finely chopped

# Tarefa 3

- Acrescente a gelatina e misture delicadamente.
- Misture levemente a farinha com uma...
- Quando a cebola estiver ligeiramente dourada
- e (levemente) polvilhada com farinha
- 110 g de presunto cozido bem picado (55 ocor.)
- 2 cebolas (médias) picadinhas

# CorTrad – projeto conjunto

- **Projeto CoMET** – DLM/USP, São Paulo (<http://www.fflch.usp.br/dlm/comet/>)
  - ✓ idealização
  - ✓ coleta e preparo dos textos
  - ✓ disponibilização do corpus para a comunidade
- **Linguateca** – SINTEF, Oslo (<http://www.linguateca.pt/>)
  - ✓ idealização
  - ✓ desenvolvimento e implementação da parte computacional - DISPARA
  - ✓ alinhamento e anotação iniciais
- **NILC** – USP, São Carlos (<http://www.nilc.icmc.usp.br/>)
  - ✓ hospedagem
  - ✓ futuramente: manutenção e melhoria da parte computacional

# CorTrad

[http://www.fflch.usp.br/dlm/comet/consulta\\_cortrad.html](http://www.fflch.usp.br/dlm/comet/consulta_cortrad.html)

- Projeto **em andamento**, iniciado em maio de 2008
- Conta atualmente com 3 subcorpora:
  - ✓ **Literário** (por ora, 27 contos australianos e 28 canadenses)
  - ✓ **Técnico-científico** (por ora, 1 livro de culinária de 350 páginas – aprox. 130.000 palavras)
  - ✓ **Jornalístico** (por ora, 1.076 textos provenientes de vários números da Revista Pesquisa Fapesp (2001 a 2003))
- Critério de coleta: **disponibilidade**
- **Diferenciais:**
  - ✓ **Multiversão** – permite comparar várias etapas da tradução / revisão
  - ✓ **Sistema de busca refinado** – permite buscas específicas para cada subcorpus



# Sistema de busca e anotação

- **DISPARA** (Santos, 2002) - sistema parametrizável de disponibilização de corpora paralelos na rede
  - ✓ **Sistema de processamento de corpus**
    - IMS-CWB (Christ *et al.*, 1999), agora **Open CWB**
  - ✓ **Etiquetagem morfossintática**
    - Português: **PALAVRAS** (Bick, 2000)  
<http://visl.hum.sdu.dk/visl/pt/>
    - Inglês: **CLAWS** (Rayson & Garside, 1998)  
<http://www.comp.lancs.ac.uk/computing/research/ucrel/claws/>
  - ✓ **Interface**
    - desenhada pela equipe e programada por Patricia Tagnin



Corpus Multilíngue para Ensino e Tradução

ovas publicações no link

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

Agradecimentos

Projeto conjunto

Como citar

Perguntas já respondidas

Fale conosco

## Bem-vindo ao CorTrad

O CorTrad é o corpus multilíngue para ensino e tradução (português-inglês) do COMET. Além das possibilidades de pesquisa normal, o CorTrad dispõe de pelo menos duas funcionalidades interessantes: (i) possibilidade de se comparar diferentes versões de um mesmo texto (original, versões corrigidas, etc.); (ii) mecanismos de busca diferenciados para cada gênero pesquisado - permitindo, por exemplo, a busca e a visualização de seções específicas dos diferentes tipos textuais.

O CorTrad é um corpus aberto e conta atualmente com três subcorpora:

- CorTrad **jornalístico** (por ora, divulgação científica)
- CorTrad **literário** (por ora, contos)
- CorTrad **técnico-científico** (por ora, culinária)

A lista das obras incluídas, assim como os agradecimentos devidos, encontram-se em [Agradecimentos](#).

A disponibilização do CorTrad na rede é um [projeto conjunto COMET/NILC/Linguateca](#), usando o sistema [DISPARA](#).

Comentários ou questões para a equipe do CorTrad





Corpus Multilíngue para Ensino e Tradução



ao site do Projeto Comet

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Responsáveis: **Stella E. O. Tagnin** e **Elisa D. Teixeira**

Desenho/concepção gráfica do site do CorTrad: **Patricia Tagnin**

Coleta e preparação de textos, revisão de alinhamento: **Sabrina Matuda** e **Soraia Manzela**

**Linguateca**: aplicação do sistema DISPARA ao CorTrad; criação dos corpora em formato IMS (Open) CWB; desenvolvimento de novas funcionalidades para este projeto específico, em particular a procura em várias traduções e a adaptação da interface para textos técnicos; **documentação** do processo a seguir para aumento posterior do corpus.

Responsável: **Diana Santos**

**NILC**: alojamento do CorTrad em seu servidor; **manutenção e apeneioamento da parte computacional** do serviço a partir da data de uma primeira versão estável.

Responsável: **Sandra Aluisio**

Direitos autorais

De acordo com o art. 46, inciso VIII da **LEI Nº 9.610, DE 19 DE FEVEREIRO DE 1998**, não constitui ofensa aos direitos autorais: "a reprodução, em quaisquer obras, de pequenos trechos de obras preexistentes, de qualquer natureza, (...), sempre que a reprodução em si não seja o objetivo principal da obra nova e que não prejudique a exploração normal da obra reproduzida nem cause um prejuízo injustificado aos legítimos interesses dos autores".

Se você constatou a presença de textos no CorTrad que ferem os direitos autorais, por favor, entre em contato conosco e tomaremos as providências cabíveis.

Comentários ou questões para a equipe do CorTrad





Corpus Multilíngue para Ensino e Tradução

USP

Publicações no link Artigos

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## Perguntas já respondidas

Como faço para ver os adjetivos mais frequentes na parte culinária?

Procure por [pos="ADJ.\*"] e peça **Distribuição dos lemas**

Como faço para procurar advérbios no corpus de divulgação científica?

Procure [pos="ADV"]

O que significa principal?

*Principal* é o corpus a partir do qual será feita a pesquisa. Por exemplo: se você quer saber quantas ocorrências de "vermelho" há no texto traduzido finalmente publicado, deverá marcá-lo como principal. Isto significa que, se alguma ocorrência de "vermelho" da primeira versão da tradução foi eliminada na revisão, ela não aparecerá no resultado.

**A sua pergunta ainda não foi respondida?**

**Contate-nos que tentaremos responder muito em breve.**

# Corpus Jornalístico

- [Revista Pesquisa FAPESP](#)
- originais em português
- traduções em inglês – publicadas online
- edições de 2001, 2002 e 2003
- total de 20 números.

# Corpus jornalístico: Divulgação científica



**Original**  
(Português brasileiro)



**Tradução publicada**  
(*on-line*)





Corpus Multilíngue para Ensino e Tradução

omet!

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## Textos incluídos no corpus jornalístico

A **parte jornalística do CorTrad** conta atualmente com textos das edições de 2001, 2002 e 2003 da [Revista Pesquisa FAPESP](#), totalizando 20 números. As seções incluídas foram: Humanidades, Ciência, Tecnologia, Estratégias, Laboratório, Linha de Produção e Política de C&T.

Lista de revistas FAPESP incluídas:

- 2001: edições 62 (mar-01), 63 (apr-01), 64 (mai-01), 65 (jun-01), 67 (ago-01), 68 (set-01), 69 (out-01) e 70 (nov-01)
- 2002: edições 76 (jun-02), 77 (jul-02), 78 (ago-02), 79 (set-02), 80 (out-02), 81 (nov-02) e 82 (dez-02)
- 2003: edições 83 (jan-03), 84 (fev-03), 85 (mar-03), 86 (abr-03) e 87 (mai-03)

Atenção, esta página ainda é preliminar, em particular a classificação quanto ao gênero e assunto, segundo a [tabela](#) proposta pelo Lácio-Web, está sendo revista neste momento.

ID	Gênero	Assunto
ca-abr01v01	Carta	Direito
ca-abr01v02	Carta	Cultura
ca-abr01v03	Carta	Cultura

# Tarefa 4

*Situação: Você quer diversificar as formas de traduzir pesquisa para o inglês.*

- Use o corpus de divulgação científica (Revista Pesquisa FAPESP)





Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad jornalístico divulgação científica

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus

Original

principal pesquisa

ver

Tradução publicada

principal

ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

Concordância

Distribuição das formas

Distribuição dos lemas

Distribuição da categoria gramatical (PoS)



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Ignorar maiúsculas/minúsculas

Pesquisar

## Resultado

Concordância

Distribuição das formas

Distribuição dos lemas

Distribuição da categoria gramatical (PoS)

Distribuição do tempo verbal e/ou do caso pronominal

Distribuição de pessoa e/ou número

Distribuição do gênero morfológico

Distribuição da função sintática

Distribuição por documento

Distribuição por data de publicação

Distribuição por gênero de texto

Distribuição por tema

Distribuição por campo semântico

Distribuição por grupo (de cor, de vestuário, etc.)

## Opções

Resultados por ordem alfabética (só distribuições)

Resultados numerados (só concordância)

Pesquisar



Corpus Multilíngue para Ensino e Tradução

bem vindo ao site do Pr

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Pesquisar

## Exemplos de pesquisas

	Procurar:	Resultado:
a palavra <i>violino</i>	violino	concordância
palavras começando por <i>açúc</i>	"açúc.*"	concordância
palavras antecedidas por dois pontos	":" @ [ ]	distribuição das formas
formas do verbo <i>reunir</i> em contexto	reunir.*	concordância
palavras modificadas por <i>muito</i>	"muito" @ [ ]	distribuição de formas

Veja [dados quantitativos](#) para mais informações.

Comentários ou questões para a equipe do CorTrad



Última atualização da página: 06 de dezembro de 2009

Última atualização do corpus: 13 de outubro de 2009



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad jornalístico divulgação científica

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

[Pesquisar no corpus](#) ?



Original

principal pesquisa

ver

Tradução publicada

principal

ver

Ignorar maiúsculas/minúsculas

Pesquisar

Resultado

Concordância

Distribuição das formas

Distribuição dos lemas

Distribuição da categoria gramatical (PoS)



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

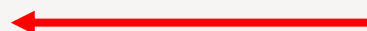
Literário

Técnico-científico

## Ajuda

Se tem dúvidas sobre a sintaxe da pesquisa, pode ver por ora as seguintes páginas feitas para projetos relacionados:

- [Exemplos de pesquisa no AC/DC](#)
- [Ajuda à pesquisa no COMPARA](#)



Em breve mais informação específica de ajuda ao próprio CorTrad será colocada aqui.

Comentários ou questões para a equipe do CorTrad





Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP



## Ajuda para pesquisar

▶ [Início](#)

▶ [Pesquisa](#)

▶ [Ajuda](#)

[Pesquisa](#)

[Aula prática](#)

▶ [Textos do COMPARA](#)

▶ [Informações gerais](#)

▶ [Documentação específica](#)

▶ [Linguateca](#)

### Tópicos de ajuda

[Imprimir](#)

- [1. Seleccione a direcção de pesquisa](#)
- [2. Preencha com o tipo de pesquisa pretendido](#)
- [2.1 Pesquisa por palavra ou expressão](#) ←
- [2.2 Mais opções de pesquisa](#)
- [3. Utilize apenas uma parte específica do corpus \(opcional\)](#)
- [3.1 Especificar variantes do português e do inglês](#)
- [3.2 Especificar datas de publicação](#)
- [3.3 Diferenciar originais de traduções](#)
- [3.4 Pesquisar apenas em textos específicos](#)
- [3.5 Pesquisar autores específicos](#)
- [4. Especifique outros tipos de resultados \(opcional\)](#)

Notícias [19 de Set de 2008]



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

O que quer pesquisar

Exemplo do que deve digitar

Possíveis resultados

*palavra isolada*

"assim"

assim

*sequência de mais de uma palavra*

"a" "pensar"

a pensar

*duas palavras intercaladas por uma palavra qualquer*

"por" ".\*" "tempo"

por um tempo, por muito tempo, por pouco tempo, etc.

*duas palavras intercaladas por zero, uma, duas ou três outras quaisquer*

"bebeu" []\* "vinho" within 3

bebeu vinho, bebeu um vinho, bebeu dois vinhos, bebeu muito vinho...

*palavra com grafias diferentes*

"a(c)?ção"

acção, ação

"lo[iu]ça"

loiça, louça

*palavras alternativas*

"(enfim|finalmente)"

enfim, finalmente

*palavras começadas com "des"*

"des.\*"

desrespeito, desfavorecer, desânimo, desta, destaque...

*palavras terminadas em "ção"*

".\*ção"

canção, lição, relação, Conceição,



Corpus Multilíngue para Ensino e Tradução

; novas publicações no lir

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad jornalístico divulgação científica

Expressão de busca: "pesquisa"

Resultado escolhido: concordância em contexto

Corpus pesquisado: originais

**Apresenta-se uma amostra aleatória de 500 das 1322 ocorrências encontradas.**



Voltar

Nova pesquisa

Se quiser mais de 500 resultados da sua pesquisa, contate-nos.

### Original

A iniciativa privada tem flexibilidade e agilidade necessárias para levar avante um projeto como esse, já que as instituições públicas não têm cultura de mercado, diz Camargo. O CAT é um dos dez Centros de Pesquisa, Inovação e Difusão (Cepids) qualificados pela FAPESP e a parceria com a indústria farmacêutica para o desenvolvimento do novo medicamento atende a um dos principais quesitos do programa que é o de

### Tradução publicada

says Dr. Camargo. The CAT is one of the ten Research Centers, Innovation and Diffusion (Cepids) certified by FAPESP and the partnership with the pharmaceutical industry for the development of the new medicine responds to one of the main requirements of the program which is that of putting together the academic activities of research with the market forces.





Corpus Multilíngue para Ensino e Tradução

Português / Inglês

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

### Original

A iniciativa privada tem flexibilidade e agilidade necessárias para levar avante um projeto como esse, já que as instituições públicas não têm cultura de mercado, diz Camargo. O CAT é um dos dez Centros de Pesquisa, Inovação e Difusão (Cepids) qualificados pela FAPESP e a parceria com a indústria farmacêutica para o desenvolvimento do novo medicamento atende a um dos principais quesitos do programa que é o de aproximar as atividades acadêmicas de **pesquisa** com o mercado.

Para um país cuja academia manteve corporativamente protegida da sociedade, a criação da revista deu visibilidade a linhas de **pesquisa** que não teriam sobrevivido ao massacre intramuros, avalia Neves.

Os exemplos citados pelo pesquisador são das mais importantes obras encontradas na **pesquisa** do século 16.

Em pauta, os dramáticos efeitos da conjuntura econômica do país, particularmente da crise cambial, sobre o fomento à **pesquisa** científica em São Paulo.

### Tradução publicada

says Dr. Camargo. The CAT is one of the ten Research Centers, Innovation and Diffusion (Cepids) certified by FAPESP and the partnership with the pharmaceutical industry for the development of the new medicine responds to one of the main requirements of the program which is that of putting together the academic activities of **research** with the market forces.

«For a country whose academic circles have remained as a body protected from society, the creation of the magazine brought visibility to the lines of **research** that would not have survived the intramural massacre», reckons Neves.

Other examples cited by the researcher are among the most important works found in the **research** of the 16th century.

The subject of discussion was the dramatic effects of the economic crisis in the country, particularly the **exchange** crisis, on the development of scientific **research** in São Paulo.

# CorTrad – Pesquisas

E quando ela **NÃO** é traduzida  
por “research”?

**!“research”**



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

## Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input type="text" value="pesquisa"/>	<input checked="" type="checkbox"/> ver
Tradução publicada	<input type="radio"/> principal	<input research\""="" type="text" value="!\"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- |  |   |
|--|---|
| <input checked="" type="radio"/> Concordância                              | <input type="radio"/> Distribuição das formas                             |
| <input type="radio"/> Distribuição dos lemas                               | <input type="radio"/> Distribuição da categoria gramatical (PoS)          |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número                  |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática                    |
| <input type="radio"/> Distribuição por documento                           | <input type="radio"/> Distribuição por data de publicação                 |
| <input type="radio"/> Distribuição por gênero de texto                     | <input type="radio"/> Distribuição por tema                               |
| <input type="radio"/> Distribuição por campo semântico                     | <input type="radio"/> Distribuição por grupo (do cor, do vestuário, etc.) |



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

barco a motor, refrigerador, balanças, estufa, câmara climática, pagamento dos catadores e um equipamento de análise de imagens por computador, usado na mensuração de mais de 3 mil caranguejos.

boat, a refrigerator, scales, an oven, an air conditioned chamber, payment of the gatherers, and an item of equipment for the computer analysis of images, used in measuring more than 3,000 crabs.

O fato é que a abrangência da **pesquisa** foi crescendo na medida em que o projeto avançava.

In fact the coverage of the **survey** increased as the project went on.

O guaxinim pode permanecer sentado, imóvel, em frente à abertura da toca, esperando o caranguejo sair, conta Gláucia. A **pesquisa** revelou, contudo, uma falta de conhecimento dos catadores sobre características importantes do uçá.

The raccoon can remain seated, motionless, in front of the entrance to the burrow, **waiting** for the crab to come out», says Gláucia. The **survey** showed, however, some gaps in the gatherers' knowledge of certain important characteristics of the uçá.

Grupos de **pesquisa** abrem perspectivas para prevenção e tratamento mais eficaz.

**Research** groups open up prospects for a more efficient method of prevention and treatment

Em relação aos animais domésticos incluídos na **pesquisa**, o teste com a técnica da reação de cadeia de polimerase (PCR) mostrou que cerca de 20 % dos 61 cães e 14 % dos 58 eqüinos testados apresentaram infecção por Leishmania.

In relation to the domestic animals included in the **study**, the test with the technique of Polymerase Chain Reaction (PCR) showed that around 20% of the 61 dogs and 14% of the 58 equines tested showed infection by Leishmania.

Na outra ponta, entre os comportamentos que podem reduzir a incidência desse problema cardiovascular, a **pesquisa** chegou a um resultado surpreendente: o consumo de álcool se mostrou a mais efetiva forma de proteção contra o ataque cardíaco.

At the other extreme, amongst the components that can reduce the **incidence** of this cardiovascular problem, the **study** arrived at a surprising result: the consumption of alcohol has proved to be the most effective form of protection against a heart attack.

# Tarefa 5

**Situação: Você não sabe se, numa entrevista, é melhor usar o verbo acreditar ou achar.**

- Usar o corpus jornalístico.
- [lema= “acreditar”] + Distribuição por gênero de texto
- [lema=“achar”] + Distribuição por gênero de texto



Corpus Multilíngue para Ensino e Tradução



Conheça o CorTrad - no

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAPrend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad jornalístico divulgação científica

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input acreditar"]"="" type="text" value="[lema="/>	<input checked="" type="checkbox"/> ver
Tradução publicada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

#### Resultado

- |  |   |
|--|---|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                             |
| <input type="radio"/> Distribuição dos lemas                               | <input type="radio"/> Distribuição da categoria gramatical (PoS)          |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número                  |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática                    |
| <input type="radio"/> Distribuição por documento                           | <input type="radio"/> Distribuição por data de publicação                 |
| <input checked="" type="radio"/> Distribuição por gênero de texto          | <input type="radio"/> Distribuição por tema                               |
| <input type="radio"/> Distribuição por campo semântico                     | <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.) |

Opcões



Início

Jornalístico

Literário

Técnico-científico

## CorTrad jornalístico divulgação científica

196 casos.

Expressão de busca: [lema="acreditar"]

Resultado escolhido: distribuição de genero

Corpus pesquisado: originais



Voltar

Nova pesquisa

Houve 11 valores diferentes de genero.

Reportagem	92
Notícia	40
article	23
Artigo	21
Entrevista	11
Notícia-OU-section	3
section	2
opinion	1
article-OU-section	1
notícia-OU-section	1
Notícia -OU-Notícia	1

Para informação sobre os códigos internos do atributo **genero**, consulte a página do subcorpus respectivo.

8-janeiro-2010:Esperamos que o CorTrad jornalístico divulgação científica lhe tenha sido útil!



## CorTrad jornalístico divulgação científica

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus ?

Original	<input checked="" type="radio"/> principal [lema="achar"]	<input checked="" type="checkbox"/> ver
Tradução publicada	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

- |  |   |
|--|---|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                             |
| <input type="radio"/> Distribuição dos lemas                               | <input type="radio"/> Distribuição da categoria gramatical (PoS)          |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número                  |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática                    |
| <input type="radio"/> Distribuição por documento                           | <input type="radio"/> Distribuição por data de publicação                 |
| <input checked="" type="radio"/> Distribuição por gênero de texto          | <input type="radio"/> Distribuição por tema                               |
| <input type="radio"/> Distribuição por campo semântico                     | <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.) |

#### Opções





Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad jornalístico divulgação científica

160 casos.

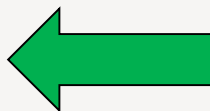
Expressão de busca: [lema="achar"]

Resultado escolhido: distribuição de genero

Corpus pesquisado: originais

Houve 11 valores diferentes de genero.

Entrevista	57
Reportagem	50
article	24
Notícia	14
Artigo	8
Carta	2
-OU-Editorial	1
editorial	1
Editorial	1
Notícia-OU-section	1
section	1



Para informação sobre os códigos internos do atributo **genero**, consulte a página do subcorpus respectivo.

8-janeiro-2010:Esperamos que o CorTrad jornalístico divulgação científica lhe tenha sido útil!

Vo  
Nova pesq

# Corpus técnico-científico: Culinária



**Original**  
(português  
brasileiro)



**Primeira  
versão da  
tradução**  
(inglês)



**Texto  
revisado**  
(por falante  
nativo)



**Tradução  
publicada**  
(ainda não  
digitalizada)





Corpus Multilíngue para Ensino e Tradução

USP

ções no link Artigos

Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## Obras no CorTrad técnico-científico

### Lista de obras incluídas no subcorpus de culinária

A **parte culinária do CorTrad** conta atualmente com o conteúdo integral de um livro de 350 páginas em quatro versões: o texto original, escrito em português brasileiro, a tradução para o inglês, feita por duas tradutoras brasileiras, a versão revisada dessa tradução, feita por uma falante nativa do inglês, e a versão final, ainda em preparação.

- Bacellar, Heloísa. *Cozinhando para amigos*, São Paulo: DBA, 2005. / *Cooking for friends*. Traduzido por Elisa Duarte Teixeira e Alvamar H. de Campos Andrade Lamparelli; revisado por Virgínia Klie. DBA, 2008.

Comentários ou questões para a equipe do CorTrad



Última atualização da página: 06 de novembro de 2009

Última atualização do corpus: 26 de outubro de 2009

## Tarefa 6

Quais adjetivos  
ocorrem  
no corpus de culinária?

**[pos="ADJ"]**



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

## Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input adj.*"]"="" type="text" value="[pos="/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- |  |  |
|--|--|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                    |
| <input checked="" type="radio"/> Distribuição dos lemas                    | <input type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número         |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática           |
| <input type="radio"/> Distribuição por parte de obra                       | <input type="radio"/> Distribuição por tipo de texto             |
| <input type="radio"/> Distribuição por receita                             | <input type="radio"/> Distribuição por campo semântico           |



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

## CorTrad técnico-científico culinária

8024 casos.

Expressão de busca: [pos="ADJ.\*"]

Resultado escolhido: **distribuição de lema**

Corpus pesquisado: **originais**



Voltar

Nova pesquisa

**Houve 761 valores diferentes de lema.**

grande	447
médio	339
leve	263
fresco	168
macio	162
restante	156
branco	138
bom	131
fino	130
pequeno	117

## Tarefa 7

Quais são os advérbios em -ly que ocorrem na tradução?

```
[pos="RR.*" & word=".*ly"]
```



Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus

Original	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input checked="" type="radio"/> principal	<input &amp;="" ]"="" rr.*"="" type="text" value="[pos=" word="*ly"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- |  |  |
|--|--|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                    |
| <input checked="" type="radio"/> Distribuição dos lemas                    | <input type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número         |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática           |
| <input type="radio"/> Distribuição por parte de obra                       | <input type="radio"/> Distribuição por tipo de texto             |
| <input type="radio"/> Distribuição por receita                             | <input type="radio"/> Distribuição por campo semântico           |





Corpus Multilíngue para Ensino e Tradução



Seja bem vindo ao site do

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad técnico-científico culinária

1598 casos.

Expressão de busca: [pos="RR.\*" & word=".\*ly"]

Resultado escolhido: distribuição de lema

Corpus pesquisado: tradução revisada



[Voltar](#)

[Nova pesquisa](#)

Houve 142 valores diferentes de lema.

finely	176
approximately	99
only	97
lightly	83
slightly	67
thoroughly	58
thinly	56
fully	55
usually	52
really	50
coarsely	49
moderately	49
constantly	46
simply	36
evenly	32
cently	22

# Tarefa 8

*Situação: Você precisa traduzir uma receita para o inglês.*

- Descubra como é traduzido “suco de laranja natural”.
- Há outras colocações com “natural” em português?
- Como são traduzidas?

# Casos em que “natural” não é vertido como “natural”



Principal

O Projeto

Equipe

CorTec

CoMAprend

CorTrad

Artigos, etc.

Links

Informativo

Contato

Site FFLCH

Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

Contos

Lista de textos

### Pesquisar no corpus

Ajuda

Original	<input checked="" type="radio"/> principal	<input type="text" value="natural"/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text" value="lnatural"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text" value="lnatural"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

- Concordância
- Distribuição dos lemas
- Distribuição do tempo verbal e/ou do caso pronominal
- Distribuição do gênero morfológico
- Distribuição por parte de obra
- Distribuição por receita
- Distribuição por grupo (de cor, de vestuário, etc.)
- Distribuição das formas
- Distribuição da categoria gramatical (PoS)
- Distribuição de pessoa e/ou número
- Distribuição da função sintáctica
- Distribuição por tipo de texto
- Distribuição por campo semântico

Opções

# Tarefa 8

1/2 xícara de iogurte **natural**

cup **whole milk plain** yogurt

cup **plain whole milk** yogurt

1 litro de suco de laranja **natural**

1 L (1.1 qt) freshly squeezed orange juice

1 L (1 qt) freshly squeezed orange juice

Quando estive no Egito, vi que era comum amaciar tâmaras mais velhas e secas deixando-as de molho no leite, e daí surgiu a idéia do smoothie, um iogurte **natural** batido com tâmara e laranja, bem aveludado, de sabor misterioso.

When I was in Egypt I noticed it was quite common softening older and drier dates by soaking them in milk, and so the idea of a smoothie was brought about: a natural yogurt combined in the blender with dates and orange, pretty velvety, with a mysterious taste.

When I was in Egypt I noticed it was quite common to soften older and drier dates by soaking them in milk, and so the idea of a smoothie came about: a natural yogurt mixed in the blender with dates and orange, velvety, with a mysterious taste.

# Tarefa 8

1/2 xícara de iogurte **natural**

1 litro de suco de laranja **natural**

Quando estive no Egito, vi que era comum amaciar tâmaras mais velhas e secas deixando-as de molho no leite, e daí surgiu a idéia do smoothie, um iogurte **natural** batido com tâmara e laranja, bem aveludado, de sabor misterioso.

cup **whole milk plain** yogurt

1 L (1.1 qt) freshly squeezed orange juice

When I was in Egypt I noticed it was quite common softening older and drier dates by soaking them in milk, and so the idea of smoothie was brought about: a natural yogurt combined in the blender with dates and orange, pretty velvety, with a mysterious taste.

cup **plain whole milk** yogurt

1 L (1 qt) **freshly squeezed orange juice**

When I was in Egypt I noticed it was quite common to soften older and drier dates by soaking them in milk, and so the idea of a smoothie came about: a **natural yogurt** mixed in the blender with dates and orange, velvety, with a mysterious taste.

# Tarefa 9: Colocações na culinária

1. Descubra as colocações com

- **finely**
- **lightly**

2. Descubra os equivalentes em português

# Colocações na culinária

1. Descubra as colocações com

- **finely**: CHOP, diced, minced, shredded
- **lightly**: VERB [imperative]

2. Descubra os equivalentes em português:

- **finely** → bem + diminutivo
- **lightly** → ligeiramente

# Corpus literário: Contos australianos

(\*corpus de aprendizes)



**Original**  
(inglês  
australiano)



**Tradução dos  
alunos**  
(português  
brasileiro)



**Tradução  
revisada**  
(incorporando  
sugestões do  
professor)



**Tradução  
publicada**





Comparando  
diversas versões  
de uma tradução



- Início
- Jornalístico
- Literário
- Técnico-científico

### CorTrad literário contos

Expressão de busca: "suit" %c  
 Resultado escolhido: concordância em contexto  
 Corpus pesquisado: originais (versão 4.2)

[Voltar](#)[Nova pesquisa](#)

12 ocorrências.

Original	Primeira tradução	Tradução revisada	Tradução publicada
She was by then a woman of forty-nine or fifth -- small, straight, business -- like, in a tailored suit and with her hair cut in a silver helmet.	Era, àquela altura, uma mulher de quarenta e nove ou cinquenta anos -- pequena, ereta, metódica, trajando um conjunto feito sob encomenda e com o cabelo cortado em forma de elmo prateado.	Era, àquela altura, uma mulher de quarenta e nove ou cinquenta anos -- pequena, ereta, estilo executivo, trajando um conjunto feito sob medida e com o cabelo cortado em forma de elmo prateado.	Àquela altura, ela era uma mulher de quarenta e nove ou cinquenta anos -- pequena, ereta, com ar profissional, trajando um conjunto feito sob medida e com o cabelo cortado em forma de elmo prateado.
In the glow of candlelight off bronze, at three-thirty in the afternoon, when the city outside lay sweltering in the glare of a blue-black thundercloud, a being I could no longer think of as the woman in the lift, with her sensible shoes and her well-cut navy suit, was seated cross-legged among cushions, eyes closed, head rolled back with all the throat exposed as for a knife stroke.	Sob a luz de um candelabro de bronze, às três e meia da tarde, quando a cidade lá fora transpirava sob a claridade de uma nuvem tempestuosa negra-azulada, um ser que eu não conseguia mais ver como a mulher do elevador, com seus sapatos confortáveis e conjunto azul-marinho bem talhado, estava sentado de pernas cruzadas entre almofadas, olhos fechados, cabeça jogada para trás, com todo o pescoço exposto como para receber uma facada.	Sob a luz de um candelabro de bronze, às três e meia da tarde, quando a cidade lá fora derretia sob a claridade de uma nuvem de tempestade negro-azulada, um ser que eu não conseguia mais ver como a mulher do elevador, com seus sapatos confortáveis e conjunto azul-marinho bem talhado, estava sentado de pernas cruzadas entre almofadas, olhos fechados, cabeça jogada para trás, com todo o pescoço exposto como para receber uma facada.	À luz de um candelabro de bronze, às três e meia da tarde, enquanto a cidade lá fora derretia sob a claridade de uma nuvem de chuva negro-azulada, um ser que eu não conseguia mais ver como a mulher do elevador, com seus sapatos sóbrios e conjunto azul-marinho bem talhado, estava sentado de pernas cruzadas entre almofadas, os olhos fechados, a cabeça jogada para trás, com todo o pescoço exposto como para receber uma facada.

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

# Tarefa 10

*Situação: No seu texto aparece “then” com muita frequência. Você quer evitar repetir “então” na sua tradução.*

- principal: then
- tradução publicada: !então



Início

Jornalístico

Literário

Técnico-científico

## CorTrad literário contos

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input type="text" value="then"/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução publicada	<input type="radio"/> principal	<input type="text" value="lentão"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

Concordância

Distribuição das formas

Distribuição por obra

Distribuição por autor

#### Opções

Resultados por ordem alfabética (só distribuições)

Resultados numerados (só concordância)



- Principal
- O Projeto
- Equipe
- CorTec
- OMApred
- CorTrad
- rtigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

▼ Início    ▼ Jornalístico    ▼ Literário    ▼ Técnico-científico

## CorTrad literário contos

Expressão de busca: "then" :CORTRAD\_LIT\_TRADE !"então"

Resultado escolhido: concordância em contexto

Corpus pesquisado: originais



Voltar

Nova pesquisa

71 ocorrências.

Original	Primeira tradução	Tradução revisada	Tradução publicada
She was by <b>then</b> a woman of forty-nine or fifty - small, straight, <b>business-like</b> , in a tailored suit and with her hair cut in a silver helmet.	Era, àquela altura, uma mulher de quarenta e nove ou <b>cinquenta</b> anos - pequena, ereta, <b>metódica</b> , trajando um conjunto <b>feito sob</b> encomenda e com o cabelo cortado em forma de elmo prateado.	Era, àquela altura, uma mulher de quarenta e nove ou <b>cinquenta</b> anos - pequena, ereta, <b>estilo executivo</b> , trajando um conjunto feito sob medida e com o cabelo cortado em forma de elmo prateado.	<b>Àquela altura</b> , ela era uma mulher de quarenta e nove ou cinquenta anos - pequena, ereta, <b>com ar profissional</b> , trajando um conjunto feito sob medida e com o cabelo cortado em forma de elmo prateado.
All the more alarming <b>then</b> , as I sat waiting on one of the cane-bottomed chairs in the corridor, while Ben Steinberg, Miss McIntyre's star pupil, played the Max Bruch, to hear the same voice oddly transmuted.	Ainda mais alarmante — quando eu estava sentado esperando em uma das cadeiras de assento de palha no corredor enquanto Ben Steinberg, o aluno modelo de Miss McIntyre, tocava Max Bruch — era escutar a mesma voz estranhamente alterada.	Ainda mais alarmante — quando eu estava sentado esperando em uma das cadeiras de assento de palha no corredor enquanto Ben Steinberg, o aluno modelo de Miss McIntyre, tocava Max Bruch — era escutar a mesma voz estranhamente alterada.	Mais alarmante ainda — um dia, em que eu esperava sentado numa das cadeiras de palha no corredor enquanto Ben Steinberg, o aluno modelo de Miss McIntyre, tocava Max Bruch — foi escutar a mesma voz estranhamente alterada.
We were alone for a little <b>then</b> .	SEM TRADUÇÃO	Depois disso, ficamos sozinhos por um tempo.	<b>Depois disso</b> , ficamos sozinhos por um tempo.
' Which one are you <b>then</b> ?	- Ora, ora, qual deles é você?	- Ora, ora, qual deles é você?	<b>- Ora, ora</b> , qual deles é você?
Just <b>then</b> my Aunt Ourania came into the room, to which arrived	Nesse momento, minha Tia Ourania entrou no cômodo, para	Nesse momento, minha Tia Ourania entrou no cômodo, para	<b>Nesse momento</b> , minha Tia Ourania entrou no cômodo, para

# Anotação semântica para



- **cor**

- **roupa**





- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

## Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input roupa.*"]"="" type="text" value="[sema="/>	<input checked="" type="checkbox"/> ver
Tradução publicada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

[Pesquisar](#)

### Resultado

- |   |   |
|---|---|
| <input type="radio"/> Concordância  | <input type="radio"/> Distribuição das formas                             |
| <input checked="" type="radio"/> Distribuição dos lemas  | <input type="radio"/> Distribuição da categoria gramatical (PoS)          |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal  | <input type="radio"/> Distribuição de pessoa e/ou número                  |
| <input type="radio"/> Distribuição do gênero morfológico  | <input type="radio"/> Distribuição da função sintática                    |
| <input type="radio"/> Distribuição por documento  | <input type="radio"/> Distribuição por data de publicação                 |
| <input type="radio"/> Distribuição por gênero de texto  | <input type="radio"/> Distribuição por tema                               |
| <input type="radio"/> Distribuição por campo semântico  | <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.) |

### Opções

- Resultados por ordem alfabética (só distribuições)
- Resultados numerados (só concordância)

[Pesquisar](#)



- Principal
- O Projeto
- Equipe
- CorTec
- oMAprend
- CorTrad
- rtigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad jornalístico **divulgação científica**

93 casos.

Expressão de busca: [sema="roupa.\*"]  
Resultado escolhido: **distribuição de lema**  
Corpus pesquisado: **originais**



[Voltar](#)

[Nova pesquisa](#)

**Houve 26 valores diferentes de lema.**

roupa	17
coroa	12
sapato	9
manta	7
malha	7
carteira	5
manto	5
calçado	4
camiseta	4
gabão	3
acessório	2
bota	2
meia	2
vestido	2
camisa	1
sujeito	1



# Tarefa 11

Agora faça a mesma pesquisa  
no corpus de culinária

[sema="roupa.\*"]

distribuição dos lemas



## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em atualizações para saber mais.

### Pesquisar no corpus ?

Original	<input checked="" type="radio"/> principal [sema="roupa.*"]	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- |  |  |
|--|--|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                    |
| <input checked="" type="radio"/> Distribuição dos lemas                    | <input type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número         |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática           |
| <input type="radio"/> Distribuição por parte de obra                       | <input type="radio"/> Distribuição por tipo de texto             |
| <input type="radio"/> Distribuição por receita                             | <input type="radio"/> Distribuição por campo semântico           |



Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico **culinária**

96 casos.

Expressão de busca: [sema="roupa.\*"]  
Resultado escolhido: **distribuição de lema**  
Corpus pesquisado: **originais**



Voltar

Nova pesquisa

Houve 12 valores diferentes de lema.

bata	68
saia	10
roupa	7
avental	2
chapéu	2
cinto	1
gravata	1
casaco	1
meia	1
lenço	1
boné	1
vestido	1

Para informação sobre os códigos internos do atributo **lema**, consulte por ora a página do projeto AC/DC sobre [anotação](#).

5-janeiro-2010: *Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!*



Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico **culinária**

96 casos.

Expressão de busca: [sema="roupa.\*"]  
Resultado escolhido: **distribuição de lema**  
Corpus pesquisado: **originais**



Voltar

Nova pesquisa

Houve 12 valores diferentes de lema.

bata 68  
saia 10  
roupa 7  
avental 2  
chapéu 2  
cinto 1  
gravata 1  
casaco 1  
meia 1  
lenço 1  
boné 1  
vestido 1

# Algo estranho nesta lista?

Para informação sobre os códigos internos do atributo **lema**, consulte por ora a página do projeto AC/DC sobre [anotação](#).

5-janeiro-2010: *Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!*



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

			with flour. In a large bowl, combine flour, ground cinnamon, baking soda, salt, grated carrots, golden raisins, and chopped walnuts; set aside.
	Na batedeira <b>bata</b> o açúcar, os ovos, o óleo e a baunilha até conseguir um creme bem fofo.	For the muffins, preheat the oven to 200C (400F / moderately hot / Gas 6), grease 20 medium size muffin / cupcake tins with butter and dust with flour. In a large bowl, combine flour, ground cinnamon, baking soda, salt, grated carrots, golden raisins, and chopped walnuts; set aside.	In an electric mixer, beat sugar, eggs, oil and vanilla extract until a very fluffy. Gently fold in dry ingredients, spoon into pans and bake for about 30 minutes, until muffins are risen, golden brown, and an inserted toothpick comes out clean.
	minutos, até que os bolinhos estejam crescidos, dourados e, enfando um palito no centro, <b>ele saia limpo</b> .	In a electric mixer, beat sugar, eggs, oil and vanilla extract until a very fluffy cream is obtained.	In an electric mixer, beat sugar, eggs, oil and vanilla extract until a very fluffy. Gently fold in dry ingredients, spoon into pans and bake for about 30 minutes, until muffins are risen, golden brown, and an inserted toothpick comes out clean.
	Deixe engrossar por uns 5 minutos, junte as ervas, o suco do limão, sal, pimenta e as nozes e <b>bata</b> no processador até obter uma pasta não muito lisa.	Let bubble for 5 minutes or so, mix in herbs, lime juice, salt, pepper and walnuts, place in a food processor and process until a well mixed, but not too smooth paste is obtained.	Let bubble for 5 minutes or so, mix in herbs, lime juice, salt, pepper and walnuts, place in a food processor and process until a well mixed, but not too smooth paste is formed.
	Junte o óleo em fio pela abertura da tampa e <b>bata</b> até conseguir um molho cremoso.	Transfer turnovers to a baking sheet, brush with egg yolk mixed with a dash of water and refrigerate for 15 minutes, while you preheat the oven to 180C (350F / moderate / Gas 4).	Moisten the edges slightly with egg white, fold in half to enclose the filling and pinch the edges together to seal.
	Adicione a salsinha, <b>bata</b> até esverdear, acerte o sal e a pimenta e leve à geladeira por até 24 horas.	Transfer turnovers to a baking sheet, brush with egg yolk mixed with a dash of water and refrigerate for 15 minutes, while you preheat the oven to 180C (350F / moderate / Gas 4).	Transfer turnovers to a baking sheet, brush with egg yolk mixed with just a little water and refrigerate for 15 minutes, while you preheat the oven to 180C (350F / moderate / Gas 4).
	Numa terceira tigela, com a batedeira <b>bata</b> a manteiga, o cream cheese, a baunilha e o açúcar até obter um creme fofo e junte os ovos um a um.	In a third bowl, using a electric mixer, combine butter, cream cheese, vanilla extract, and sugar until a fluffy cream is obtained, and then stir in eggs, one at a time.	In a third bowl, using an electric mixer, combine butter, cream cheese, vanilla extract, and sugar until fluffy, and then stir in eggs, one at a time.
	Em seguida, adicione o leite de coco e 1/3 da mistura de farinha, <b>bata</b> até incorporar o	Next, add in coconut milk and 1/3 of the flour mixture, beat until thickened; repeat with	Next, add coconut milk and? of the flour mixture, beat until thickened; repeat with



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico

96 casos.

Expressão de busca: [sema  
Resultado escolhido: **distrib**  
Corpus pesquisado: **origina**

Houve 12 valores c

bata	68
saia	10
roupa	7
avental	2
chapéu	2
cinto	1
gravata	1
casaco	1
meia	1
lenço	1
boné	1
vestido	1

## CorTrad técnico-científico culinária

20 casos.

Expressão de busca: [sema="roupa.\*" %c]  
Resultado escolhido: **distribuição de lema**  
Corpus pesquisado: **originais (versão 2.3)**

Houve 11 valores diferentes de lema.

roupa	7
avental	2
vestir	2
chapéu	2
cinto	1
gravata	1
casaco	1
meia	1
lenço	1
boné	1
vestido	1



Voltar

Nova pesquisa

Para informação sobre os códigos internos do atributo **lema**, consulte por ora a página do projeto AC/DC sobre [anotação](#).

5-janeiro-2010: *Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!*



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

Original	<input checked="" type="radio"/> principal	<input cor.*"]"="" type="text" value="[sema="/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

### Resultado

- Concordância
- Distribuição dos lemas
- Distribuição do tempo verbal e/ou do caso pronominal
- Distribuição do gênero morfológico
- Distribuição por parte de obra
- Distribuição por receita
- Distribuição por grupo (de cor, de vestuário, etc.)
- Distribuição das formas
- Distribuição da categoria gramatical (PoS)
- Distribuição de pessoa e/ou número
- Distribuição da função sintática
- Distribuição por tipo de texto
- Distribuição por campo semântico

### Opções

- Resultados por ordem alfabética (só distribuições)
- Resultados numerados (só concordância)

Exemplos de pesquisas



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

▼ Início    ▼ Jornalístico    ▼ Literário    ▼ Técnico-científico  
Culinária  
Lista de textos

## CorTrad técnico-científico culinária

1139 casos.

Expressão de busca: [sema="cor.\*"]  
Resultado escolhido: distribuição de lema  
Corpus pesquisado: originais



Voltar

Houve 43 valores diferentes de lema.

Nova pesquisa

dourar	341
branco	144
vermelho	122
verde	118
louro	79
cru	45
preto	39
cor	39
laranja	32
dourado	27
amarelo	15
roxo	14
rosado	12
creme	12
castanho	10
castanho de peru	0





- Principal
- O Projeto
- Equipe
- CorTec
- CoMAPrend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad técnico-científico **culinária**

1139 casos.

Expressão de busca: [sema="cor.\*"]  
Resultado escolhido: **distribuição das formas**  
Corpus pesquisado: **originais**



[Voltar](#)

**Houve 87 valores diferentes de forma.**

[Nova pesquisa](#)

dourar	142
branco	92
verde	85
louro	79
doure	79
vermelhos	47
dourados	41
vermelhas	38
branca	37
verdes	33
laranja	30
douradas	30
vermelho	29
cru	28
dourado	28
dourada	28



- Principal
- Projeto
- Equipe
- CorTec
- OMApred
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Instituto FFLCH
- Site USP

## CorTrad jornalístico **divulgação científica**

650 casos.

Expressão de busca: [sema="cor.\*"]  
Resultado escolhido: **distribuição de lema**  
Corpus pesquisado: **originais**



Voltar

**Houve 83 valores diferentes de lema.**

Nova pesquisa

cor	87
negro	71
branco	61
vermelho	47
verde	45
amarelo	44
infravermelho	27
azul	25
ultravioleta	25
amarelinho	22
preto	21
transparente	15
roseo	9
colorido	8
corante	8
verdura	7
avermelhar	6



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus ?

Original	<input checked="" type="radio"/> principal [sema="cor.*"]	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- Concordância
- Distribuição dos lemas
- Distribuição do tempo verbal e/ou do caso pronominal
- Distribuição do gênero morfológico
- Distribuição por parte de obra
- Distribuição por receita
- Distribuição das formas
- Distribuição da categoria gramatical (PoS)
- Distribuição de pessoa e/ou número
- Distribuição da função sintática
- Distribuição por tipo de texto
- Distribuição por campo semântico



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAPrend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico culinária

1139 casos.

Expressão de busca: [sema="cor.\*"]  
Resultado escolhido: **distribuição de func**  
Corpus pesquisado: originais



Voltar

Nova pesquisa

Houve 40 valores diferentes de func.

N<	475
FMV	98
P<	92
>N	84
IMV_#ICL-P<	82
<SC	72
<ACC	42
IMV_#ICL-AUX<	31
IMV_#ICL-<SC	30
IMV_#ICL-<ACC	24
<PRED	19
NPHR	13
IMV	9
N<PRED	9
IMV_#ICL-N<	6
<SUBJ	5



Corpus Multilíngue para Ensino e Tradução

Conheça o CorTrad

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

## Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input "ambiente"="" type="text" value="meio"/>	<input checked="" type="checkbox"/> ver
Tradução publicada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

[Pesquisar](#)

### Resultado

- |  |   |
|--|---|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                             |
| <input type="radio"/> Distribuição dos lemas                               | <input type="radio"/> Distribuição da categoria gramatical (PoS)          |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número                  |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática                    |
| <input checked="" type="radio"/> Distribuição por documento                | <input type="radio"/> Distribuição por data de publicação                 |
| <input type="radio"/> Distribuição por gênero de texto                     | <input type="radio"/> Distribuição por tema                               |
| <input type="radio"/> Distribuição por campo semântico                     | <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.) |

### Opções

- Resultados por ordem alfabética (só distribuições)
- Resultados numerados (só concordância)

[Pesquisar](#)



## CorTrad jornalístico divulgação científica

38 casos.

Expressão de busca: "meio" "ambiente"  
Resultado escolhido: **distribuição de docid**  
Corpus pesquisado: **originais**



Voltar

Nova pesquisa

Houve 29 valores diferentes de **docid**.

su-set01v24	4
po-nov02v04	2
po-ago01v13	2
hu-mai01v09	2
ci-jul02v12	2
ci-ago01v08	2
ci-abr03v17	2
tec-mar03v04	1
tec-mai03v02	1
tec-mai01v10	1
tec-abr03v04	1
su-set01v32	1
su-set01v29	1
su-set01v27	1
su-set01v02	1
po-jul02v02	1
po-ago02v05	1
po-ago02v03	1
...	...



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

## Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input type="text" value='[sema="cor.*"]'/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

### Resultado

- |  |   |
|--|---|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                               |
| <input type="radio"/> Distribuição dos lemas                               | <input checked="" type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número                    |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática                      |
| <input type="radio"/> Distribuição por parte de obra                       | <input type="radio"/> Distribuição por tipo de texto                        |
| <input type="radio"/> Distribuição por receita                             | <input type="radio"/> Distribuição por campo semântico                      |
| <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.)  |   |

### Opções

- Resultados por ordem alfabética (só distribuições)
- Resultados numerados (só concordância)



Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico culinária

1139 casos.

Expressão de busca: [sema="cor.\*"]  
Resultado escolhido: **distribuição de pos**  
Corpus pesquisado: **originais**



Houve 6 valores diferentes de pos.

*pos = part-of-speech*

Voltar

Nova pesquisa

ADJ 619  
V 270  
N 139  
V\_fmc 93  
ADJ\_n 14  
V\_n 4

Para informação sobre os códigos internos do atributo **pos**, consulte por ora a página do projeto AC/DC sobre [anotação](#).

5-janeiro-2010: *Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!*

Comentários ou questões para a equipe do CorTrad







- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad técnico-científico culinária

1139 casos.

Expressão de busca: [sema="cor.\*"]  
 Resultado escolhido: **distribuição de pos**  
 Corpus pesquisado: **originais**



*pos = part-of-speech*

Voltar

Nova pesquisa

Houve 6 valores diferentes de pos.

<http://beta.visl.sdu.dk/visl/pt/info/symbolset-manual.html>

ADJ 619  
 V 270  
 N 139  
 V\_fmc 93  
 ADJ\_n 14  
 V\_n 4

V\_fmc = verb heading finite main clause

Para informação sobre os códigos internos do atributo **pos**, consulte por ora a página do projeto AC/DC sobre [anotação](#).

5-janeiro-2010: *Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!*

Comentários ou questões para a equipe do CorTrad





Corpus Multilíngue para Ensino e Tradução



Confira as novas publica

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus ?

Original	<input checked="" type="radio"/> principal	<input type="text" value='[sema="cor.*"]'/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

- |  |  |
|--|--|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                    |
| <input type="radio"/> Distribuição dos lemas   | <input type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal           | <input type="radio"/> Distribuição de pessoa e/ou número         |
| <input type="radio"/> Distribuição do gênero morfológico                             | <input type="radio"/> Distribuição da função sintática           |
| <input type="radio"/> Distribuição por parte de obra                                 | <input type="radio"/> Distribuição por tipo de texto             |
| <input type="radio"/> Distribuição por receita                                       | <input type="radio"/> Distribuição por campo semântico           |
| <input checked="" type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.) |  |

#### Opções



Corpus Multilíngue para Ensino e Tradução

USP

português / inglês

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAPrend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico culinária

1139 casos.

Expressão de busca: [sema="cor.\*"]

Resultado escolhido: distribuição de grupo

Corpus pesquisado: originais



Voltar

Nova pesquisa

Houve 17 valores diferentes de grupo.

Dourado	368
Branco	202
Verde	132
Vermelho	127
0	98
Nãoespecificada	47
Laranja	44
Preto	40
Amarelo	23
Castanho	19
Roxo	14
Creme	13
Outras_cor:caqui	6
Azul	2
Cinzentos	2
Rosa	1



Corpus Multilíngue para Ensino e Tradução



Seja bem

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus

Original	<input checked="" type="radio"/> principal	<input cor.*"]"="" type="text" value="[sema="/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

- |  |  |
|--|--|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                    |
| <input type="radio"/> Distribuição dos lemas                               | <input type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número         |
| <input checked="" type="radio"/> Distribuição do gênero morfológico        | <input type="radio"/> Distribuição da função sintática           |
| <input type="radio"/> Distribuição por parte de obra                       | <input type="radio"/> Distribuição por tipo de texto             |
| <input type="radio"/> Distribuição por receita                             | <input type="radio"/> Distribuição por campo semântico           |
| <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.)  |  |

#### Opcões



Corpus Multilíngue para Ensino e Tradução

USP

no link Artigos

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

Início

Jornalístico

Literário

Técnico-científico

## CorTrad técnico-científico culinária

1139 casos.

Expressão de busca: [sema="cor.\*"]  
Resultado escolhido: distribuição de gen  
Corpus pesquisado: originais



Voltar

Nova pesquisa

Houve 4 valores diferentes de género.

M 512  
F 352  
O 247  
M/F 28

5-janeiro-2010: Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!

Comentários ou questões para a equipe do CorTrad





- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus ?

Original	<input checked="" type="radio"/> principal	<input cor.*"]"="" type="text" value="[sema="/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- Concordância
- Distribuição dos lemas
- Distribuição do tempo verbal e/ou do caso pronominal
- Distribuição do gênero morfológico
- Distribuição por parte de obra
- Distribuição por receita
- Distribuição por grupo (de cor, de vestuário, etc.)
- Distribuição das formas
- Distribuição da categoria gramatical (PoS)
- Distribuição de pessoa e/ou número
- Distribuição da função sintática
- Distribuição por tipo de texto
- Distribuição por campo semântico

### Opções



Principal  
O Projeto  
Equipe  
CorTec  
CoMAprend  
CorTrad  
Artigos, etc.  
Links  
Informativo  
Contato  
Site FFLCH  
Site USP

## CorTrad técnico-científico culinária

991 casos.

Expressão de busca: [sema="cor.\*\*" %c]

Resultado escolhido: distribuição de idparte

Corpus pesquisado: originais (versão 2.3)

Houve 3 valores diferentes de idparte.

recheio 883

introd 64

concl 44

1-outubro-2014: Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!



Corpus Multilíngue para Ensino e Tradução



Seja bem vindo ao site

- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

## CorTrad tecnico-cientifico culinaria

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus

Original	<input checked="" type="radio"/> principal <input type="text" value='[sema="cor.*"]'/>	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal <input type="text"/>	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal <input type="text"/>	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

#### Resultado

- |  |  |
|--|--|
| <input type="radio"/> Concordância   | <input type="radio"/> Distribuição das formas                    |
| <input type="radio"/> Distribuição dos lemas                               | <input type="radio"/> Distribuição da categoria gramatical (PoS) |
| <input type="radio"/> Distribuição do tempo verbal e/ou do caso pronominal | <input type="radio"/> Distribuição de pessoa e/ou número         |
| <input type="radio"/> Distribuição do gênero morfológico                   | <input type="radio"/> Distribuição da função sintática           |
| <input type="radio"/> Distribuição por parte de obra                       | <input checked="" type="radio"/> Distribuição por tipo de texto  |
| <input type="radio"/> Distribuição por receita                             | <input type="radio"/> Distribuição por campo semântico           |
| <input type="radio"/> Distribuição por grupo (de cor, de vestuário, etc.)  |  |

#### Opções

Resultados por ordem alfabética (só distribuições)





## CorTrad técnico-científico culinária

991 casos.

Expressão de busca: [sema="cor.\*" %c]

Resultado escolhido: distribuição de tipotexto

Corpus pesquisado: originais (versão 2.3)

Houve 4 valores diferentes de tipo de texto.

Modo de fazer	430
Lista de ingredientes	272
Outro tipo	271
Título de receita	18

1-outubro-2014: Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!



- Principal
- O Projeto
- Equipe
- CorTec
- CoMAprend
- CorTrad
- Artigos, etc.
- Links
- Informativo
- Contato
- Site FFLCH
- Site USP

- Início
- Jornalístico
- Literário
- Técnico-científico

## CorTrad técnico-científico culinária

O CorTrad é um corpus aberto, sujeito a alterações. Clique em [atualizações](#) para saber mais.

### Pesquisar no corpus ?

Original	<input checked="" type="radio"/> principal [sema="cor.*"]	<input checked="" type="checkbox"/> ver
Primeira tradução	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver
Tradução revisada	<input type="radio"/> principal	<input checked="" type="checkbox"/> ver

Ignorar maiúsculas/minúsculas

Pesquisar

### Resultado

- Concordância
- Distribuição dos lemas
- Distribuição do tempo verbal e/ou do caso pronominal
- Distribuição do gênero morfológico
- Distribuição por parte de obra
- Distribuição por receita
- Distribuição das formas
- Distribuição da categoria gramatical (PoS)
- Distribuição de pessoa e/ou número
- Distribuição da função sintática
- Distribuição por tipo de texto
- Distribuição por campo semântico



## CorTrad técnico-científico culinária

991 casos.

Expressão de busca: [sema="cor.\*" %c]  
Resultado escolhido: **distribuição de sema**  
Corpus pesquisado: **originais** (versão 2.3)

Houve 4 valores diferentes de sema.

cor	583
cor:original	317
cor:vinho	87
cor_0	4

**Campo semântico**

1-outubro-2014: *Esperamos que o CorTrad técnico-científico culinária lhe tenha sido útil!*

# Relembrando...

- A LC é uma abordagem empírica:
  - **até pode, mas não é assim que se diz!**
- Estudos baseados em exemplos autênticos da **língua em uso** → corpus
- **Corpus** para a L.C.:
  - **textos autênticos**
  - **formato eletrônico**
  - **coletado segundo critérios específicos**
  - **representativo da língua ou variedade linguística em estudo**

# Corpora na Tradução

# Estratégias de tradução

- Como diversificar o vocabulário usando um corpus paralelo - COMPARA

O caso do “can”

# O caso de “can”

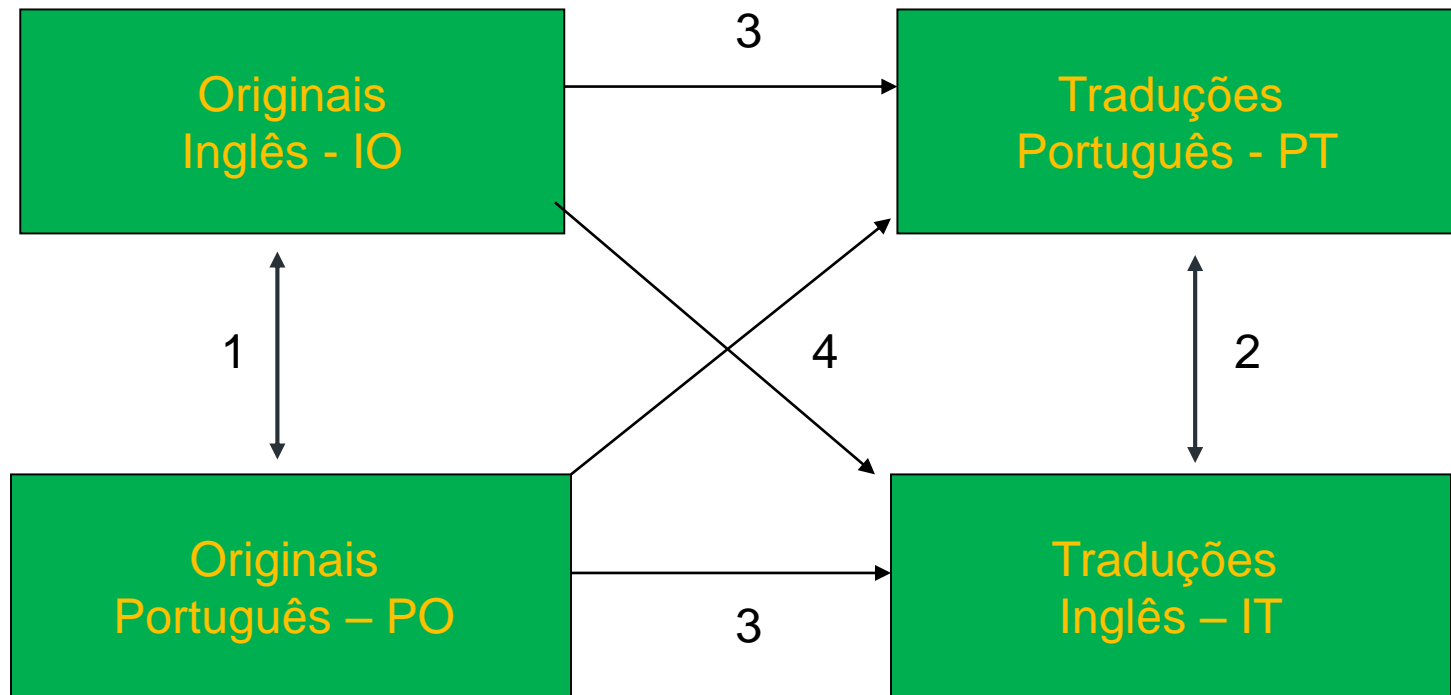
<a href="#">EBDL1T 1(19):</a>	I took a few paces forward without any ill-effects, shrugged, and put it down to some freakish twitch of a nerve, like the sudden excruciating crick you <b>can get</b> in your neck sometimes, twisting round to get something from the back seat of a car.	Dei uns passos e, sem sentir nada de especial, encolhi os ombros e atribuí a dor a uma veneta qualquer de um nervo, como aqueles estalos excruciantes que às vezes <b>sentimos</b> no pescoço quando nos voltamos para tirar qualquer coisa do banco de trás do carro.
<a href="#">EBDL1T 1(38):</a>	Admittedly, there are worse things that <b>can happen</b> to you, physically.	Reconheço que <b>podem acontecer</b> a uma pessoa coisas piores, pelo menos do ponto de vista físico.
<a href="#">EBDL1T 1(40):</a>	Not to mention the things you <b>can be born with</b> , like muscular dystrophy, cerebral palsy, haemophilia and epilepsy.	para não falar das coisas <b>congénitas</b> , como distrofia muscular, paralisia cerebral, hemofilia e epilepsia;
<a href="#">EBDL1T 1(43):</a>	Perhaps it's what they call «compassion fatigue», the idea that we get so much human suffering thrust in our faces every day from the media that we've become sort of numbed, we've used up all our reserves of pity, anger, outrage, and <b>can only think</b> of the pain in our own knee.	Se calhar, é aquilo a que costumam chamar «fadiga da compaixão», ou seja, a ideia de que os meios de comunicação nos atiram todos os dias à cara com tanto sofrimento humano que ficamos como que entorpecidos, esgotamos as nossas reservas de piedade, raiva, indignação, e <b>não conseguimos pensar em mais nada senão</b> nas dores dos nossos joelhos.

# COMPARA

- textos literários
- 30% de cada texto
- Inglês: África do Sul, E.U.A., Reino Unido
- Português: Angola, Brasil, Moçambique, Portugal.



# Pesquisas Possíveis



1. Linguística contrastiva

3. Estratégias e normas de tradução

2. Traduções contrastivas

4. “Tradutês”

# Especificidades de cada língua

# Metodologia “Ping-Pong”

1. **IO** → **PT**: levantamento das traduções em português da expressão de busca;
2. **PO** → **IT**: levantamento desses equivalentes e suas traduções para o inglês
3. **IO** → **PT**: levantamento desses equivalentes e suas traduções para o português
4. **PO** → **IT**: e assim por diante

# O caso de “said”

IO	→	PT
<b>said</b> (310)	→	<b>disse</b> (203) – 65.48% zero (21) – 6,77% respondi (14) – 3,2% perguntou (11) respondeu (7) perguntei (7) retrunquei (5) retrucou/exclamou/sugeriou (2) outras 7 + 15 estruturas diferentes

# O caso de “said”

<b>PO</b>	→	<b>IT</b>
respondi (17)	→	replied (9) answered (6) zero (2)
respondeu (104)	→	answered (38) replied (31) <b>said (14)</b> responded (6) zero (4) made no reply (3) made no response (1)

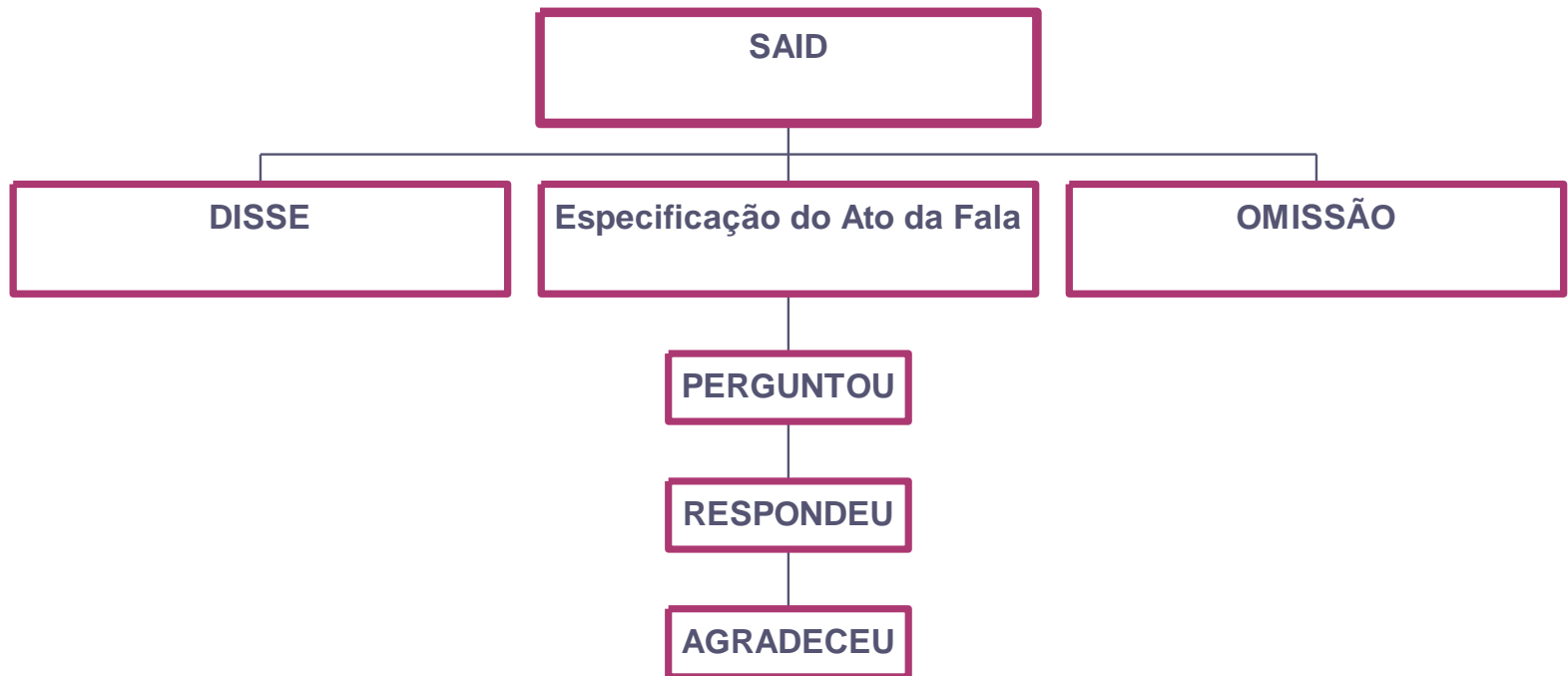
# O caso de “said”

PO	→	IT
perguntei (45)	→	asked (43) zero (2)
perguntou (137)	→	asked (136) zero (8) <b>said (3)</b>
retruquei (0) retrucou (3)	→	replied/retorted/explained

# O caso do verbo “said”

<a href="#">EBDL1T2(240):</a>	«Well, frankly old bean, I'm baffled,» he <b>said</b> .	-- Bem, francamente, meu velho camarada, estou surpreso -- disse ele.
<a href="#">EBDL1T2(247):</a>	«Particularly for me,» I <b>said</b> .	-- Principalmente para mim.
<a href="#">EBDL1T2(167):</a>	«Got a smoke on you?» <b>said</b> the nurse.	-- Tem um cigarro aí? -- perguntou a enfermeira.
<a href="#">EBDL1T2(98):</a>	«Ward 3J,» I <b>said</b> .	«Ala 3J», respondi.
<a href="#">EBDL1T2(319):</a>	She <b>said</b> , «When I don't feel the need to see Karl any more.»	Respondeu, «Quando eu não sentir mais necessidade de ver Karl».
<a href="#">EBDL1T2(357):</a>	«Pathological?» she <b>said</b> .	-- Patológica? -- perguntou ela.
<a href="#">EBDL1T2(914):</a>	«But I am writing,» I <b>said</b> .	-- Mas estou escrevendo -- retruquei.
<a href="#">EBDL1T2(1053):</a>	«This doesn't affect Rhoda, believe me,» he <b>said</b> earnestly.	-- Isso não tem nada a ver com Rhoda, acredite em mim.
<a href="#">EBDL1T2(1660):</a>	I <b>said</b> I could.	Respondi que sim.
<a href="#">EBDL1T2(1728):</a>	«I wont stop, thanks,» she <b>said</b> .	Obrigada, mas não vou ficar -- agradeceu.

# Said vs. Disse





# Resultados

**said** IO – 310=0,4% → **disse** PT – 203=0,25%

**disse** PO – 936=0,23% → **said** IT – 772=0,18%

**told** IT - 59=0,013%

# Resultados

said IO – 310=0,4% → disse PT – 203=0,25%  
disse PT total - 226

disse PO – 936=0,23% → said IT – 772=0,18%  
told IT - 59=0,013% } 0,193%

Conclusão:

1. PT tem maior variação de verbos *dicendi*, mas mantém a forma natural da língua de chegada
2. IT muito pouca variação, mas ainda assim mantém-se mais distante da forma natural da língua de chegada:

IT – 1106 said: 772 ← disse – 69,8%  
334 ← outras – 30,19%

$1106/421.725 = 0,26\%$



Palavras ou  
referências  
culturais

---

# Palavras culturais: cachaça

COMPARA - Resultados da pesquisa cachaça - Microsoft Internet Explorer

Arquivo Editar Exibir Favoritos Ferramentas Ajuda

Endereço <http://www.linguateca.pt/cgi-bin/paralelo.pl>

Google Pesquisar na Web PageRank 80 bloqueado(s) Opções

Procura: "cachaça".

<a href="#">PBAA1</a> (181):	Afinal, já lhe não bastava sortir o seu estabelecimento nos armazéns fornecedores; começou a receber alguns gêneros diretamente da Europa: o vinho, por exemplo, que ele dantes comprava aos quintos nas casas de atacado, vinha-lhe agora de Portugal às pipas, e de cada uma fazia três com água e <b>cachaça</b> ; e despachava faturas de barris de manteiga, de caixas de conserva, caixões de fósforos, azeite, queijos, louça e muitas outras mercadorias.	After a while, he began to buy less from wholesalers and ordered some products directly from Europe -- wine, for example. Before, he had purchased it in demijohns, but now he bought barrels straight from Portugal. He turned each barrel into three by adding water and rum. Likewise, he ordered kegs of butter, crates of canned goods, big boxes of matches, oil, cheese, crockery, and much else besides.
<a href="#">PBAA2</a> (24):	O leiloeiro tinha piscos de olhos significativos; de martelo em punho, entusiasmado, o ar trágico, mostrava com o braço erguido um cálice de <b>cachaça</b> , ou, comicamente acororado esbrocava com o furador os paneiros de farinha e de milho.	The auctioneer winked his eyes meaningfully, excited, with the gavel always in one fist, his countenance serious, he would raise his arm to display a jug of crude brandy, or squat comically to poke at the baskets of flour and corn with a sharp-pointed rod.
<a href="#">PBAA2</a> (1389):	Adorava os perfumes ativos, as jóias e as cores vivas, para ele, nada havia, porém, como um passeio ao sitio embarcado, à fresca da madrugada, bebericando o seu trago de <b>cachaça</b> e pitando o seu fumo do Codó.	He adored strong colognes, jewels, and lively colors. But there was nothing, in his view, equal to a short trip out to his rural property. He would set off in the fresh early morning, swigging gulps of sugarcane brandy and smoking his Codó mixture.
<a href="#">PBAD2</a> (447):	Apesar da ordem para que as vendas se mantivessem fechadas, muita gente fizera de véspera a sua provisão de <b>cachaça</b> e patifaria, e cana corria alegre, bebida mesmo na boca da botija, no bafo e no arrote.	Despite an order that the stores should stay closed, many people had stocked up with spirits and baccy the day before, and the drink was running freely, drunk straight from the bottle, accompanied by burps and stinking breath.
<a href="#">PBAD2</a> (770):	Januário tirou a rolha de sabugo com os dentes, deixou a <b>cachaça</b> cair quente goela abaixo.	Januario pulled the corn-cob stopper out with his teeth and poured the spirit hot down his throat.
<a href="#">PBMR1</a> (1622):	Um ano depois da tragédia, eu substituí a uísque pelo finalô, imediatamente substituído pelo gim e mesmo pela <b>cachaça</b> com limão.	A year after the tragedy, I had replaced whiskey with brandy and water, which was immediately supplanted by gin, and then by sugarcane rum with lime.

Esperamos que o COMPARA lhe tenha sido útil!

[UTILIZAR O COMPARA](#) [MAIS INFORMAÇÕES SOBRE O CORPUS](#) [AGRADECIMENTOS](#) [COMO COLABORAR](#)

Internet

Iniciiar Explora... Microsof... Microsof... COMP... Mapa d... Explora... 10:05

# Cachaça no COMPARA

1(18	<p>Afinal, já-lhe não bastava sortir o seu estabelecimento nos armazéns fornecedores; começou a receber alguns gêneros diretamente da Europa: o vinho, por exemplo, que ele dantes comprava aos quintos nas casas de atacado, vinha-lhe agora de Portugal às pipas, e de cada uma fazia três com água e <b>cacheça</b>; e despachava faturas de barris de manteiga, de caixas de conserva, caixões de fósforos, azeite, queijos, louça e muitas outras mercadorias.</p>	<p>After a while, he began to buy less from wholesalers and ordered some products directly from Europe -- wine, for example. Before, he had purchased it in demijohns, but now he bought barrels straight from Portugal. He turned each barrel into three by adding water and <b>rum</b>. Likewise, he ordered kegs of butter, crates of canned goods, big boxes of matches, oil, cheese, crockery, and much else besides.</p>
2(24	<p>O leiloeiro tinha piscos de olhos significativos; de martelo em punho, entusiasmado, o ar trágico, mostrava com o braço erguido um cálice de <b>cacheça</b>, ou, comicamente acocorado esbrocava com o furador os paneiros de farinha e de milho.</p>	<p>The auctioneer winked his eyes meaningfully; excited, with the gavel always in one fist, his countenance serious, he would raise his arm to display a jug of <b>crude brandy</b>, or squat comically to poke at the baskets of flour and corn with a sharp-pointed rod.</p>
2(13	<p>Adorava os perfumes ativos, as jóias e as cores vivas, para ele, nada havia, porém, como um passeio ao sitio embarcado, à fresca da madrugada, bebendo o seu trago de <b>cacheça</b> e pitando o seu fumo do Codó.</p>	<p>He adored strong colognes, jewels, and lively colors. But there was nothing, in his view, equal to a short trip out to his rural property. He would set off in the fresh early morning, swigging gulps of <b>sugarcane brandy</b> and smoking his Codó mixture.</p>

# Cachaça no COMPARA

AD2 7):	Apesar da ordem para que as vendas se mantivessem fechadas, muita gente fizera de véspera a sua provisão de <b>caçaça</b> e patifaria, e cana corria alegre, bebida mesmo na boca da botija, no bafo e no arroteo.	Despite an order that the stores should stay closed, many people had stocked up with <b>spirits</b> and baccy the day before, and the drink was running freely, drunk straight from the bottle, accompanied by burps and stinking breath.
AD2 0):	Januário tirou a rolha de sabugo com os dentes, deixou a <b>caçaça</b> cair quente goela abaixo.	Januario pulled the corn-cob stopper out with his teeth and poured the <b>spirit</b> hot down his throat.
MR 622	Um ano depois da tragédia, eu substituía o uísque pelo finalô, imediatamente substituído pelo gim e mesmo pela <b>caçaça</b> com limão.	A year after the tragedy, I had replaced whiskey with brandy and water, which was immediately supplanted by gin, and then by <b>sugarcane rum</b> with lime.

# Palavras culturais: pinga

COMPARA - Resultados da pesquisa pinga - Microsoft Internet Explorer

Arquivo Editar Exibir Favoritos Ferramentas Ajuda

Endereço <http://www.linguateca.pt/cgi-bin/paralelo.pl>

Google Pesquisar na Web PageRank 80 bloqueado(s) Opções

**Breve descrição do corpus usado nesta procura:**

Palavras portuguesas	Palavras inglesas	Unidades de alinhamento
487877	524951	3111

6 ocorrências.

**Concordância**

Procura: `[word="pinga" & ((_.texto="PB.*") | (_.variantetrad="brasileira" ))]`

<b>PBAD2</b> (85):	Ali no escuro, à luz esbranquiçada do luar, ele sentia que a <b>pinga</b> de sangue branco que herdara do pai como que o tinha abandonado, de todo restituído à noite selvagem da sua raça.	There in the dark, by the whitish light of the moon, he felt that the drop of white blood he had inherited from his father had deserted him, restoring him completely to the savage night of his race.
<b>PBAD2</b> (448):	E alguns mais altos, a <b>pinga</b> subida na cabeça, já riam e antegozavam, na névoa estúpida da bebida, o grã-guinhol, a fantástica ópera de títeres.	And those who were already high, the drink having gone to their heads, were already laughing in the stupefying cloud of drunkenness and looking forward to the Grand Guignol, the fantastic puppet theatre.
<b>PBAD2</b> (768):	Nhonhô quer um gole de <b>pinga</b> ? disse o negro sentindo frio.	Do you want a drop of rum, massa? asked the Negro, feeling the cold.
<b>PBAD2</b> (769):	Como Januário custasse a responder, ele acrescentou é <b>pinga</b> da cabeça, Nhonhô, esquentá os peitos.	As Januario was slow to answer, he added it's a drop of the best, massa, warms the cockles of the heart.
<b>PBAD2</b> (1026):	A <b>pinga</b> desceu áspera arranhando a goela, foi queimar o estômago vazio.	The spirit was sharp going down, rasping his throat, burning his empty stomach.
<b>PBPM2</b> (407):	Quando abri a porta, encontrei um monte de pacotes na soleira: cigarros, carne moída, cerveja, <b>pinga</b> e flores.	When I opened the door I found a pile of packages on the threshold: cigarettes, ground beef, rum and flowers.

Esperamos que o COMPARA lhe tenha sido útil!

Concluído Internet

Iniciador Explora... Microsoft... Microsoft... COMP... Mapa d... Explora... 10:12

# Pinga no COMPARA

AD2(3):	E alguns mais altos, a <b>pinga</b> subida na cabeça, já riam e antegozavam, na névoa estúpida da bebida, o grã-guinhol, a fantástica ópera de títeres.	And those who were already high, the <b>drink</b> having gone to their heads, were already laughing in the stupefying cloud of drunkenness and looking forward to the Grand Guignol, the fantastic puppet theatre.
AD2(3):	Nhonhô quer um gole de <b>pinga</b> ? disse o negro sentindo frio.	Do you want a drop of <b>rum</b> , massa? asked the Negro, feeling the cold.
AD2(3):	Como Januário custasse a responder, ele acrescentou é <b>pinga</b> da cabeça, Nhonhô, esquentá os peitos.	As Januario was slow to answer, he added it's a <b>drop</b> of the best, massa, warms the cockles of the heart.
AD2(26):	A <b>pinga</b> desceu áspera arranhando a goela, foi queimar o estômago vazio.	The <b>spirit</b> was sharp going down, rasping his throat, burning his empty stomach.
PM2(7):	Quando abri a porta, encontrei um monte de pacotes na soleira: cigarros, carne moída, cerveja, <b>pinga</b> e flores.	When I opened the door I found a pile of packages on the threshold: cigarettes, ground beef, <b>rum</b> and flowers.



# Dubliners

Lourdes Bernardes Gonçalves

Composição dos Corpora

**Corpus de Estudo:** 15 contos de **Dubliners** (James Joyce) - 67.940 palavras

**Corpus de Referência:**

**Crítérios:**

- Contos (mesmo gênero: compressão, densidade de linguagem)
- Sincronia, relevância: Joyce (1882–1941); Woolf (1882–1941); Mansfield (1888–1923); Lawrence (1885–1930)

**Composição:**

- 23 contos de Virginia Woolf (total de 53.659 palavras)
- 12 contos de D. H. Lawrence (total de 79.087 palavras)
- 17 contos de Katherine Mansfield (total de 79.845 palavras)
- Total: 212.591 palavras

## Pesquisa usando *KeyWords* e *Concord*

- **KeyWords:** Importância da Chavicidade negativa:  
*she* (-376,6) ; *her* (-112,5) ; *herself* (-45,2)
- **Concord:** Verificação dos verbos com *she*, *aunt*, *Mrs.* e *Miss(es)*

### **she:**

- ações definidas e independentes, desfrutando a vida, argumentando : 48,76% do total de verbos distintos
- conotação neutra: 42,29% do total de verbos distintos
- indicando atitudes submissas ou angústia: 8,95% do total de verbos distintos

## Pesquisa usando *KeyWords* e *Concord*

### ***aunt***

- ações definidas e independentes, desfrutando a vida, argumentando : 58,54% do *total de verbos distintos*
- conotação neutra: 36,58% do *total de verbos distintos*
- indicando atitudes submissas ou angústia: 4,88% do *total de verbos distintos*

### ***Mrs.***

- ações definidas e independentes, desfrutando a vida, argumentando : 67,35% do *total de verbos distintos*
- conotação neutra: 20,41% do *total de verbos distintos*
- indicando atitudes submissas ou angústia: 12,24% do *total de verbos distintos*

# Estudos da Tradução

(*Lourdes Bernardes Gonçalves*)

## ○ **Corpus paralelo**

- 15 contos de ***Dubliners***
- 15 traduções dos contos de ***Dublinenses***:
  - Hamilton Trevisan
  - José Roberto O'Shea

# Estudos da Tradução

*(Lourdes Bernardes Gonçalves)*

## **Comparação**

**Original vs Trad 1**

**Original vs. Trad 2**

**Trad 1 vs Trad 2**

# Corpora na Tradução

## Vantagens dos corpora:

- podem ser construídos de acordo com as necessidades do tradutor
- podem ser constantemente atualizados
- conferem segurança ao tradutor na escolha do termo a empregar
- fornecem exemplos autênticos de uso →  
**naturalidade para o texto traduzido**

# Pausa para o almoço

*Bom apetite!*

# Projeto e Compilação de um Corpus

Stella E. O. Tagnin

Estudos em Tradução – Teorias, Práticas e Tecnologias

PUC-RS / 10-11/10/2014

2ª. Parte



# Projeto

## 1. Objetivo do corpus

perguntas de pesquisa

- como montar um glossário sobre aviação?
- como estudar a tradução em legendagem?
- como montar um corpus de aprendizes de tradução?
- como um corpus pode ajudar o tradutor?
- como comparar uma tradução com seu original?
- como comparar duas traduções de um mesmo original?

# Projeto

## 2. Decisões sobre o corpus

- a. estático ou dinâmico
- b. falado ou escrito
- c. monolíngüe ou multilíngüe  
(comparável ou paralelo)
- d. gêneros e tipos de textos a serem  
incluídos
- e. domínios a serem incluídos
- f. proporção dos textos

# Projeto

## 2. Decisões sobre o corpus

- g. quantidade de textos
- h. completos ou trechos
- i. extensão dos textos
- j. fonte dos textos
- k. tamanho do corpus

# Projeto

## 2. Decisões sobre o corpus

### I. Balanceamento

- Por número de palavras?
- Por número de textos?

### m. Representatividade

# Compilação de um Corpus

## 3. Coleta dos textos

### a. Fonte

- Internet
- Livro → escanear → revisar

### b. Limpeza

- Eliminar figuras, gráficos etc. – legendas??
- Eliminar (ou etiquetar) URLs
- Eliminar “Referências bibliográficas” (quando for o caso)

### c. Gravar em .txt

# Compilação de um Corpus

## 4. Cabeçalho

- a. quais informações são relevantes para o projeto?
- b. que outras informações poderiam interessar a outros pesquisadores? → **REUSABILIDADE**

# Cabeçalho

```
<Header>
  <title>
    <filename> </filename>
  </title>
  <author>
    <name></name>
  </author>
  <sourceText>
    <language></language>
    <mode>[mode of delivery of textual
      content]</mode>
    <publisher></publisher>
    <pubPlace>[place of publication]</pubPlace>
    <date></date>
    <copyright>[copyrights holder]</copyright>
  </sourceText>
</Header>
```

# Cabeçalho

```
<Header>
  <title>
    <filename> Adv05red001 </filename>
  </title>
  <author>
    <name> Mariazinha da Silva </name>
  </author>
  <sourceText>
    <language> inglês </language>
    <mode> written </mode>
    <publisher> Internet </publisher>
    <pubPlace> http://www.xyz.com </pubPlace>
    <date> September 2, 2012 </date>
    <copyright> [copyrights holder] </copyright>
  </sourceText>
</Header>
```



# Compilação de um Corpus

## 5. Código de nomeação

(Manual Lácio-Web)

**JO-IF-FSP-mu-05fev99\_01**

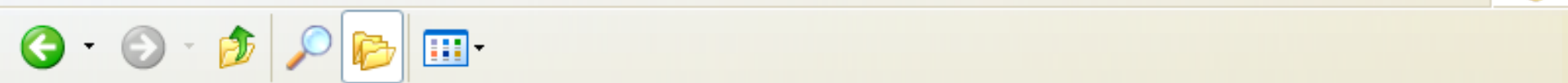
- Texto divulgado em jornal
- Gênero textual: informativo
- Nome do periódico: Jornal “Folha de São Paulo”
- Caderno: “Mundo”
- Data: 05 de fevereiro de 1999
- Primeiro texto (deste caderno, nesta data de publicação)

# Compilação de um Corpus

## 5. Código de nomeação

**RE-IF-NE-cea-mar01\_05**

- Texto divulgado em revista
- Gênero textual: informativo
- Nome do periódico: Revista “Nova Escola”
- Caderno: “Cresça e Aconteça”
- Data: mês de março de 2001
- Quinto texto (deste caderno, nesta data de publicação)



Endereço C:\Documents and Settings\Tagnin\Meus documentos\Corpora\Informática\Inglês\InfoIngAc\Orig Ir

- Pastas
- Corpora
    - Austrália Completo
    - Australia ptg
    - Chico
    - CorRefIng
    - CorRefPtg
    - CorTec
    - Culinaria
      - Culinária Gourmand
    - Fotografia
    - Hipertensão
    - Informática
      - Inglês
        - InfoIngAc
          - Orig
        - InfoIngEmp
        - InfoIngJor
      - Português
    - Investments
    - Journalism
    - Kidney corpus
    - LW Parcial Geral
    - Lw Ref Total
    - MAC-MORPHO\_PESQUISA

- InfoIngAcO 1
- InfoIngAcO 2
- InfoIngAcO 3
- InfoIngAcO 4
- InfoIngAcO 6
- InfoIngAcO 7
- InfoIngAcO 8
- InfoIngAcO 9
- InfoIngAcO 10
- InfoIngAcO 11
- InfoIngAcO 12
- InfoIngAcO 13
- InfoIngAcO 14
- InfoIngAcO 15

Estruturação de um corpus

# Enriquecimento de um Corpus

## 6. Etiquetagem

- a. **morfossintática (POS-tagging)**
- b. **sintática (parsing)**
- c. **semântica: campo semântico**
- d. **discursiva: ingredientes, modo de fazer, rendimento**
- e. **terminológica: termo**

# Etiquetagem morfossintática

- <s>
- Foi\_VAUX
- cercada\_PCP
- de\_PREP|+
- o\_ART
- maior\_ADJ
- sigilo\_N
- a\_ART
- chegada\_N
- de\_PREP|+
- a\_ART
- agência=de=publicidade\_N
- Saatchi\_NPROP
- \$&\_NPROP
- Saatchi\_NPROP
- a\_PREP|+
- o\_ART
- Brasil\_NPROP
- .\_.
- </s>

## Etiquetagem semântica

**For the soup**, preheat the oven to 160°C (350°F / moderate / Gas 4). **<cut>Cut</cut>** **<veg>tomatoes</veg>** lengthwise, discard seeds, place in a medium heatproof dish with **<season>garlic</season>**, olive oil, **<season>salt</season>**, **<season>pepper</season>**, and **<herb>parsley</herb>** and **<herb>basil</herb>** sprigs tied by the stems. **<cook>Bake</cook>** for approximately 1 hour, until **<veg>tomatoes</veg>** are soft and fragrant, let cool and refrigerate for 2 hours, or up to 2 days. **Discard** wilted herbs and blistered tomato skin and puree in a **<appl>blender</appl>** until a smooth paste is obtained (if you want a soup with a more delicate texture, press mixture through a sieve). **Complete** with cold water as to obtain 1 L (1 qt) of soup, adjust **<season>salt</season>** and **<season>pepper</season>**, correct the acidity by adding a pinch of **<season>sugar</season>**, and refrigerate for at least 1 hour, or overnight.

# Etiquetagem

## semântica/terminológica

Caponata (1 hour and 30 minutes)

1 onion

2 **<term>**celery stalks**</term>**

1 **<term>**red bell pepper**</term>**

4 fully ripe tomatoes, peeled and seeded

1 small deep green zucchini (courgette)

2 medium eggplants (aubergines)

2 tablespoons **<term>**pine nuts**</term>**

2 garlic cloves, **<term>**finely chopped**</term>**

1 **<term>**bay leaf**</term>**

1 teaspoon oregano

¼ cup **<term>**red wine vinegar**</term>**

1 tablespoon sugar

2 tablespoons capers

2 tablespoons **<term>**dark raisins**</term>**

½ cup slivered green olives

1 cup flat-leaf parsley leaves

½ cup basil leaves

olive oil

salt and black pepper

**<term>**to taste**</term>**

# Etiquetagem discursiva

**<titRec>** Pudim de Leite Condensado **</titRec>**

**<coment>** Pudim de leite condensado é uma sobremesa que dispensa elogios, ou qualquer palavra para defini-lo. É simplesmente o máximo!! **</coment>**

**<ingr>** Ingredientes:

1 lata de leite condensado

1 lata de leite

3 ovos

essência de baunilha

3 colheres de açúcar **</ingr>**

**<modFaz>** Modo de Preparo:

Coloque o açúcar numa forma própria para pudim e leve ao fogo brando para caramelizar a forma. Bater todos os outros ingredientes no liquidificador. Despeje o conteúdo na forma caramelizada. Levar ao forno em banho-maria. **</modFaz>**

**<coment>** Dica: para verificar se o pudim está pronto, fure o pudim com um palito de dente, se o palito sair limpo, é que está pronto, espere esfriar, desenforme e sirva. **</coment>**



# Etiquetador Brill (LácioWeb)

<\_NN Tit\_NPROP >\_V A\_ART força\_N de\_PREP o\_ART tsunami\_N  
<\_V \_/N

Tit\_NPROP >=<\_N Text\_NPROP >\_NN em\_PREP|+ o\_ART dia\_N  
26\_NUM de\_PREP dezembro\_N ,\_, um\_ART fenômeno\_N de\_PREP  
a\_ART natureza\_N assolou\_V diversas\_PROADJ regiões\_N  
de\_PREP a\_ART Ásia\_NPROP e\_KC África\_NPROP ,\_, comoveu\_V  
e\_KC mobilizou\_V o\_ART mundo\_N inteiro\_ADJ e\_KC trouxe\_V  
para\_PREP o\_ART vocabulário=cotidiano\_N uma\_ART palavra\_N :\_:  
tsunami\_N .\_.A\_ART união\_N dos\_NPROP termos=japoneses\_N  
tsu\_V (\_( porto\_N )\_) e\_KC nami\_N (\_( onda\_N )\_) tem\_V o\_ART  
significado\_N aproximado\_PCP de\_PREP “\_V onda\_N que\_PRO-KS-  
REL atinge\_V o\_ART porto=”\_N ,\_, buscando=dar\_V idéia\_N  
de\_PREP o\_ART alcance\_N e\_KC de\_PREP a\_ART força\_N  
das\_NPROP águas\_N .\_.

# Busca num corpus etiquetado

COCA: [nn\*] and [nn\*]

## Binômios com substantivos

1	MEN AND WOMEN	12813
2	SALT AND PEPPER	1787
3	FAMILY AND FRIENDS	2749
4	LADIES AND GENTLEMEN	2735
5	WOMEN AND CHILDREN	2615
6	BOYS AND GIRLS	2497
7	COLLEGES AND UNIVERSITIES	2494
8	OIL AND GAS	2449
9	RESEARCH AND DEVELOPMENT	2338
10	FRIENDS AND FAMILY	2286
11	FOOD AND DRUG	2050
12	MOM AND DAD	1917
13	FRUITS AND VEGETABLES	1907
14	SCIENCE AND TECHNOLOGY	1825
15	BROTHERS AND SISTERS	1811
16	MOTHER AND FATHER	1785
17	GOODS AND SERVICES	1676
18	LIFE AND DEATH	1625
19	DAY AND NIGHT	1567
20	ARMS AND LEGS	1501
21	WOMEN AND MEN	1494
22	MALES AND FEMALES	1375
23	ARTS AND CRAFTS	1325
24	CONTROL AND PREVENTION	1266
25	UPS AND DOWNS	1249

# Busca por padrão

## Cobuild: N + and + N

young people should be enjoying the loving care **and** support of a family,  
such as the use of the computer, oral history **and** film as historical evidence.  
properties have a balancing effect on body **and** mind. It blends happily with  
singing and brilliant acting combine with humour **and** pathos in this haunting,  
the highest calibre. [h] south [/h] [p] Garlic **and** spice and all things nice, lassi  
in pursuit of its main objectives to save life **and** property from fire. [p] How  
battery pack (see below). [c] model number **and** price [/c] [c] model name and  
These could include the `uncertainty, confusion, **and** insecurity" that Burden (1979-  
me. In return I'll cater for your needs. Photo **and** phone ess. Box 26944. [p]  
salopettes, at £ 77.25, now have elastic **and** Velcro fastenings, double  
American insistence on the separation of church **and** state-or mosque and state-  
stilled used technology developed from the horse **and** cart. [p] Still, Leyland sold  
colours associated with irons. Anthracite finish **and** control lights add to hi-tech feel.  
administration in restoring and maintaining law **and** order. This weekend's military  
She closed her eyes and surrendered her mind **and** body to him. [p] In a matter of

# Cobuild: N + and + N

young people should be enjoying the loving **care and support** of a family,  
such as the use of the computer, oral history **and** film as historical evidence.  
properties have a balancing effect on **body and mind**. It blends happily with  
singing and brilliant acting combine with humour **and** pathos in this haunting,  
the highest calibre. [h] south [/h] [p] Garlic **and** spice and all things nice,  
pursuit of its main objectives to save **life and property** from fire. [p] How  
battery pack (see below). [c] model number **and** price [/c] [c] model name and  
These could include the `uncertainty, confusion, **and** insecurity" that Burden (1979-  
me. In return I'll cater for your needs. Photo **and** phone ess. Box 26944. [p]  
salopettes, at £ 77.25, now have elastic **and** Velcro fastenings, double  
American insistence on the separation of **church and state**-or mosque and state-  
stilled used technology developed from the **horse and cart**. [p] Still, Leyland sold  
colours associated with irons. Anthracite finish **and** control lights add to hi-tech feel.  
administration in restoring and maintaining **law and order**. This weekend's military  
She closed her eyes and surrendered her **mind and body** to him. [p] In a matter of

# Como delimitar uma unidade fraseológica?

## Cobuild: N + and + N

young people should be enjoying the **loving care** **and** **support** of a family,  
such as the use of the computer, oral history **and** film as historical evidence.  
properties have a balancing effect on **body** **and** **mind**. It blends happily with  
singing and brilliant acting combine with humour **and** pathos in this haunting,  
the highest calibre. [h] south [/h] [p] **Garlic** **and** **spice** **and** **all things nice**,  
in pursuit of its main objectives to save **life** **and** **property** from fire. [p] How  
battery pack (see below). [c] model number **and** price [/c] [c] model name and  
These could include the `uncertainty, confusion, **and** insecurity" that Burden (1979-  
me. In return I'll cater for your needs. Photo **and** phone ess. Box 26944. [p]  
salopettes, at £ 77.25, now have elastic **and** Velcro fastenings, double  
American insistence on the separation of **church** **and** **state**-or mosque and state-  
stilled used technology developed from the **horse** **and** **cart**. [p] Still, Leyland sold  
colours associated with irons. Anthracite finish **and** control lights add to hi-tech feel.  
administration in restoring and maintaining **law** **and** **order**. This weekend's military  
She closed her eyes and surrendered her **mind** **and** **body** to him. [p] In a matter of

# Loving care no COCA

- 1 have, you've appreciated how wonderful life is and that it deserves your **loving care**. " # Hanh looks serious. I now realize she is not the cousin
- 2 # Dr. Byock comforts Asorevich's husband, assuring him that he's taking **loving care** of his dying wife # Caption: # "Our first priority is to relieve
- 3 that toddlers and teenagers who must submit to parental guidance, scrutiny, and **loving care** grow into adult children who must negotiate how to guide, advise, and lovingly
- 4 did not scream " new money. " They whispered " slender means but **loving care**. " The freshly painted walls pointed to some aesthetic consideration, but that was
- 5 look up and see him, regular as clockwork, tending his sheep with **loving care** on the side of Ben Fennon.He was the heart and soul of Castle Vandemere.Hester,
- 6 a daycare program that will nurture your child's development while providing warm, **loving care**: Toddler-friendly schedule Ask how regimented the day is. Is there enough flexibility to
- 7 and dull behavior. She might develop seizures. The best treatment is tender **loving care** and appropriate medications as symptoms surface. # Dear Dr. Fox: # My 4-year-old
- 8 Tell them you're lost. Ask for help. JOHN-QUIONES-1-# (Voiceover) But tender **loving care** can come in all shapes and sizes. Watch what happens when this man,
- 9 showed great patience and understanding. I was so tired of this so-called tender **loving care** that I wanted to scream, but instead, I turned myself into a sponge
- 10 But most of the other trees and bushes needed a good deal of tender **loving care** to ensure their continuing survival. In the historical district entire blocks had been razed
- 11 while twirling a stray lock of hair with her finger. " All the **loving care** of home, huh? " Rook asked. " Well, this idiot's
- 12 and given up for dead, yet thanks to the new owners' tender **loving care**, reborn with soul. So now that Garguilo and Demos have created a comfortable
- 13 . " The Results So which company treats your packages with the most tender **loving care**? After crunching the data and averaging the number of spikes recorded by each carrier
- 14 longer version of this column at suntimes.com.) # Brighton Garden' angels provided **loving care** # Dear Fixer: Four years ago, as my father was dying of colon
- 15 SUSAN-SPENCER-1vo: Each of them, says Masters lawyer David Wymore, positioned with **loving care**. DAVID-WYMORE: It shows shower cam and toilet cam. So you sit on

# Loving care no COCA

1 have, you've appreciated how wonderful life is and that it deserves your **loving care**. " # Hanh looks serious. I now realize she is not the cousin

2 # Dr. Byock comforts Asorevich's husband, assuring him that he's taking **loving care** of his dying wife #  
Caption: # " Our first priority is to relieve

3 that toddlers and teenagers who must submit to parental guidance, scrutiny, and **loving care** grow into adult children who must negotiate how to guide, advise, and lovingly

4 did not scream " new money. " They whispered " slender means but **loving care**. " The freshly painted walls pointed to some aesthetic consideration, but that was

5 look up and see him, regular as clockwork, tending his sheep with **loving care** on the side of Ben Fennon. He was the heart and soul of Castle Vandemere. Hester,

5 a daycare program that will nurture your child's development while providing warm, **loving care**: Toddler-friendly schedule Ask how regimented the day is. Is there enough flexibility to

7 and dull behavior. She might develop seizures. The best treatment is **tender loving care** and appropriate medications as symptoms surface. # Dear Dr. Fox: # My 4-year-old

8 Tell them you're lost. Ask for help. JOHN-QUIONES-1-# (Voiceover) But **tender loving care** can come in all shapes and sizes. Watch what happens when this man,

9 showed great patience and understanding. I was so tired of this so-called **tender loving care** that I wanted to scream, but instead, I turned myself into a sponge

10 But most of the other trees and bushes needed a good deal of **tender loving care** to ensure their continuing survival. In the historical district entire blocks had been razed

11 while twirling a stray lock of hair with her finger. " All the **loving care** of home, huh? " Rook asked. " Well, this idiot's

12 and given up for dead, yet thanks to the new owners' **tender loving care**, reborn with soul. So now that Garguilo and Demos have created a comfortable

13 . " The Results So which company treats your packages with the most **tender loving care**? After crunching the data and averaging the number of spikes recorded by each carrier

14 longer version of this column at [suntimes.com](http://suntimes.com).) # Brighton Garden' angels provided **loving care** # Dear Fixer: Four years ago, as my father was dying of colon

15 SUSAN-SPENCER-1vo: Each of them, says Masters lawyer David Wymore, positioned with **loving care**. DAVID-WYMORE: It shows shower cam and toilet cam. So you sit on

# Cobuild: loving + care

se young people should be enjoying the **loving** care and support of a family, the  
are the men who offer her such tender, **loving** care? Are they real or are they  
nse of timeless mystery. A great deal of **loving** care has gone into the creation of  
that it has been hand made with tender **loving** care. [p] Milky Bar - It's just so  
children have found security and tender **loving** care from our dedicated staff.  
ts, shrubs and fruit for sale, grown with **loving** care in our nurseries, including  
e given a new lease of life with a little **loving** care and a certain amount of  
a new lease of life with a little tender **loving** care and a certain amount of  
able, especially if they're surrounded by **loving** care. However tough it is at the  
play in their heyday, a club carved with **loving** care from a block of persimmon  
w of the most common: [p] TLC: Tender **Loving** Care [p] TW's: Time  
regulations have taken over from tender **loving** care and turned well meaning  
of the most common: TLC: Tender **Loving** Care TW's: Time Waters  
of the most common: [p] TLC: Tender **Loving** Care [p] TW's: Time  
that she attempted to get a little tender **loving** care from Russell Bissell. [p] So if I  
but what is really needed is tender **loving** care, which can sometimes be  
simply a matter of gaining some tender **loving** care for a few days, having a  
Throughout history `tender, **loving** care" has uniformly been recognized



# Cobuild: loving + care

these young people should be enjoying the loving care and support of a family, the  
is. Who are the men who offer her such tender loving care? Are they real or are they  
a sense of timeless mystery. A great deal of loving care has gone into the creation of  
that it has been hand made with tender loving care. [p] Milky Bar - It's just so  
1,000 children have found security and tender loving care from our dedicated staff.  
plants, shrubs and fruit for sale, grown with loving care in our nurseries, including  
can be given a new lease of life with a little loving care and a certain amount of  
given a new lease of life with a little tender loving care and a certain amount of  
bearable, especially if they're surrounded by loving care. However tough it is at the  
to play in their heyday, a club carved with loving care from a block of persimmon  
of a few of the most common: [p] TLC: Tender Loving Care [p] TW's: Time  
and regulations have taken over from tender loving care and turned well meaning  
of a few of the most common: TLC: Tender Loving Care TW's: Time Waters  
of a few of the most common: [p] TLC: Tender Loving Care [p] TW's: Time  
that she attempted to get a little tender loving care from Russell Bissell. [p] So if I  
skills, but what is really needed is tender loving care, which can sometimes be  
may be simply a matter of gaining some tender loving care for a few days, having a  
Throughout history "tender, loving care" has uniformly been recognized as a

# Binômios formados por verbos

COCA: [v\*] and [v\*]

1	COME AND GO	1250
2	WAIT AND SEE	1074
3	BORN AND RAISED	923
4	CAME AND WENT	755
5	SHOT AND KILLED	700
6	READ AND WRITE	690
7	CAN AND SHOULD	625
8	SAID AND DONE	625
9	TRY AND GET	624
10	COME AND GONE	520
11	SEE AND HEAR	472
12	TURNED AND LOOKED	464
13	TURNED AND WALKED	463
14	LIVE AND WORK	444
15	COMING AND GOING	425
16	BUY AND SELL	408
17	COME AND GET	397
18	PICK AND CHOOSE	380
19	COVER AND COOK	377
20	GO AND GET	366

## Cobuild: V+or+V

to protest Japan's continued whaling, call **or** write to the Japanese Embassy, 101-104  
the police know of anything suspicious you see **or** hear, you are helping to reduce  
in breast tissue long before a woman can feel **or** notice anything unusual or before  
Neither does Midlands Electricity plc confirm **or** deny any of the claims made in  
Uvistat Sun Cream spf15), so the sun won't burn **or** damage her skin. Unfortunately,  
aircraft are a scourge to those who live, work **or** relax in the country. [p] For several  
which you lay brick or stone paths can shorten **or** increase their apparent distance.  
was pressed by mps of all parties to confirm **or** deny reports that the cost of  
in town. Except for Ella, no one would know **or** care where Sue Anne Norton  
humankind or the goodwill of God can do to avert **or** transform the effects of  
was my experience that a patient's will to live (**or** die) made a substantial difference  
be inadequate because it was unable to predict **or** suggest how managers'  
credit cards are often useless. You can't bribe **or** barter with plastic, sometimes can't  
alcohol or other sedatives. 4. Not drive **or** operate machinery whilst under the  
should use any opportunity that he has to visit **or** contact friends or relations he has  
two major companies and their decisions can make **or** break small publications. [p]  
years are very important and whether I stay **or** go remains to be seen. [p] But I've  
thoroughly

## Cobuild: V+or+V

to protest Japan's continued whaling, call **or** write to the Japanese Embassy, 101-104  
the police know of anything suspicious you **see or hear**, you are helping to reduce  
in breast tissue long before a woman can feel **or** notice anything unusual or before  
Neither does Midlands Electricity plc **confirm or deny** any of the claims made in  
Uvistat Sun Cream spf15), so the sun won't burn **or** damage her skin. Unfortunately,  
aircraft are a scourge to those who live, work **or** relax in the country. [p] For several  
which you lay brick or stone paths can shorten **or** increase their apparent distance.  
was pressed by mps of all parties to **confirm or deny** reports that the cost of  
in town. Except for Ella, no one would know **or** care where Sue Anne Norton  
humankind or the goodwill of God can do to avert **or** transform the effects of  
was my experience that a patient's will to **live (or die)** made a substantial difference  
be inadequate because it was unable to predict **or** suggest how managers'  
credit cards are often useless. You can't **bribe or barter** with plastic, sometimes  
a alcohol or other sedatives. 4. Not **drive or operate** machinery whilst under the  
should use any opportunity that he has to visit **or** contact friends or relations he has  
two major companies and their decisions can **make or break** small publications. [p]  
years are very important and whether I **stay or go** remains to be seen. [p] But I've

## Cobuild: confirm or deny

any way. Neither does Midlands Electricity plc **confirm or deny** any of the claims  
Wakeham was pressed by mps of all parties to **confirm or deny** reports that the cost  
and flown to America. Dr Swire has refused to **confirm or deny** reports that the  
political prisoners. He has always refused to **confirm or deny** whether he's a  
J. Dore which the Princess refused either to **confirm or deny** she wore during the  
Household had seen them, it was impossible to **confirm or deny** that they were any  
of the reason and no one is prepared to **confirm or deny** that the Princess forced  
to show on his face that would either **confirm or deny**. When's the next time  
merely smiled and nodded and refused to **confirm or deny** anything. Of the three  
were, ended there and then. Neither side will **confirm or deny** that they took place  
[p] The Prison Service yesterday refused to **confirm or deny** whether women were  
the Hong Kong police were last night able to **confirm or deny** the report, which  
by Lord Stevens of Ludgate, has refused to **confirm or deny** claims that it is  
soon. [p] The Agriculture Minister refused to **confirm or deny** reports that he had  
s Playboy magazine Elizabeth Ward refuses to **confirm or deny** whether she has  
[p] But Downing Street later refused to **confirm or deny** whether Maxwell's

## Cobuild: confirm or deny

any way. Neither does Midlands Electricity plc **confirm or deny** any of the claims  
Wakeham was pressed by mps of all parties to **confirm or deny** reports that the cost  
and flown to America. Dr Swire has **refused to confirm or deny** reports that the  
political prisoners. He has always **refused to confirm or deny** whether he's a  
J. Dore which the Princess **refused** either to **confirm or deny** she wore during the  
Household had seen them, it was impossible to **confirm or deny** that they were any  
of the reason and no one is prepared to **confirm or deny** that the Princess forced  
to show on his face that would either **confirm or deny**. When's the next time  
merely smiled and nodded and **refused to confirm or deny** anything. Of the  
were, ended there and then. Neither side will **confirm or deny** that they took place  
[p] The Prison Service yesterday **refused to confirm or deny** whether women were  
the Hong Kong police were last night able to **confirm or deny** the report, which  
by Lord Stevens of Ludgate, has **refused to confirm or deny** claims that it is  
soon. [p] The Agriculture Minister **refused to confirm or deny** reports that he had  
s Playboy magazine Elizabeth Ward **refused to confirm or deny** whether she has  
[p] But Downing Street later **refused to confirm or deny** whether Maxwell's

## Cobuild: confirm or deny

any way. **Neither** does Midlands Electricity plc **confirm or deny** any of the claims  
Wakeham was pressed by mps of all parties to **confirm or deny** reports that the cost  
and flown to America. Dr Swire has **refused to confirm or deny** reports that the  
political prisoners. He has always **refused to confirm or deny** whether he's a  
J. Dore which the Princess **refused either to confirm or deny** she wore during the  
Household had seen them, it was **impossible** to **confirm or deny** that they were any  
of the reason and **no one** is prepared to **confirm or deny** that the Princess forced  
to show on his face that would either **confirm or deny**. When's the next time  
merely smiled and nodded and **refused to confirm or deny** anything. Of the  
were, ended there and then. **Neither** side will **confirm or deny** that they took place  
[p] The Prison Service yesterday **refused to confirm or deny** whether women were  
the Hong Kong police were last night able to **confirm or deny** the report, which  
by Lord Stevens of Ludgate, has **refused to confirm or deny** claims that it is  
soon. [p] The Agriculture Minister **refused to confirm or deny** reports that he had  
s Playboy magazine Elizabeth Ward **refuses to confirm or deny** whether she has  
[p] But Downing Street later **refused to confirm or deny** whether Maxwell's

# Exemplos de Corpora

- **Projeto CoMET / CorTec:**  
[www.fflch.usp.br/dlm/comet/consulta\\_cortec.html](http://www.fflch.usp.br/dlm/comet/consulta_cortec.html)
- **Para estudos terminológicos**
  - Informática – só hardware
  - Contratos – número igual de documentos
- **Para estudos socioterminológicos**
  - Ecoturismo – agências governamentais, agências de turismo, instituições ambientais: Brasil e Nova Zelândia



# Turismo Cultural

## *(Jeanette Lammell)*

- **Cidades “similares”:** Rio de Janeiro, São Paulo, Salvador vs Londres, Liverpool, Edimburgo
- **Setores:**
  - **Accomodation**
  - **Fairs and Exhibitions**
  - **Festivals**
  - **General Information**
  - **Museums**
  - **Music**
  - **Theater and Musicals**
  - **Tours**

# Turismo Cultural:

## <header>**Cabeçalho**

<city>**London**</city>

<type>**Accomodation**</type>

<title>**The Landmark London - About**</title>

<author>**unknown**</author>

<date>**unknown**</date>

<source>**Expedia**</source>

<publ>**[http://www.expedia.co.uk/pub/agent.dll?  
qscr=dspv&&itid=&itdx=&itty=&from=f](http://www.expedia.co.uk/pub/agent.dll?qscr=dspv&&itid=&itdx=&itty=&from=f)**</publ>

<access>**May 11, 2007**</access>

</header>

<text>

# Parte 3: Wordsmi th Tools

---

# **Como usar o Wordsmith Tools para investigar um corpus**

# Primeiro

**Baixar versão demo do  
Wordsmith Tools 6.0 do site de  
Mike Scott:**

<http://www.lexically.net/wordsmith/version6/index.html>

# Em seguida

- Siga as instruções para instalar no seu computador

**C** Concord

**K** KeyWords

**W** WordList

- Previous results
- Main settings
- Print settings
- Colour settings
- Folder settings
- Language settings
- Concord
- KeyWords
- WordList
- WSConcgram
- Utilities
- About
- Characters

On startup

- remember screen position
- restore last work saved
- show toolbar in Tools
- show statusbar in Tools

Updates

monthly

Version 6.0.0.120 (09/12/2012)  
latest version =

*type/token ratio is too sensitive to corpus size*

System

Windows default

[Get Started Guide](#) [Support](#)

**WordList**

# WordSmith Tools

- **Statistics**
- **Frequency**
- **Alphabetical**
- **Filenames**
- **Notes**



# WordList

- **Statistics:** visão geral do corpus / texto
- **Frequency:** palavras mais frequentes podem indicar tópico
- **Alphabetical:** facilita a lematização
- **Filenames:** relaciona todos os arquivos no corpus
- **Notes:** anotações pessoais

# WordList - Statistics

## Identificação de peculiaridades

- o corpus (Geral)
- o cada texto

# WORDLIST - STATISTICS

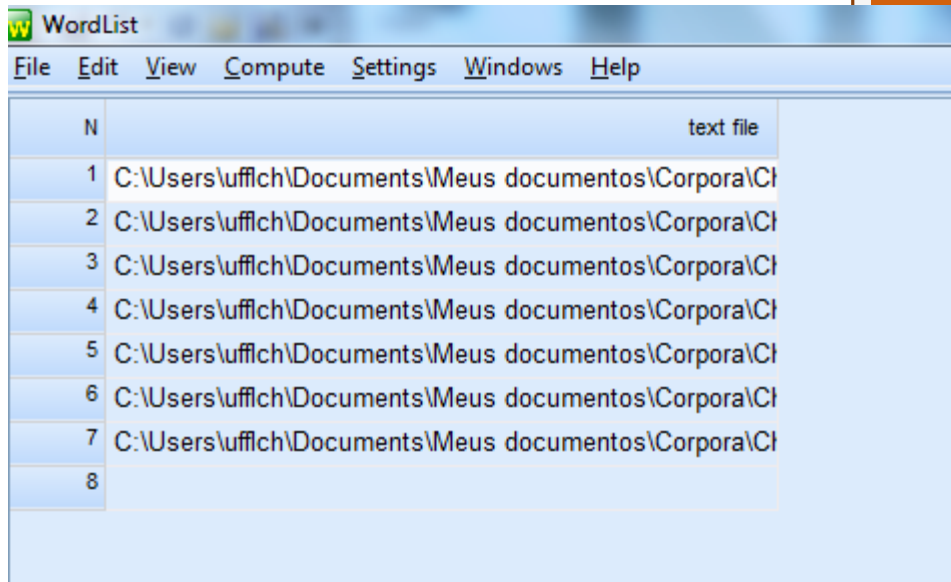
N	text file	file s	tokens (running words)	tokens used for	sum of e	types (disti)	type/ratio	stands T	SSTR s	SSTR b	mean word length	word length s	s mean (in w)	s	p	mean (in w)	s	
1	Overall	2,520,4	4,223	4,155		858	20.65			1,000	5.53	2.15	1,402	2.96	1.23	1,400	2.97	1.23
2	TBacarajé01_I.txt	2,725	4	4		4	100.00			1,000	4.50	0.58	1	4.00		1	4.00	
3	TBalcachofras01_I.txt	1,729	2	2		2	100.00			1,000	9.00	2.83	1	2.00		1	2.00	
4	TBalmôndegas01_I.txt	1,383	3	3		3	100.00			1,000	6.00	2.65	1	3.00		1	3.00	
5	TBalmôndegas02_I.txt	988	2	2		2	100.00			1,000	4.00		1	2.00		1	2.00	
6	TBangu01_I.txt	1,038	1	1		1	100.00			1,000	7.00		1	1.00		1	1.00	
7	TBarroz01_I.txt	1,393	2	2		2	100.00			1,000	6.50	3.54	1	2.00		1	2.00	
8	TBarroz02_I.txt	844	3	3		3	100.00			1,000	5.00	3.61	1	3.00		1	3.00	
9	TBarroz03_I.txt	999	3	3		3	100.00			1,000	6.00	3.46	1	3.00		1	3.00	
10	TBarroz04_I.txt	1,174	3	3		3	100.00			1,000	5.00	1.73	1	3.00		1	3.00	
11	TBarroz05_I.txt	1,278	3	3		3	100.00			1,000	4.67	1.15	1	3.00		1	3.00	
12	TBarroz06_I.txt	1,239	2	2		2	100.00			1,000	5.50	2.12	1	2.00		1	2.00	
13	TBbaba01_I.txt	1,072	2	2		2	100.00			1,000	7.00	1.41	1	2.00		1	2.00	
14	TBbacialhau01_I.txt	1,529	5	5		5	100.00			1,000	5.40	2.70	1	5.00		1	5.00	
15	TBbanana01_I.txt	986	2	2		2	100.00			1,000	6.00	1.41	1	2.00		1	2.00	
16	TBbanana02_I.txt	974	5	5		5	100.00			1,000	4.40	1.52	1	5.00		1	5.00	
17	TBbanana03_I.txt	833	2	2		2	100.00			1,000	5.50	0.71	1	2.00		1	2.00	
18	TBbanana04_I.txt	909	2	2		2	100.00			1,000	6.00	1.41	1	2.00		1	2.00	
19	TBbatata01_I.txt	1,032	4	4		4	100.00			1,000	5.75	1.71	1	4.00		1	4.00	
20	TBbatata02_I.txt	787	3	3		3	100.00			1,000	5.67	2.08	1	3.00		1	3.00	
21	TBbeijinho01_I.txt	803	3	3		3	100.00			1,000	6.33	0.58	1	3.00		1	3.00	

# WORDLIST - FREQUENCY

WordList							
File Edit View Compute Settings Windows Help							
N	Word	Freq.	%	Texts	%	Lemmas	Set
1	DE	1,612	4.33	7	100.00		
2	A	1,319	3.54	7	100.00		
3	E	1,284	3.45	7	100.00		
4	O	1,154	3.10	7	100.00		
5	QUE	1,114	2.99	7	100.00		
6	UM	701	1.88	7	100.00		
7	SE	667	1.79	7	100.00		
8	DO	502	1.35	7	100.00		
9	COM	476	1.28	7	100.00		
10	NO	430	1.16	7	100.00		
11	BENJAMIM	409	1.10	7	100.00		
12	PARA	401	1.08	7	100.00		
13	EM	387	1.04	7	100.00		
14	DA	378	1.02	7	100.00		
15	UMA	372	1.00	7	100.00		
16	ARIELA	361	0.97	7	100.00		
17	NA	352	0.95	7	100.00		
18	NÃO	330	0.89	7	100.00		
19	OS	268	0.72	7	100.00		
20	AS	236	0.63	7	100.00		
21	AO	227	0.61	7	100.00		
22	POR	212	0.57	7	100.00		
23	MAS	196	0.53	7	100.00		
24	É	188	0.51	7	100.00		
25	ELA	177	0.48	7	100.00		
26	SUA	172	0.46	7	100.00		
27	ELE	170	0.46	7	100.00		
28	COMO	166	0.45	7	100.00		

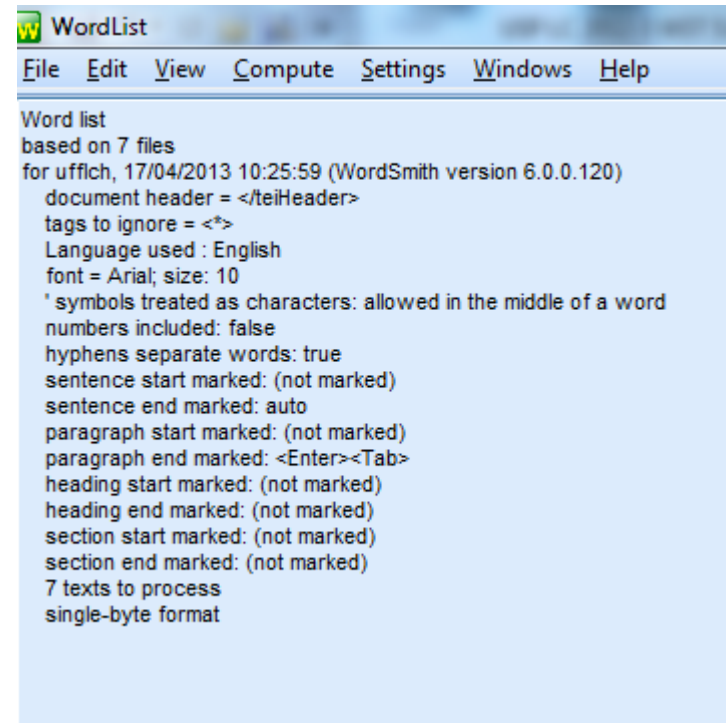
# WORDLIST - ALPHABETICAL

WordList							
File Edit View Compute Settings Windows Help							
N	Word	Freq.	%	Texts	%	Lemmas	Set
1	#	72	0.19	7	100.00		
2	A	1,319	3.54	7	100.00		
3	À	165	0.44	7	100.00		
4	ABAFADA	1		1	14.29		
5	ABAFADA	1		1	14.29		
6	ABAFADOS	1		1	14.29		
7	ABAFASSE	1		1	14.29		
8	ABAIXO	7	0.02	5	71.43		
9	ABAJUR	3		2	28.57		
10	ABALOU	1		1	14.29		
11	ABANA	1		1	14.29		
12	ABANANDO	1		1	14.29		
13	ABANDONADA	2		2	28.57		
14	ABANDONADO	2		2	28.57		
15	ABANDONAR	1		1	14.29		
16	ABANDONARA	1		1	14.29		
17	ABANDONARIA	1		1	14.29		
18	ABANDONE	1		1	14.29		
19	ABANOS	1		1	14.29		
20	ABANOU	1		1	14.29		
21	ABARROTADA	1		1	14.29		
22	ABARROTADO	1		1	14.29		
23	ABATATADO	1		1	14.29		
24	ABATIDOS	1		1	14.29		
25	ABATIMENTO	1		1	14.29		
26	ABDOME	1		1	14.29		
27	ABERTA	10	0.03	4	57.14		
28	ABERTAS	1		1	14.29		



## WORDLIST - FILENAMES

## WORDLIST - NOTES



# Frequency WordList

- **Dicas sobre tópicos do corpus/texto**
- **Palavras mais recorrentes no corpus/texto**

## WordList

File Edit View Compute Settings Windows Help

N	Word	Freq.	%	Texts	% Len
1	#	2,925	4.22	140	97.22
2	A	2,408	3.47	144	100.00
3	O	2,373	3.42	144	100.00
4	DE	2,330	3.36	144	100.00
5	E	1,791	2.58	143	99.31
6	DO	1,344	1.94	142	98.61
7	QUE	1,247	1.80	138	95.83
8	DA	1,163	1.68	144	100.00
9	EM	903	1.30	139	96.53
10	COM	812	1.17	131	90.97
11	PARA	780	1.13	132	91.67
12	NA	770	1.11	140	97.22
13	NO	653	0.94	133	92.36
14	OS	618	0.89	129	89.58
15	COPA	523	0.75	139	96.53
16	UM	515	0.74	128	88.89
17	NÃO	491	0.71	122	84.72
18	É	447	0.64	100	69.44
19	POR	445	0.64	117	81.25
20	UMA	380	0.55	115	79.86
21	MAIS	344	0.50	113	78.47
22	SE	344	0.50	120	83.33
23	FOLHA	324	0.47	39	27.08
24	SELEÇÃO	304	0.44	102	70.83
25	FOI	285	0.41	103	71.53
26	BRASIL	278	0.40	70	48.61

# Alphabetical WordList

## Facilita a

- Localização de palavras
- Lematização
- Pesquisa por sufixos



# LEMATIZAÇÃO

WordList

File Edit View Compute Settings Windows Help

N	Word	Freq.	%	Texts	%	Lemmas	Se
1	#	2,925	4.22	140	97.22		
2	A	2,408	3.47	144	100.00		
3	Á	2		2	1.39		
4	À	107	0.15	69	47.92		
5	Ã	1		1	0.69		
6	Ã	19	0.03	1	0.69		
7	Ã NCIAS	1		1	0.69		
8	AARON	5		2	1.39		
9	ABAIXO	9	0.01	7	4.86		
10	ABAIXOU	1		1	0.69		
11	ABALADA	1		1	0.69		
12	ABALADO	1		1	0.69		
13	ABALOU	2		2	1.39		
14	ABANDONADO	9	0.01	0		abandonado[2] abandonar[4] abandonou[3]	
15	ABANDONAR	4		4	2.78		
16	ABANDONOU	3		3	2.08		
17	ABATIMENTO	1		1	0.69		

# KeyWords


- **Identificação do vocabulário predominante**
- **Corpus de Estudo vs Corpus de Referência**

## Getting Started...

Key words KW Database

### Wordlists

To make a keyword list, you need to choose one or more wordlists ...

 Swap


C:\Users\uffich\Documents\wsmith6\Chico ptg.lst

... and a reference corpus wordlist.

C:\Users\uffich\Documents\wsmith6\LV generalidades\_files.lst

### Keywords...

You can make one keyword list...

 Make a keyword list now

... or a batch of keyword lists, one per wordlist...

Make a batch now

processing 1 files  
English  
p value = 0,0000001

 Help



# KEYWORDS

KeyWords

File Edit View Compute Settings Windows Help

N	Key word	Freq.	%	RC. f	RC. %	t	P	Lemmas	Set
1	COPA	523	0.75	327		3,306.0	0.00000		
2	SELEÇÃO	304	0.44	687	0.01	1,371.7	0.00000		
3	ALEMANHA	267	0.39	620	0.01	1,193.1	0.00000		
4	FOLHA	324	0.47	1,344	0.03	1,139.8	0.00000		
5	UOL	141	0.20	15		1,095.0	0.00000		
6	GOL	187	0.27	177		1,083.6	0.00000		
7	GOLS	159	0.23	72		1,061.2	0.00000		
8	JOGO	250	0.36	814	0.02	979.84	0.00000		
9	CROÁCIA	120	0.17	10		945.37	0.00000		
10	ONLINE	125	0.18	29		909.76	0.00000		

# Estudos Literários

## (*Lourdes Bernardes Gonçalves*)

- Estudo de ***Dubliners***
- **15 contos de *Dubliners***: (James Joyce :1882-1941)

### **Corpus de referência:**

Virginia Woolf (1882-1941): 23 contos – 53.659 pals.

Katherine Mansfield (1888-1923): 12 contos – 79.087 pals.

D. H. Lawrence (1885-1930): 17 contos – 79.845 pals.

**Total 52 contos – 212.591 pals.**

# Estudos Literários

(*Lourdes Bernardes Gonçalves*)

Composição do corpus de referência:

<b>Autores</b>	<b>nº de contos</b>	<b>Total de palavras</b>
Virginia Woolf	23	53.659
Katherine Mansfield	12	79.087
D. H. Lawrence	17	79.845
<b>TOTAL</b>	<b>52</b>	<b>212.591</b>

**Obs.:** Corpus de Estudo: 67.936 palavras  
(3,13 vezes menor)

# Estudos Literários

(*Lourdes Bernardes Gonçalves*)

## Comparação

Joyce vs Corpus de Referência =  
Características de Joyce

## Comparação

Conto X vs [Contos Totais – X]=  
Características do Conto X



# Keywords

- Palavras-chave positivas  
(vocabulário predominante)
- Palavras-chave negativas  
(vocabulário que NÃO ocorre)

N	Key word	Freq.	%	. Freq.	RC. %	eyness	P	emmas	Set
817	IMATURIDADE	8	0,01	19		24,79	006356		
818	TV	19	0,03	147		24,78	006408		
819	CALMA	9	0,01	27		24,66	006827		
820	SENTIU	10	0,01	36		24,57	007142		
821	RITMO	18	0,03	134		24,46	007549		
822	AGORA	50	0,07	738	0,03	24,39	007826		
823	ALEMÃO	33	0,05	391	0,02	24,36	007960		
824	CLUBE	9	0,01	28		24,15	008890		
825	PREFERE	9	0,01	28		24,15	008890		
826	DUAS	79	0,11	1.408	0,06	24,03	009452		
827	OUTROS	23	0,03	1.843	0,08	-24,11	009079		
828	PAPEL	4		801	0,03	-25,49	004426		
829	D	3		728	0,03	-25,54	004311		
830	ESTA	12	0,02	1.305	0,06	-25,72	003908		
831	TEMA	3		739	0,03	-26,10	003210		
832	QUANTO	10	0,01	1.207	0,05	-26,47	002644		
833	IMAGEM	5		896	0,04	-26,70	002343		
834	CRISE	5		905	0,04	-27,14	001864		
835	FORMAÇÃO	3		769	0,03	-27,65	001425		
836	SUAS	30	0,04	2.293	0,10	-27,66	001414		
837	ISTO	3		781	0,03	-28,27	001026		
838	SER	140	0,20	7.078	0,31	-28,33	000994		

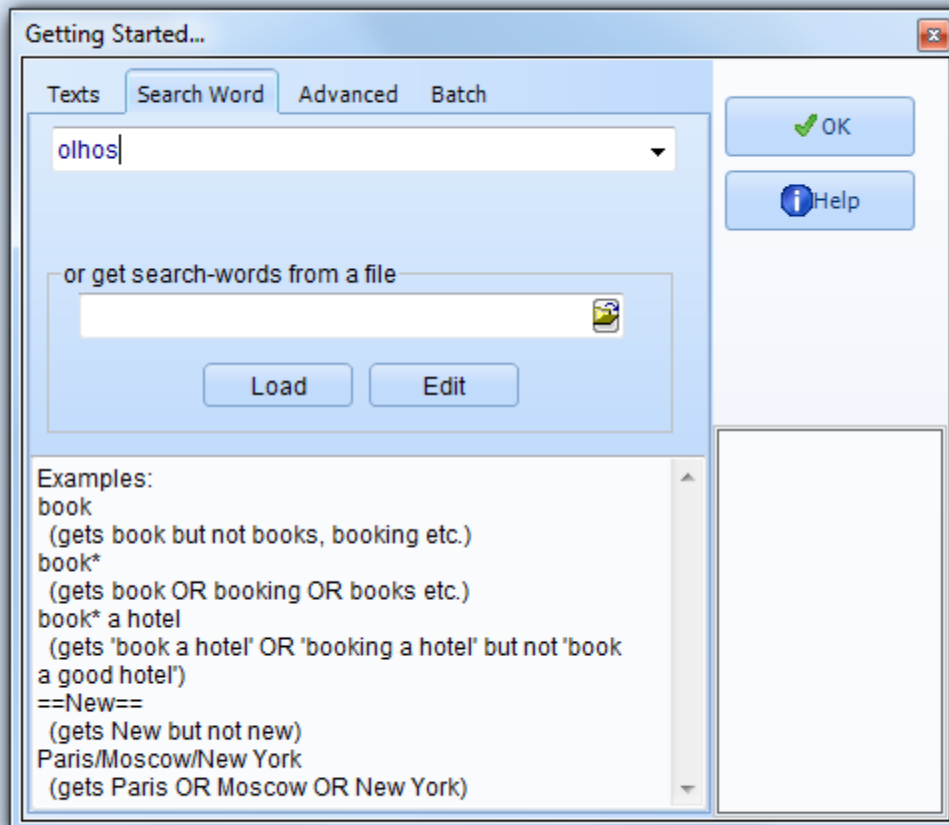
# Concord

- KWIC = Key Word In Context
- Collocates
- Clusters

# Identificação de padrões

## Contexto

- **Linhas de concordância**
  - **Padrões lexicais – colocações**
  - **Padrões gramaticais – coligações**



Concord														
File Edit View Compute Settings Windows Help														
N	Concordance	Set	Word #	Sen	Sen	Para	Para	Hea	Hea	Sec	Sec	File	Date	%
1	de maneira despachada o movimento que, <b>aos olhos</b> de Zorza, transcorre minuciosamente: seu		2,739	115	35%	0	57%			0	57%	Ben03Po.txt	2007/nov/09 00:	57%
2	a tiros, outra explicação não haveria, <b>aos olhos</b> de Jeovan e de seus colegas, senão que		2,175	94	69%	0	40%			0	40%	Ben07Po.txt	2007/nov/09 00:	39%
3	girafa. Nua, a girafa nem parece tão grande <b>aos olhos</b> de Benjamim. Mas Benjamim já notou que		2,797	144	83%	0	49%			0	49%	Ben02Po.txt	2007/nov/09 00:	50%
4	ali pensando em outro homem. Porém, <b>aos olhos</b> de Ariela aqueles dois homens são tão		2,755	150	17%	0	64%			0	64%	Ben04Po.txt	2007/nov/09 00:	64%
5	que não tarda em usurpar-lhe a namorada. <b>Com olhos</b> trinta anos mais velhos, Benjamim		4,739	183	13%	0	96%			0	96%	Ben01Po.txt	2007/nov/09 00:	96%
6	os espectadores transfiguravam-se, e <b>com olhos</b> mais e mais esbugalhados pareciam		2,136	89	56%	0	39%			0	39%	Ben07Po.txt	2007/nov/09 00:	38%
7	479, destino largo do Elefante, e faz o trajeto <b>de olhos</b> fechados. Viu a moça pela segunda vez		3,545	135	95%	0	72%			0	72%	Ben01Po.txt	2007/nov/09 00:	71%
8	destaca o par de páginas com a manequim <b>de olhos</b> pintados e leva-a à janela, para analisá-la		4,503	173	59%	0	91%			0	91%	Ben01Po.txt	2007/nov/09 00:	91%
9	em dezesseis cubos e põe-se a mastigá-los <b>de olhos</b> fechados, sorrindo. O pianista começa a		3,506	193	87%	0	81%			0	81%	Ben04Po.txt	2007/nov/09 00:	81%
10	mas ele prossegue na rota de Ariela, com <b>dois olhos</b> exorbitantes. Não é um olhar igual ao de		3,932	191	94%	0	69%			0	69%	Ben02Po.txt	2007/nov/09 00:	70%
11	do início ao fim, tal qual um filme, na venda <b>dos olhos</b> . Mais rápido que uma bala, o filme poderia		114	3	100	0	2%			0	2%	Ben01Po.txt	2007/nov/09 00:	2%
12	dois filetes fluindo sem cessar da cavidade <b>dos olhos</b> em direção aos ouvidos. E Ariela chorou		2,500	107	82%	0	46%			0	46%	Ben07Po.txt	2007/nov/09 00:	45%
13	e viu as lágrimas que brotavam <b>dos olhos</b> de Jeovan, e se acumulavam, e formavam		2,465	106	42%	0	45%			0	45%	Ben07Po.txt	2007/nov/09 00:	45%
14	por enquanto uma íris embaralhada diante <b>dos olhos</b> e não dispõe de qualquer critério para		4,135	160	48%	0	84%			0	84%	Ben01Po.txt	2007/nov/09 00:	84%
15	nada porque a cusparada o atingiu dentro <b>dos olhos</b> . Nem Benjamim nem ninguém saberá		1,736	70	100	0	38%			0	38%	Ben05Po.txt	2007/nov/09 00:	38%

The image shows a screenshot of the Concord software interface. The main window displays a concordance table with columns for 'Set' and 'v'. The 'Edit' menu is open, showing options like Undo, Redo, Copy, and Resort (F6). The 'Resort' option is highlighted. A 'Concordance Sort' dialog box is overlaid on the main window, showing a list of sort options: L5, L4, L3, L2, L1, Centre, R1, R2, R3, R4, R5, Context word, Set order, File order, and Date. The 'R1' option is selected and highlighted in blue. A yellow tooltip labeled 'primary sort' points to the 'R1' option. The dialog box also includes 'OK' and 'Help' buttons, a dropdown menu set to 'All', and checkboxes for 'Case sensitive' and 'Ascending'. The current sort settings are displayed as 'SortAll' and 'R1'.

Concord

File Edit View Compute Settings Windows Help

Undo Ctrl+U

Redo

Set column

Copy Ctrl+C

Deleting

Restore deleted

Insert tags

**Resort F6**

Zap Ctrl+Z

Find... Ctrl+Alt+F

Find next deleted entry Alt+D

Remove duplicates

Change language

Concordance Sort

Main Sort Sort 2 Sort 3

Main Sort ...

L5

L4

L3

L2

L1

Centre

**R1**

R2

R3

R4

R5

Context word

Set order

File order

Date

primary sort

OK

Help

All

Current sort settings:  
SortAll  
R1

Case sensitive

Ascending

back (Suécia), o  
para Gana por 2  
or número de  
or com maior  
ITA) - 5 - Equipe  
mentem. Confira  
s de Trinidad e  
s dez minutos,  
ol como este.  
te não tem mais  
marcado, quem  
or de imprensa da  
Ãrum  
o e pela  
travessão aos  
io. E vozes com

14 ao México. Os mexicanos até procura  
15 talento de seu melhor jogador para che  
16 em média-- brincavam e riam. De repe  
17 partiu com tudo para o ataque e quase chegou ao gol de empate aos quatro minutos. Depois de uma  
18 A Itália sem ter chegado uma única vez com perigo ao gol de Cech, saiu na frente. Totti bateu escanteio aos

# PADRÕES - COLOCAÇÕES

N	Concordance	Set
23	fazer o gol. GE.Net – A defesa cochilou no segundo gol argentino? Zinha – Foi um descuido, mas se esse	
24	mas viu a bola pegar efeito e entrar em seu próprio gol. Assim como na etapa inicial, apenas a Austrália	
25	Os espanhóis ainda foram os que mais chutaram a gol até aqui, 33 vezes. Os caribenhos de Trinidad e	
26	Em um lance confuso, a Austrália chegou perto do gol. Atento em uma bola espirrada, Chipperfield	
27	chegando, sufocando, e Trezeguet perdendo um gol atrás do outro. Ribery, com "ciúmes", repetiu o	
28	Bildes mit der Bildunterschrift: Iaquina vibra ao marcar gol "Azzurra" deixa crise de lado contra Gana, supera	
29	(ING), Valencia, Carlos Tenorio e De la Cruz (EQU) Gol: Beckham (25'/2ºT) INGLATERRA: Robinson,	
30	por Lúcio, a bola passou por Kaká e Cafu, o segundo gol brasileiro surgiu. Após cruzamento da direita,	
31	caiu fora na primeira fase sem marcar um único gol. Com a vitória, o time francês garantiu a segunda	
32	a 0 e eliminar a Suíça do Mundial sem levar um único gol com bola rolando.	
33	grande movimentação, o Brasil chegou ao segundo gol com Juninho. O meia, que aproveitou muito bem a	
34	a monotonia em campo e passaram a buscar o gol com mais ênfase a partir dos 20 minutos. Os	
35	permitindo que a Croácia criasse várias chances de gol. Com a entrada de Robinho no lugar de Ronaldo, o	
36	tentar outra vez não vai fazer em dez anos um gol como aquele. GE – Foi mais sorte do que talento?	
37	– Ele chutou a gol, mas dificilmente vai fazer outro gol como este. Acho que foi mais sorte. Seleção	
38	sofreu, em 12 minutos, contra a Alemanha. Ou um gol, como a Argentina sofreu do México. Com um	
39	ele tentar outra vez, não vai fazer em dez anos um gol como aquele", desdenhou. Com uma atuação	
40	para Dida. Mas o Brasil, que, mesmo após sofrer o gol, continuou buscando o ataque, foi premiado. O	
41	contra Angola. É muito perigoso e criou a jogada do gol contra Angola jogando atrás", lembrou. "É	
42	gols e ainda deu o chute que foi desviado por Petit no gol contra. É ótimo marcar dois gols para a seleção.	
43	Franck Ribery, o argentino Maxi Rodríguez - cujo gol contra o México foi o mais bonito da Copa,	
44	a Azzurra só só foi vazada uma única vez: um gol contra do lateral Zaccardo, no empate diante dos	
45	O jogo Os croatas começaram na frente, com um gol contra de Pablo Ibañez, logo aos 14min do	
46	Zinha – Todos deram o máximo, mas faltou fazer o gol. Criamos chances, mas não soubemos fazer o gol.	



# Collocates

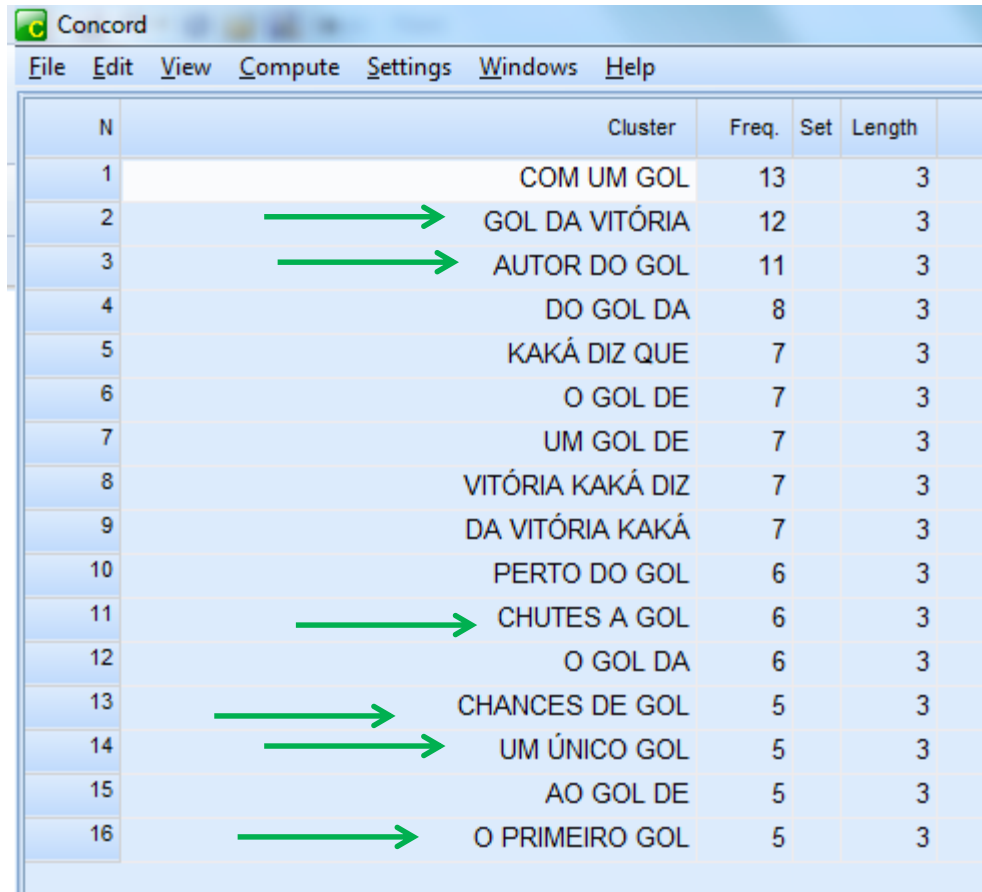
- **Posição dos colocados mais frequentes**



# Clusters

- Agrupamentos de palavras: 2, 3, 4
- “Lexical bundles”
- não são necessariamente unidades completas de sentido

# CLUSTERS



N	Cluster	Freq.	Set	Length
1	COM UM GOL	13		3
2	GOL DA VITÓRIA	12		3
3	AUTOR DO GOL	11		3
4	DO GOL DA	8		3
5	KAKÁ DIZ QUE	7		3
6	O GOL DE	7		3
7	UM GOL DE	7		3
8	VITÓRIA KAKÁ DIZ	7		3
9	DA VITÓRIA KAKÁ	7		3
10	PERTO DO GOL	6		3
11	CHUTES A GOL	6		3
12	O GOL DA	6		3
13	CHANCES DE GOL	5		3
14	UM ÚNICO GOL	5		3
15	AO GOL DE	5		3
16	O PRIMEIRO GOL	5		3

# Plot

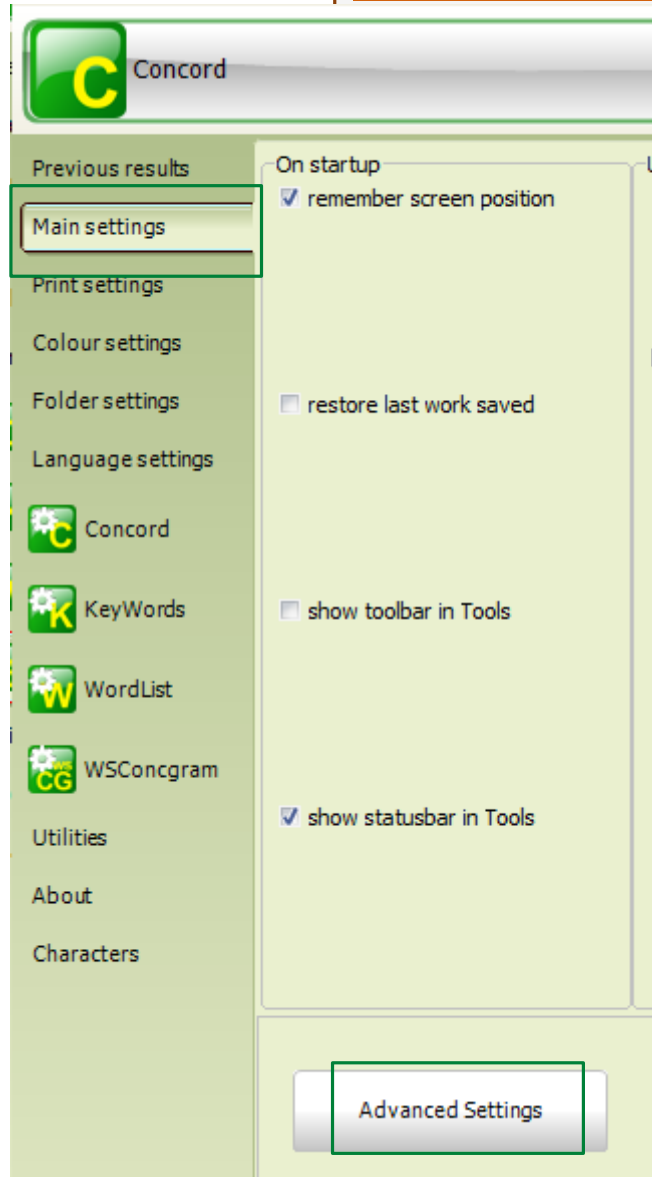
- Distribuição da palavra no texto

# PLOT – “olhos”

Concord

File Edit View Compute Settings Windows Help

N	File	Words	Hits	per		Plot
1	Overall	36,434	63	1.73	0.871	
2	Ben01Po.txt	4,934	13	2.63	0.697	
3	Ben02Po.txt	5,679	8	1.41	0.596	
4	Ben03Po.txt	4,829	9	1.86	0.667	
5	Ben04Po.txt	4,309	6	1.39	0.644	
6	past demo	past de	past de	past de	past de	
7	past demo	past de	past de	past de	past de	
8	past demo	past de	past de	past de	past de	



# Como ignorar texto indesejado

## Pela etiquetagem:

- Título
- Subtítulo
- Figura
- Data
- URL
- etc.

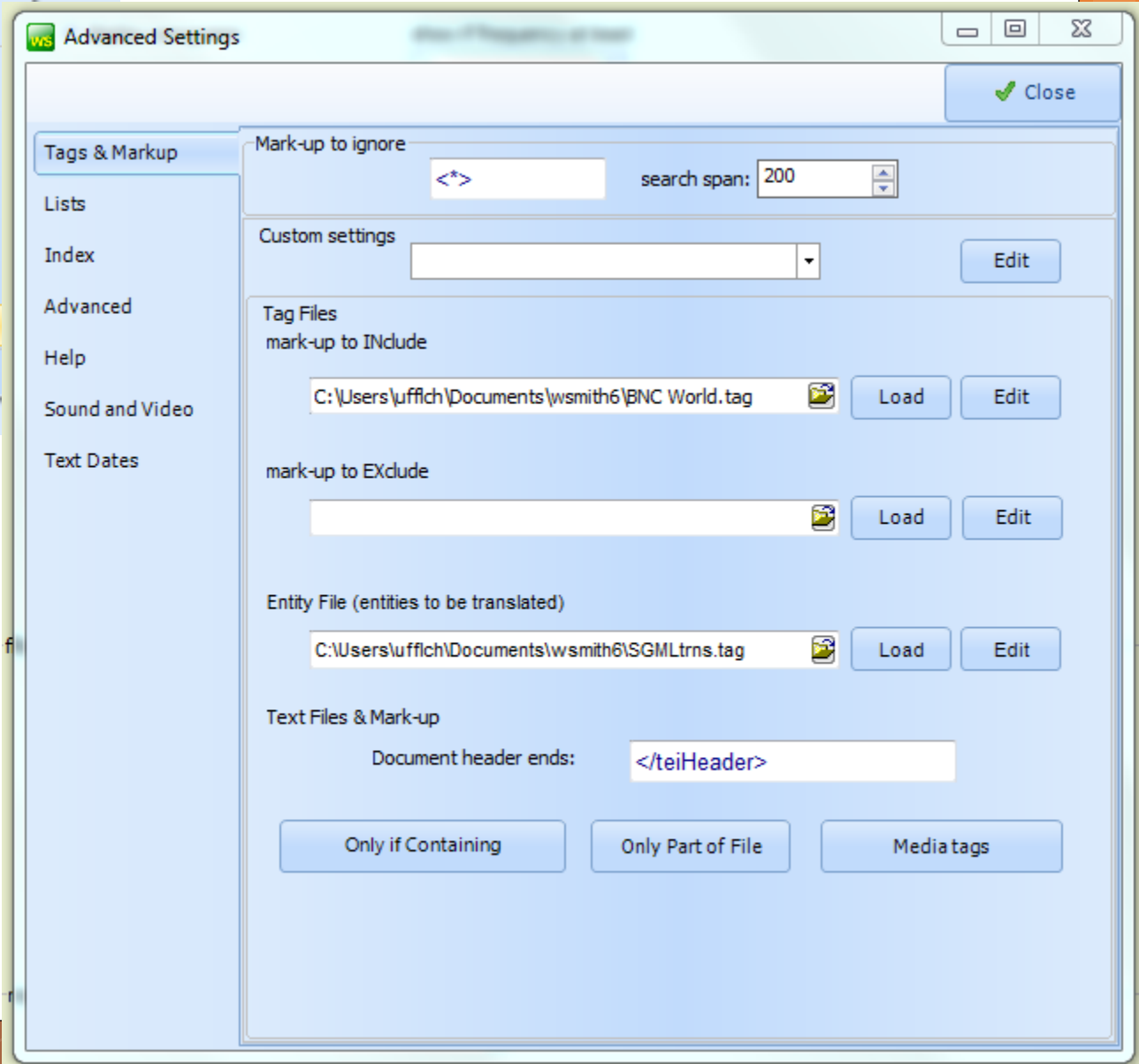
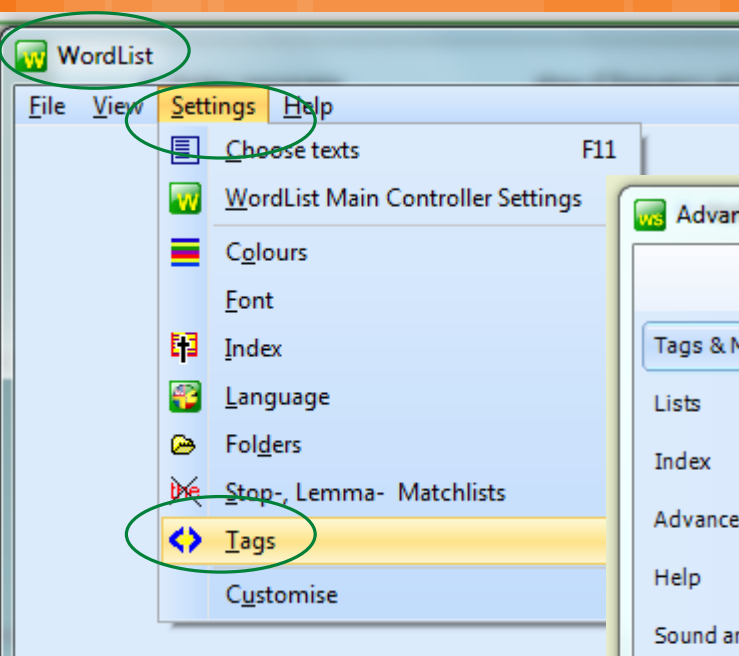


Para só aparecer determinado trecho na concordância

## Adjust Settings

### Wordlist

- → Settings
- → Tags
- → Only part of file
  - Sections to cut out
  - Sections to keep



Advanced Settings



Close

Tags & Markup

Lists

Index

Advanced

Help

Sound and Video

Text Dates

Mark-up to ignore

search span: 200

Custom settings

Tag Files

mark-up to Include

C:\Users\uffich\Documents\wsmith6\BNC World.tag

mark-up to EXclude

Entity File (entities to be translated)

C:\Users\uffich\Documents\wsmith6\SGMLtrns.tag

Text Files & Mark-up

Document header ends:

</teiHeader>

Only if Containing

Only Part of File

Media tags

Select Only sections of text files

Sections to cut out

Cut start of each line/paragraph

0

characters

 to 

----- or -----

 to 

----- or -----

 to 

Sections to keep

<Tit> to </Tit>

----- or -----

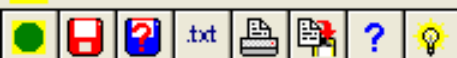
 to 

----- or -----

 to 

OK

Help



N	Concordance	Set	Tag	Word No.	File	%
1	istória</Tit> <Tit>Sismo que causou tsunami foi o 2.º maior da história</Tit>			14	\tsup3.txt	29
2	<Tit>Rescaldo tóxico do tsunami na tsunami na África</Tit>			24	\tsup3.txt	58
3	<Tit>tsunami			10	\tsup3.txt	18
4	ortos pelo <Tit>Número de mortos pelo tsunami chega a 280 mil</Tit>			5	\tsup3.txt	1
5	<Tit>Tsunami Tsunami (onda)</Tit>			0	\tsup2.txt	0
6	<Tit>A força do força do tsunami</Tit>			3	\tsup1.txt	0
7	<Tit>O que é um tsunamiO que é um tsunami</Tit>			13	\tsup1.txt	24

- **Visualiza todos os textos**
- **Alinha textos paralelos**

- **Utilities**

- **Viewer & Aligner**

WordSmith Tools 6.0

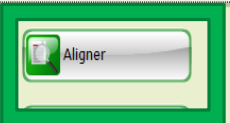
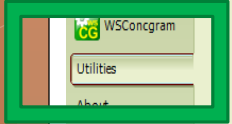
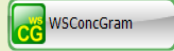
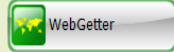
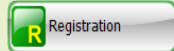
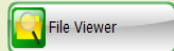
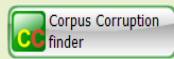
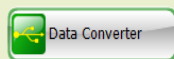
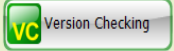
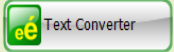
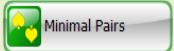
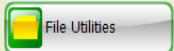
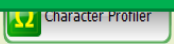
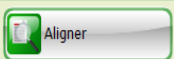
unregistered



save all settings



- Previous results
- Main settings
- Print settings
- Colour settings
- Folder settings
- Language settings
- Concord
- KeyWords
- WordList
- WSConcGram**
- Utilities
- About
- Characters



# Passo a passo – viewer & aligner

1. Language Settings → língua do texto → **Save all settings**
2. File → Open → Escolhe o texto 1 (original)
3. Salvar no Viewer (vwr)
4. Language Settings → língua do texto (se for diferente)
5. File → Open → Escolhe o texto 2 (traduzido)
6. Salvar no Viewer (vwr)
7. Abre texto 1
8. File → Merge with → texto 2
9. Salvar no formato .ali

# Dicas para o alinhamento

A partir de problemas no alinhamento

1. Abrir cada texto no Viewer
2. Fazer as correções separadamente
3. Dá para colocar as duas telas juntas para ir comparando
4. **SALVAR SEMPRE**
5. **ALINHAR (Merge)**
6. Repetir até estar tudo perfeito.



1 Benjamim

2 Chico Buarque

3 O pelotão estava em forma, a voz de comando foi enérgica e a fuzilaria produziu um único estrondo. Mas para Benjamim Zambraia soou como um rufo, e ele seria capaz de dizer em que ordem haviam disparado as doze armas ali defronte. Cego, identificaria cada fuzil e diria de que cano partira cada um dos projéteis que agora o atingiam no peito, no pescoço e na cara. Tudo se extinguiria com a velocidade de uma bala entre a epiderme e o primeiro alvo letal (aorta, coração, traquéia, bulbo), e naquele instante Benjamim assistiu ao que já esperava: sua existência projetou-se do início ao fim, tal qual um filme, na venda dos olhos. Mais rápido que uma bala, o filme poderia projetar-se uma outra vez por dentro das suas pálpebras, em marcha a ré, quando a sucessão dos fatos talvez resultasse mais aceitável. E ainda sobraria um fiapo de tempo para Benjamim rever-se aqui e acolá em situações que preferiria esquecer, as imagens ricocheteando no bojo do seu crânio. O prazo se esgotaria e sobreviria um ultimato, um apito, um alarme, mas Benjamim os entenderia como ameaça de criança contando até três - um... dois... dois... dois e meio... - e se deteria mais um pouco em momentos que lhe pertenciam, e que antes não soubera apreciar. Aprenderia também a penetrar em espaços que não conhecera, em tempos que não eram o seu, com o senso de outras pessoas. E súbito se surpreenderia a caminhar simultaneamente em todas as direções, e tudo alcançaria de um só olhar, e tudo o que ele percebesse jamais cessaria, e mesmo a infinitude caberia numa bolha no interior do sonho de um homem como Benjamim Zambraia, que não se lembra de alguma vez ter morrido em sonho.

4 Se uma câmera focalizasse Benjamim na hora do almoço, captaria um homem longilíneo, um pouco curvado, com vestígios de atletismo, de cabelos brancos mas bastos, prejudicado por uma barba de sete dias, camisa para fora da calça surrada aparentando desleixo e não penúria, estacionado em frente ao Bar-Restaurante Vasconcelos, tremulando os joelhos como se esperasse alguém. Benjamim entretanto não espera nada, a não ser que ele mesmo resolva o dilema: entrar num bar-restaurant ou voltar para a cama. A questão é embaraçosa porque Benjamim não tem sono, nem sede, nem apetite, nem alternativa para esta tarde. Preso ao chão, as pernas irrequietas, impacienta-se com a própria hesitação, e é nessas conjunturas que lhe costuma voltar a sensação de estar sendo filmado.

5 Adolescente, Benjamim adquiriu uma câmera invisível por entender que os colegas mais astutos já possuíam as suas. O equipamento mostrou-se tão providencial quanto um pente de bolso, e a partir daquele dia a vida dele tomou novo rumo. Benjamim passou a usar topete, e nas pendengas em que antes se descabelava, certo de estar com a razão, mantinha agora um sorriso vago e deixava o adversário a gesticular de costas para a câmera. Com isso ganhou prestígio e beijou na boca muitas garotas, cujos ombros, orelhas e rabos-de-cavalo foram imortalizados em suas películas. O acervo de Benjamim também guarda dublagens de cantor de jazz, saltos de trampolim, proezas no futebol, brigas de rua em que sangrou ou se saiu bem e a sua estréia no sexo com uma senhora de idade (trinta anos, trinta e um, trinta e três), quando ele quase estragou a cena ao olhar para a lente. Fez-se filmar durante toda a juventude, e só com o advento do primeiro cabelo branco decidiu abolir a ridícula coisa. Era tarde: a câmera criara autonomia, deu de encarapitar-se em qualquer parte para flagrar episódios medíocres, e Benjamim já teve ganas de erguer a camisa e cobrir o rosto no meio da rua, ou de investir contra o cinegrafista, à maneira dos bandidos e dos artistas principais. Hoje ele é um homem amadurecido e usa a indiferença como tática para desencorajar as filmagens. Mas quando entra enfim no Bar-Restaurante Vasconcelos, ainda o incomoda a suspeita de uma câmera, talvez acoplada ao bico do ventilador de longas pás que gira no teto. Benjamim não pode ignorar que, daquele improvável ponto de vista, os fregueses circulariam sentados num carrossel, e ele próprio seguiria num redemoinho até o centro do salão, faria piruetas, daria ordens ao garçom como a um satélite e fugiria às tontas para o banheiro.

6 Entre as pastilhas de naftalina no mijadouro há duas guimbas, na verdade dois tocos de filtro, um com bocal amarelo sarapintado e outro com bocal branco manchado de batom. Benjamim já não se lembra do que veio fazer no banheiro e pensa "fascistas", de repente indignado com os sujeitos que atiram guimbas nos mijadouros de restaurantes sem querer saber se alguém virá recolhê-las mais tarde, um empregado que seguramente não disporá de luvas de borracha ou instrumento apropriado (garfo, pinça, espeto, tridente, aspirador portátil) para tal serviço. Compadecido do empregado, decide catar ele mesmo as guimbas. Mete a mão no mijadouro, e naquele momento entra no banheiro um rapaz com o cabelo na testa e gravata frouxa (jovem repórter, operador da bolsa, jogador de pôquer, parente egresso de enterro) que olha para Benjamim e desvia o olhar. Talvez o rapaz tenha visto Benjamim com a mão no mijadouro, mas também é possível que evitasse o confronto porque essa é a etiqueta em pequenos ambientes públicos. Pudor semelhante leva o rapaz a evitar o mijadouro vizinho ao de Benjamim e a urinar na latrina, deixando a porta aberta por pudor do pudor. Benjamim permanece com as guimbas na palma da mão, pois planejava atirá-las na latrina, e quando outro sujeito entra no banheiro arrastando as sandálias, desloca-se para a pia. Procura em vão alguma cesta de lixo, recusando-se a largar as guimbas na beira da pia ou no chão. Abre a torneira e passa-as de mão para mão debaixo da água corrente. Nota que a guimba branca traz impressas em dourado quatro argolas e a marca Dam; a água não borra a mancha de batom, antes acentua o seu grená. Também nota que a guimba amarela está retorcida, mas como ninguém amassa cigarros num mijadouro, deduz que seu fumante era dado a morder o filtro. Chegam mais dois rapazes falando alto, o despenteado não libera a latrina, e Benjamim experimenta seultar as guimbas pelo ralo da pia. Mas no fundo do ralo há uma cruz de ferro e elas entalam ali. Benjamim lava as mãos com o líquido da

1	Benjamim
2	Chico Buarque
3	O pelotão estava em forma, a voz de comando foi enérgica e a fuzilaria produziu um único estrondo.
4	Mas para Benjamim Zambraia soou como um rufo, e ele seria capaz de dizer em que ordem haviam disparado as doze armas ali defronte.
5	Cego, identificaria cada fuzil e diria de que cano partira cada um dos projéteis que agora o atingiam no peito, no pescoço e na cara.
6	Tudo se extinguiria com a velocidade de uma bala entre a epiderme e o primeiro alvo letal (aorta, coração, traquéia, bulbo), e naquele instante Benjamim assistiu ao que já esperava: sua existência projetou-se do início ao fim, tal qual um filme, na venda dos olhos.
7	Mais rápido que uma bala, o filme poderia projetar-se uma outra vez por dentro das suas pálpebras, em marcha a ré, quando a sucessão dos fatos talvez resultasse mais aceitável.
8	E ainda sobraria um fiapo de tempo para Benjamim rever-se aqui e acolá em situações que preferiria esquecer, as imagens ricocheteando no bojo do seu crânio.
9	O prazo se esgotaria e sobreviria um ultimato, um apito, um alarme, mas Benjamim os entenderia como ameaça de criança contando até três - um... dois... dois... dois e meio... - e se deteria mais um pouco em momentos que lhe pertenciam, e que antes não soubera apreciar.
10	Aprenderia também a penetrar em espaços que não conhecera, em tempos que não eram o seu, com o senso de outras pessoas.
11	E súbito se surpreenderia a caminhar simultaneamente em todas as direções, e tudo alcançaria de um só olhar, e tudo o que ele percebesse jamais cessaria, e mesmo a infinitude caberia numa bolha no interior do sonho de um homem como Benjamim Zambraia, que não se lembra de alguma vez ter morrido em sonho.
12	Se uma câmera focalizasse Benjamim na hora do almoço, captaria um homem longilíneo, um pouco curvado, com vestígios de atletismo, de cabelos brancos mas bastos, prejudicado por uma barba de sete dias, camisa para fora da calça surrada aparentando desleixo e não penúria, estacionado em frente ao Bar-Restaurante Vasconcelos, tremulando os joelhos como se esperasse alguém.
13	Benjamim entretanto não espera nada, a não ser que ele mesmo resolva o dilema: entrar num bar-restaurant ou voltar para a cama.
14	A questão é embaraçosa porque Benjamim não tem sono, nem sede, nem apetite, nem alternativa para esta tarde.
15	Preso ao chão, as pernas irrequietas, impacienta-se com a própria hesitação, e é nessas conjunturas que lhe costuma voltar a sensação de estar sendo filmado.
16	Adolescente, Benjamim adquiriu uma câmera invisível por entender que os colegas mais astutos já possuíam as suas.
17	O equipamento mostrou-se tão providencial quanto um pente de bolso, e a partir daquele dia a vida dele tomou novo rumo.
18	Benjamim passou a usar topete, e nas pendengas em que antes se descabelava, certo de estar com a razão, mantinha agora um sorriso vago e deixava o adversário a gesticular de costas para a câmera.
19	Com isso ganhou prestígio e beijou na boca muitas garotas, cujos ombros, orelhas e rabos-de-cavalo foram imortalizados em suas películas.
20	O acervo de Benjamim também guarda dublagens de cantor de jazz, saltos de trampolim, proezas no futebol, brigas de rua em que sangrou ou se saiu bem e a sua estréia no sexo com uma senhora de idade (trinta anos, trinta e um, trinta e três), quando ele quase estragou a cena ao olhar para a lente.
21	Fez-se filmar durante toda a juventude, e só com o advento do primeiro cabelo branco decidiu abolir a ridícula coisa.
22	Era tarde: a câmera criara autonomia, deu de encarapitar-se em qualquer parte para flagrar episódios medíocres, e Benjamim já teve ganas de erguer a camisa e cobrir o rosto no meio da rua, ou de investir contra o cinegrafista, à maneira dos bandidos e dos artistas principais.
23	Hoje ele é um homem amadurecido e usa a indiferença como tática para desencorajar as filmagens.
24	Mas quando entra enfim no Bar-Restaurante Vasconcelos, ainda o incomoda a suspeita de uma câmera, talvez acoplada ao bico do ventilador de longas pás que gira no teto.
25	Benjamim não pode ignorar que, daquele improvável ponto de vista, os fregueses circulariam sentados num carrossel, e ele próprio seguiria num redemoinho até o centro do salão, faria piruetas, daria ordens ao garçom como a um satélite e fugiria às tontas para o banheiro.
26	Entre as pastilhas de naftalina no mijadouro há duas guimbas, na verdade dois tocos de filtro, um com bocal amarelo sarapintado e outro com bocal branco manchado de batom.
27	Benjamim já não se lembra do que veio fazer no banheiro e pensa "fascistas", de repente indignado com os sujeitos que atiram guimbas nos mijadouros de restaurantes sem querer saber se alguém virá recolhê-las mais tarde, um empregado que acuradamente não disporá de luvas de barbecho ou instrumento apropriado (caixa, pinça, caneta, tábua, esquadro portátil) para tal tarefa.

1	Benjamim
1	1
2	Chico Buarque
2	The execution squad was lined up; the order to fire was forceful and the shots produced a single report. But to Benjamin Zambraia they sounded like a drum roll, and he might even be able to tell the order in which the dozen weapons in front of him had discharged. Even if blind, he would identify each rifle and say from which barrel had come each of the projectiles that now tore into his chest, his neck, his face. Everything would be extinguished with the swiftness of a bullet between the skin and its first vital target (aorta, heart, trachea, medulla), and in that instant Benjamin witnessed what he had expected: his existence was projected from beginning to end, like a motion picture, on the blindfold of his eyes. Faster than a bullet, the film could be projected again on the inside of his eyelids, in reverse, when the succession of events might prove more acceptable. And there would still be a small thread of time for Benjamin to see himself once more here and there in situations that he would prefer to forget, the images ricocheting in the recesses of his skull. The allotted time would end to be succeeded by an ultimatum, a whistle, a cry of alarm, but Benjamin would interpret them as the threat of a child counting to three -one... two... two... two and a half... - and would pause a little longer on moments that belonged to him, and that he had not known how to appreciate before. He would also learn to penetrate into spaces he had never known, into times that were not his own, with the senses of other people. And suddenly he would catch himself walking simultaneously in all directions, and he would take in everything at a single glance, and all that he perceived would never cease, and infinity itself would fit inside a bubble in the interior of a dream of a man like 1, Benjamin Zambraia, who does not remember ever having died in a dream.
3	O pelotão estava em forma, a voz de comando foi enérgica e a fuzilaria produziu um único estrondo. Mas para Benjamim Zambraia soou como um ruído, e ele seria capaz de dizer em que ordem haviam disparado as doze armas ali defronte. Cego, identificaria cada fuzil e diria de que cano partira cada um dos projéteis que agora o atingiam no peito, no pescoço e na cara. Tudo se extinguiria com a velocidade de uma bala entre a epiderme e o primeiro alvo letal (aorta, coração, traquéia, bulbo), e naquele instante Benjamim assistiu ao que já esperava: sua existência projetou-se do início ao fim, tal qual um filme, na venda dos olhos. Mais rápido que uma bala, o filme poderia projetar-se uma outra vez por dentro das suas pálpebras, em marcha a ré, quando a sucessão dos fatos talvez resultasse mais aceitável. E ainda sobriava um fiapo de tempo para Benjamim rever-se aqui e acolá em situações que preferiria esquecer, as imagens ricocheteando no bojo do seu crânio. O prazo se esgotaria e sobreviria um ultimato, um apito, um alarme, mas Benjamim os entenderia como ameaça de criança contando até três - um... dois... dois... dois e meio... - e se deteria mais um pouco em momentos que lhe pertenciam, e que antes não soubera apreciar. Aprenderia também a penetrar em espaços que não conhecera, em tempos que não eram o seu, com o senso de outras pessoas. E súbito se surpreenderia a caminhar simultaneamente em todas as direções, e tudo alcançaria de um só olhar, e tudo o que ele percebesse jamais cessaria, e mesmo a infinitude caberia numa bolha no interior do sonho de um homem como Benjamim Zambraia, que não se lembra de alguma vez ter morrido em sonho.
3	If a camera were to focus on Benjamin at lunchtime, it would capture a tall, slender man, slightly stooped, with traces of an athletic past and white but abundant hair, hardly helped by a seven-day growth of beard, his shirt outside his worn trousers, giving the appearance of carelessness but not hardship, stationed in front of the Vasconcelos Bar and Restaurant, his knees tremulous as if he were waiting for someone. Benjamin, however, is not waiting for anything, except for himself to resolve his dilemma: whether to enter the bar or return to bed. The question is an awkward one because Benjamin isn't sleepy, or thirsty, or hungry, nor does he have any alternative for the afternoon. Rooted to the ground, his legs restive, he grows impatient with his own hesitation, and it is usually in circumstances such as these that the sensation of being filmed comes back to him.
4	Se uma câmera focalizasse Benjamim na hora do almoço, captaria um homem longilíneo, um pouco curvado, com vestígios de atletismo, de cabelos brancos mas bastos, prejudicado por uma barba de sete dias, camisa para fora da calça surrada aparentando desleixo e não penúria, estacionado em frente ao Bar-Restaurante Vasconcelos, tremulando os joelhos como se esperasse alguém. Benjamim entretanto não espera nada, a não ser que ele mesmo resolva o dilema: entrar num bar-restaurant ou voltar para a cama. A questão é embaraçosa porque Benjamim não tem sono, nem sede, nem apetite, nem alternativa para esta tarde. Preso ao chão, as pernas irrequietas, impacienta-se com a própria hesitação, e é nessas conjunturas que lhe costuma voltar a sensação de estar sendo filmado.
4	As an adolescent, Benjamin acquired an invisible camera because he knew that the cagiest of his schoolmates had theirs. The equipment proved to be as auspicious as a pocket comb, and from that day on his life took a new direction. Benjamin started wearing a pompadour, and in the squabbles in which earlier he would have torn his hair, sure that he was right, he now assumed a vague smile and allowed his opponent to gesture with his back to the camera. With this he gained prestige and got to kiss a lot of girls, whose shoulders, ears and ponytails were immortalized in his movies. Benjamin's film library also houses voice-overs for jazz singers, trampoline leaps, footballing feats, street fights in which he bled or came out well, and his initiation into sex with an older woman (thirty, thirty-one, thirty-three), when he nearly ruined the scene by looking at the camera. He filmed himself throughout his youth, and only with the coming of the first gray hair did he decide to put aside the ridiculous thing. It was too late: the camera had acquired a will of its own; everywhere, it would seek out the highest point to record mediocre episodes, and Benjamin felt like lifting his shirt and covering his face in the middle of the street, or lunging at the cinematographer, like criminals and famous actors. Today he is a mature man and uses indifference as a tactic to discourage filming. But when he finally goes into the Vasconcelos Bar and Restaurant, he is still bothered by the suspicion of a camera, perhaps one connected to the edge of the long-bladed fan whirling on the ceiling. Benjamin cannot be unaware that from that improbable angle the customers would turn in a circle as if sitting on a carousel and he himself would proceed in a vortex to the center of the room, perform pirouettes, give orders to the waiter as if to a satellite, and flee staggeringly to the lavatory.
5	Adolescente, Benjamim adquiriu uma câmera invisível por entender que os colegas mais astutos já possuíam as suas. O equipamento mostrou-se tão providencial quanto um pente de bolso, e a partir daquele dia a vida dele tomou novo rumo. Benjamim passou a usar topete, e nas pendengas em que antes se descabelava, certo de estar com a razão, mantinha agora um sorriso vago e deixava o adversário a gesticular de costas para a câmera. Com isso ganhou prestígio e beijou na boca muitas garotas, cujos ombros, orelhas e rabos-de-cavalo foram imortalizados em suas películas. O acervo de Benjamim também guarda dublagens de cantor de jazz, saltos de trampolim, proezas no futebol, brigas de rua em que sangrou ou se saiu bem e a sua estréia no sexo com uma senhora de idade (trinta anos, trinta e um, trinta e três), quando ele quase estranhou a cena ao olhar para a lente. Fez-se filmar durante toda a juventude, e só com o advento do primeiro cabelo branco decidiu abolir a ridícula coisa. Era tarde: a câmera criara autonomia, deu de

1	Benjamim Chico Buarque
1	Benjamin - Chico Buarque
2	O pelotão estava em forma, a voz de comando foi enérgica e a fuzilaria produziu um único estrondo.
2	The execution squad was lined up, the order to fire was forceful and the shots produced a single report.
3	Mas para Benjamin Zambraia soou como um rufo, e ele seria capaz de dizer em que ordem haviam disparado as doze armas ali defronte.
3	But to Benjamin Zambraia they sounded like a drum roll, and he might even be able to tell the order in which the dozen weapons in front of him had discharged.
4	Cego, identificaria cada fuzil e diria de que cano partira cada um dos projéteis que agora o atingiam no peito, no pescoço e na cara.
4	Even if blind, he would identify each rifle and say from which barrel had come each of the projectiles that now tore into his chest, his neck, his face.
5	Tudo se extinguiu com a velocidade de uma bala entre a epiderme e o primeiro alvo letal (aorta, coração, traquéia, bulbo), e naquele instante Benjamin assistiu ao que já esperava: sua existência projetou-se do início ao fim, tal qual um filme, na venda dos olhos.
5	Everything would be extinguished with the swiftness of a bullet between the skin and its first vital target (aorta, heart, trachea, medulla), and in that instant Benjamin witnessed what he had expected: his existence was projected from beginning to end, like a motion picture, on the blindfold of his eyes.
6	Mais rápido que uma bala, o filme poderia projetar-se uma outra vez por dentro das suas pálpebras, em marcha a ré, quando a sucessão dos fatos talvez resultasse mais aceitável.
6	Faster than a bullet, the film could be projected again on the inside of his eyelids, in reverse, when the succession of events might prove more acceptable.
7	E ainda sobraría um fiapo de tempo para Benjamin rever-se aqui e acolá em situações que preferiria esquecer, as imagens ricocheteando no bojo do seu crânio.
7	And there would still be a small thread of time for Benjamin to see himself once more here and there in situations that he would prefer to forget, the images ricocheting in the recesses of his skull.
8	O prazo se esgotaria e sobreviria um ultimato, um apito, um alarme, mas Benjamin os entenderia como ameaça de criança contando até três - um... dois... dois... dois e meio... - e se deteria mais um pouco em momentos que lhe pertenciam, e que antes não soubera apreciar.
8	The allotted time would end to be succeeded by an ultimatum, a whistle, a cry of alarm, but Benjamin would interpret them as the threat of a child counting to three -one... two... two... two and a half... - and would pause a little longer on moments that belonged to him, and that he had not known how to appreciate before.
9	Aprenderia também a penetrar em espaços que não conhecera, em tempos que não eram o seu, com o senso de outras pessoas.
9	He would also learn to penetrate into spaces he had never known, into times that were not his own, with the senses of other people.
10	E súbito se surpreenderia a caminhar simultaneamente em todas as direções, e tudo alcançaria de um só olhar, e tudo o que ele percebesse jamais cessaria, e mesmo a infinitude caberia numa bolha no interior do sonho de um homem como Benjamin Zambraia, que não se lembra de alguma vez ter morrido em sonho.
10	And suddenly he would catch himself walking simultaneously in all directions, and he would take in everything at a single glance, and all that he perceived would never cease, and infinity itself would fit inside a bubble in the interior of a dream of a man like 1, Benjamin Zambraia, who does not remember ever having died in a dream.
11	Se uma câmera focalizasse Benjamin na hora do almoço, captaria um homem longilíneo, um pouco curvado, com vestígios de atletismo, de cabelos brancos mas bastos, prejudicado por uma barba de sete dias, camisa para fora da calça surrada aparentando desleixo e não penúria, estacionado em frente ao Bar-Restaurante Vasconcelos, tremulando os joelhos como se esperasse alguém.
11	If a camera were to focus on Benjamin at lunchtime, it would capture a tall, slender man, slightly stooped, with traces of an athletic past and white but abundant hair, hardly helped by a seven-day growth of beard, his shirt outside his worn trousers, giving the appearance of carelessness but not hardship, stationed in front of the Vasconcelos Bar and Restaurant, his knees tremulous as if he were waiting for someone.
12	Benjamim entretanto não espera nada, a não ser que ele mesmo resolva o dilema: entrar num bar-restaurante ou voltar para a cama.
12	Benjamin, however, is not waiting for anything, except for himself to resolve his dilemma: whether to enter the bar or return to bed.
13	A questão é embaraçosa porque Benjamin não tem sono, nem sede, nem apetite, nem alternativa para esta tarde.
13	The question is an awkward one because Benjamin isn't sleepy, or thirsty, or hungry, nor does he have any alternative for the afternoon.
14	Preso ao chão, as pernas irrequietas, impacienta-se com a própria hesitação, e é nessas conjunturas que lhe costuma voltar a sensação de estar sendo filmado.
14	Rooted to the ground, his legs restive, he grows impatient with his own hesitation, and it is usually in circumstances such as these that the sensation of being filmed comes back to him.

# Outro alinhador

- **VisualTCA - Sentence alignment visualization**
- <http://www.nilc.icmc.usp.br/nilc/tools/pagina-visualtca/visualtca/tca.htm>

# VisualTCA - Sentence alignment visualization [\(Load example\)](#) [\(Help\)](#)

Paste the source and target texts which you want to align, and then click on the button

Align

David Malouf  
A Medium  
- 1985 -

WHEN I WAS ELEVEN I took violin lessons once a week from a Miss Katie McIntyre, always so called to distinguish her from Miss Pearl, her sister, who taught piano and accompanied us at exams.

Miss Katie had a big sunny studio in a building in the city, which was occupied below by dentists, paper suppliers and cheap photographers. It was on the fourth floor, and was approached by an old-fashioned cage lift that swayed precariously as it rose (beyond the smell of chemical fluid and an occasional whiff of gas) at the purer atmosphere Miss Katie shared with the only other occupant of the higher reaches, Miss E. Sampson, Spiritualist.

I knew about Miss Sampson from gossip I had heard among my mother's friends; and sometimes, if I was early, I would find myself riding up with her, the

David Malouf  
A Medium  
1985

AOS ONZE ANOS, eu tinha aulas de violino uma vez por semana com Miss Katie McIntyre, assim sempre chamada para não ser confundida com Miss Pearl, sua irmã, que lecionava piano e nos acompanhava nos exames . Miss Katie morava em um amplo estúdio ensolarado em um edifício da cidade, cujo andar de baixo era ocupado por dentistas, papelarias e fotógrafos baratos . Era no quarto andar, e servido por um antiquado elevador de cabine gradeada que oscilava perigosamente enquanto ascendia (além do odor de vapores químicos e, de vez em quando, uma lufada de gás) à atmosfera mais pura que Miss Katie compartilhava com a única outra ocupante das alturas, Miss E. Sampson, espiritualista . Eu conhecia Miss Sampson pelas fofocas que ouvira entre as amigas de minha mãe; e às vezes, quando eu chegava cedo, via-me subindo com ela, os dois com os pés pregados no assoalho enquanto a gaiola

# VisualTCA - Sentence alignment visualization

(Align another text) (Help)

David Malouf

A Medium

- 1985 -

WHEN I WAS ELEVEN I took violin lessons once a week from a Miss Katie McIntyre, always so called to distinguish her from Miss Pearl, her sister, who taught piano and accompanied us at exams.

Miss Katie had a big sunny studio in a building in the city, which was occupied below by dentists, paper suppliers and cheap photographers. It was on the fourth floor, and was approached by an old-fashioned cage lift that swayed precariously as it rose (beyond the smell of chemical fluid and an occasional whiff of gas) at the purer atmosphere Miss Katie shared with the only other occupant of the higher reaches, Miss E. Sampson, Spiritualist.

I knew about Miss Sampson from gossip I had heard among my mother's friends; and sometimes,

David Malouf

A Medium

1985

AOS ONZE ANOS, eu tinha aulas de violino uma vez por semana com Miss Katie McIntyre, assim sempre chamada para não ser confundida com Miss Pearl, sua irmã, que lecionava piano e nos acompanhava nos exames

. Miss Katie morava em um amplo estúdio ensolarado em um edifício da cidade, cujo andar de baixo era ocupado por dentistas, papelarias e fotógrafos baratos

. Era no quarto andar, e servido por um antiquado elevador de cabine gradeada que oscilava perigosamente enquanto ascendia (além do odor de vapores químicos e, de vez em quando, uma lufada de gás) à atmosfera mais pura que Miss Katie compartilhava com a única outra ocupante das

está se passando na sala ao lado, mas para medir os acordes - precisos, fixos, denomináveis como quintas diminutas ou sextas Napolitanas, mas também em certos momentos parecendo lágrimas - que estão sendo tocados em um piano de armário; e a voz que os nomeia como sendo seus.

Alignment results

Alignment types

- 48 alignments 1:1
- 2 alignments 1:2
- 5 alignments 2:1
- 7 alignments 0:1
- 2 alignments 1:0

(Click on a color to highlight alignments)



Source text

Paragraphs: 25  
Sentences: 62  
Words: 1692

Target text

Paragraphs: 17  
Sentences: 64  
Words: 1679

(Align another text)



named those. It wasn't difficult. It was simple mathematics and I had an ear, though the chords might also in other contexts, and in ways that were not explicable, move you to tears.

There is no story, no set of events that leads anywhere or proves anything - no middle, no end.

Just a glimpse through a half-open door, voices seen not heard, vibrations sensed through a wall while the trained ear strains, not to hear what is passing in the next room, but to measure the chords - precise, fixed, nameable as diminished fifths or Neapolitan sixths, but also at moments approaching tears - that are being struck out on an iron-framed upright; and the voice that names them your own.

7

## Junção de parágrafo

nova presença entrara na sala. Assumiu a forma de uma voz de criança, aguda e lamuriosa, e uma das mulheres deu um grito imediatamente seguido por um cochichar de outras vozes. A aguda, agora mais forte, transpassava-as. Miss Sampson oscilava como uma flor no talo. Minutos depois, atrás da porta do estúdio ensolarado de Miss Katie, depois de exhibir minhas escalas, meus arpejos, minhas três peças, fiquei de costas para o piano (de frente para a parede atrás da qual tanta emoção estava contida) enquanto Miss Katie tocava intervalos e eu dizia de quanto eram, ou tocava acordes e eu dizia o seu nome. Não era difícil. Era apenas matemática e eu tinha bom ouvido, embora os acordes também pudessem, em outros contextos, e em formas inexplicáveis, levar às lágrimas. Não há história, nem conjunto de acontecimentos que leve a algum lugar ou prove alguma coisa - nenhum meio, nenhum fim. Apenas um vislumbre através de uma porta semi-aberta, vozes vistas mas não ouvidas, vibrações sentidas através de uma parede enquanto o ouvido treinado se esforça, não para ouvir o que está se passando na sala ao lado, mas para medir os acordes - precisos, fixos, denomináveis como quintas diminutas ou sextas Napolitanas, mas também em certos momentos parecendo lágrimas -

eleven-year-old with scarred knees clutching at Mozart. Like most boys of that age I had much to conceal.

But she appeared to be looking at me, not through me. She smiled, I responded, and clearing my throat to find a voice, would say in a well-brought-up, Little Lord Fauntleroy manner that I hoped might fool her and leave me alone with my secrets, 'Good afternoon, Miss Sampson.'

Her own voice was as unremarkable as an aunt's: 'Good afternoon, dear.'

All the more alarming then, as I sat waiting on one of the cane-bottomed chairs in the corridor, while Ben Steinberg, Miss McIntyre's star pupil, played the Max Bruch, to hear the same voice oddly transmuted. Resonating above the slight swishing and breathing of her congregation, all those women in gloves, hats, fur-pieces, packed in among ghostly pampas-grass, it had stepped down a tone - no, several - and came from another continent. I felt a shiver go up my spine. It was the Indian, speaking through her out of another existence.

Standing at an angle to the half-open door, I caught only a segment of the scene. In the glow of

era, eu achava, apenas o peso dos corpos ("capacidade máxima: oito pessoas", dizia o aviso) que fazia o velho mecanismo ranger em seu poço, mas sim o peso de toda aquela aflição, toda aquela desesperança e derradeiras esperanças, toda aquela dignidade na privacidade da dor e o silêncio quebrado apenas por um fortuito "Não fique triste, meu bem", ou um sussurrado, "George ia gostar, tenho certeza". Subíamos lentamente. Eu achava preferível, de modo geral, chegar cedo e subir rápido, e em silêncio, com a própria Miss Sampson. Às vezes, por vã curiosidade (se é que tal motivo lhe podia ser atribuído) ela deixava que seus olhos pousassem por um instante sobre mim, e eu, ansioso, me perguntava o que ela estaria vendo além de um menino gorducho de onze anos e joelhos com cicatrizes agarrado a Mozart. Como a maioria dos meninos de minha idade, eu tinha muito a esconder. No entanto, ela parecia estar olhando para mim, não através de mim. Sorria, eu retribuía e, pigarreando para encontrar minha voz, dizia, em um estilo tão bem-educado quanto o do Pequeno Lorde, esperando enganá-la e ficar a sós com meus segredos, "Boa tarde, Miss Sampson". A voz dela, por sua vez, era tão comum quanto a de uma tia: "Boa tarde, querido". Ainda mais alarmante - quando eu estava sentado esperando

# Corpora na Tradução

## Vantagens dos corpora:

- podem ser construídos de acordo com as necessidades do tradutor
- podem ser constantemente atualizados
- conferem segurança ao tradutor na escolha do termo a empregar
- fornecem exemplos autênticos de uso →  
**naturalidade para o texto traduzido**

# Mão na massa!!!

**Stella E. O. Tagnin**

**Estudos em Tradução – Teorias,  
Práticas e Tecnologias**

**PUC-RS / 10-11/10/2014**

3ª. Parte

---

# Tutorial do BootCat

[http://comet.fflch.usp.br/sites/comet.fflch.usp.br/files/u30/Guia%20BootCat%20atual\\_out2014%20%281%29\\_0.pdf](http://comet.fflch.usp.br/sites/comet.fflch.usp.br/files/u30/Guia%20BootCat%20atual_out2014%20%281%29_0.pdf)

# Guia prático para a ferramenta BootCat

Atualizado em 06/10/2014



# BootCaT

By Joacyr T. de Oliveira

# O que você precisa para usar o BootCat?

- 1. Baixar o BootCat.
- 2. Durante a instalação do BootCat, permitir também a instalação do Strawberry Pearl.
- 3. Usar uma conta de e-mail do *Outlook.com* para gerar um número que permite usar a ferramenta de busca Bing.
- (Se você não tem uma conta Outlook, acesse o site [www.outlook.com](http://www.outlook.com) e crie seu e-mail. Além de ter acesso a uma variedade de aplicativos, você ainda terá 10Gb de armazenamento nas nuvens grátis no Onelook e acesso, também grátis, ao

*Pacote Microsoft Office – Word, Excel, Power Point*



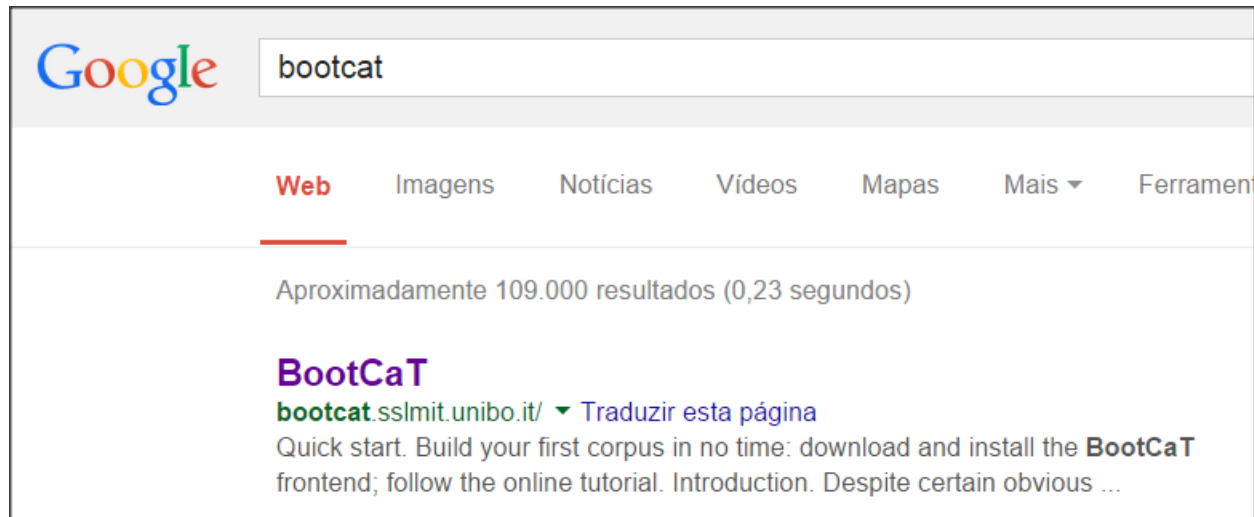
Simple Utilities to Bootstrap Corpora And Terms from the Web

**BootCaT é uma interface que serve como guia na criação de corpora simples a partir da web.**

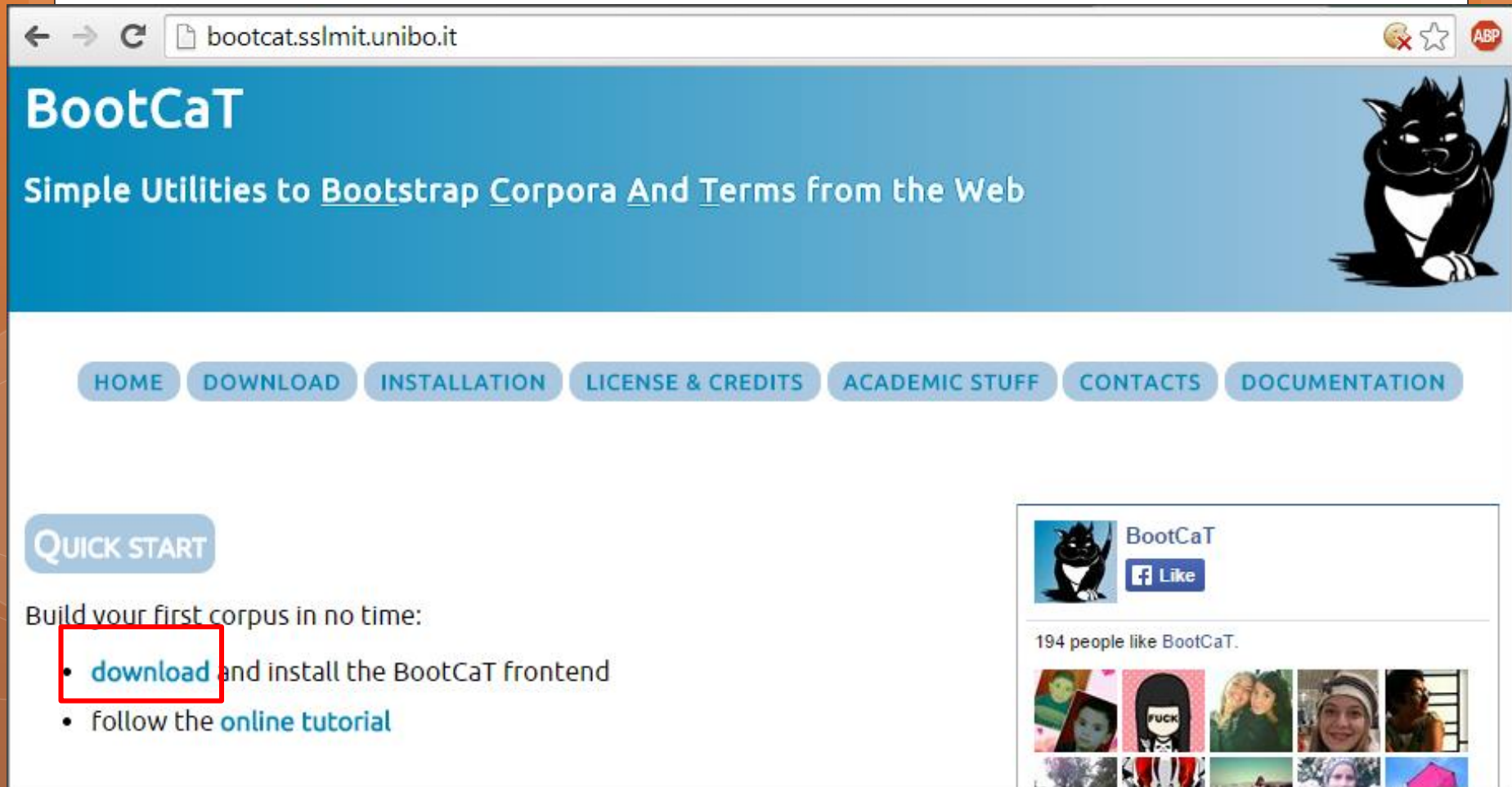


# Encontrar o programa BootCat

- No Google, digite “bootcat”. Clique no link que tenha o domínio unibo.it.



# Site BootCaT



The screenshot shows a web browser window with the address bar containing "bootcat.sslmit.unibo.it". The page has a blue header with the title "BootCaT" and the subtitle "Simple Utilities to Bootstrap Corpora And Terms from the Web". A cartoon cat logo is in the top right. Below the header is a navigation menu with buttons for "HOME", "DOWNLOAD", "INSTALLATION", "LICENSE & CREDITS", "ACADEMIC STUFF", "CONTACTS", and "DOCUMENTATION". A "QUICK START" section is visible, containing the text "Build your first corpus in no time:" followed by a list of instructions. The word "download" in the first instruction is highlighted with a red box. On the right side, there is a Facebook social media widget for "BootCaT" showing a "Like" button and a notification that "194 people like BootCaT." with a grid of profile pictures.

bootcat.sslmit.unibo.it

## BootCaT

Simple Utilities to Bootstrap Corpora And Terms from the Web

HOME DOWNLOAD INSTALLATION LICENSE & CREDITS ACADEMIC STUFF CONTACTS DOCUMENTATION

### QUICK START

Build your first corpus in no time:

- **download** and install the BootCaT frontend
- follow the [online tutorial](#)

BootCaT  
Like

194 people like BootCaT.

# Versão 0.71

LATEST RELEASE (VERSION 0.71 — SEPTEMBER 10, 2014)

See the [release notes](#) to find out what's new in version 0.71

- [BootCaT front-end \(graphical interface\)](#)  
(includes the latest version of the toolkit)

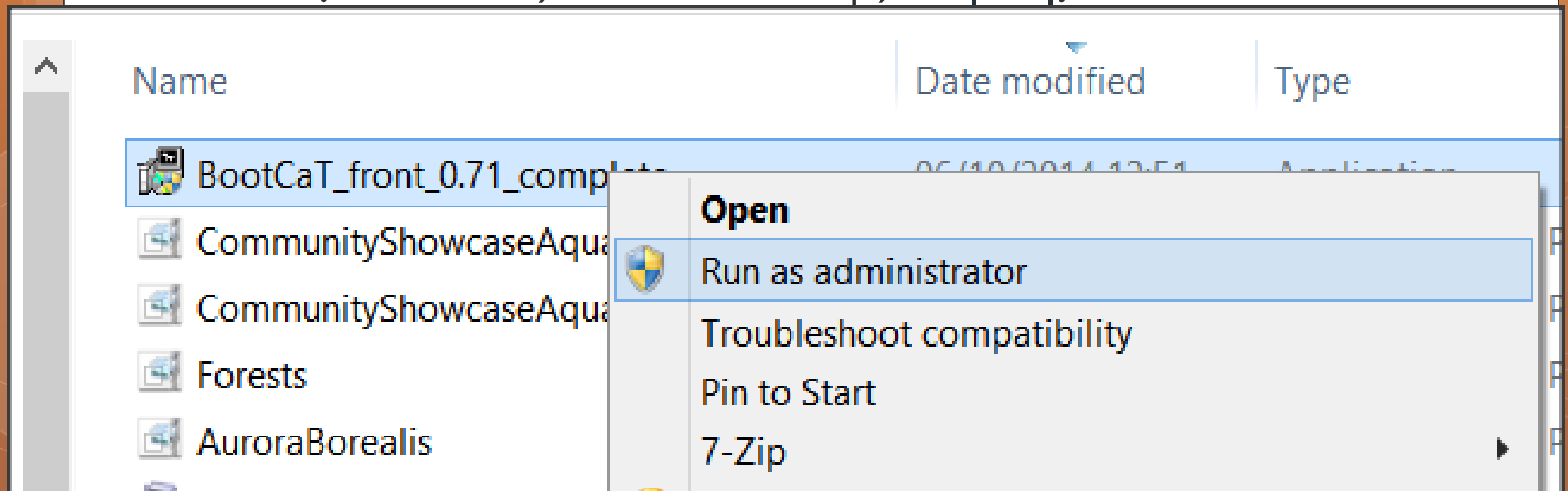


[Upgrade package for Windows](#) (use this if you're upgrading from previous versions)

[Complete installation for Windows](#) (includes BootCaT and Strawberry Perl)

# Instalação

- No Windows 8, você deve instalar como administrador. Assim que aparecer o

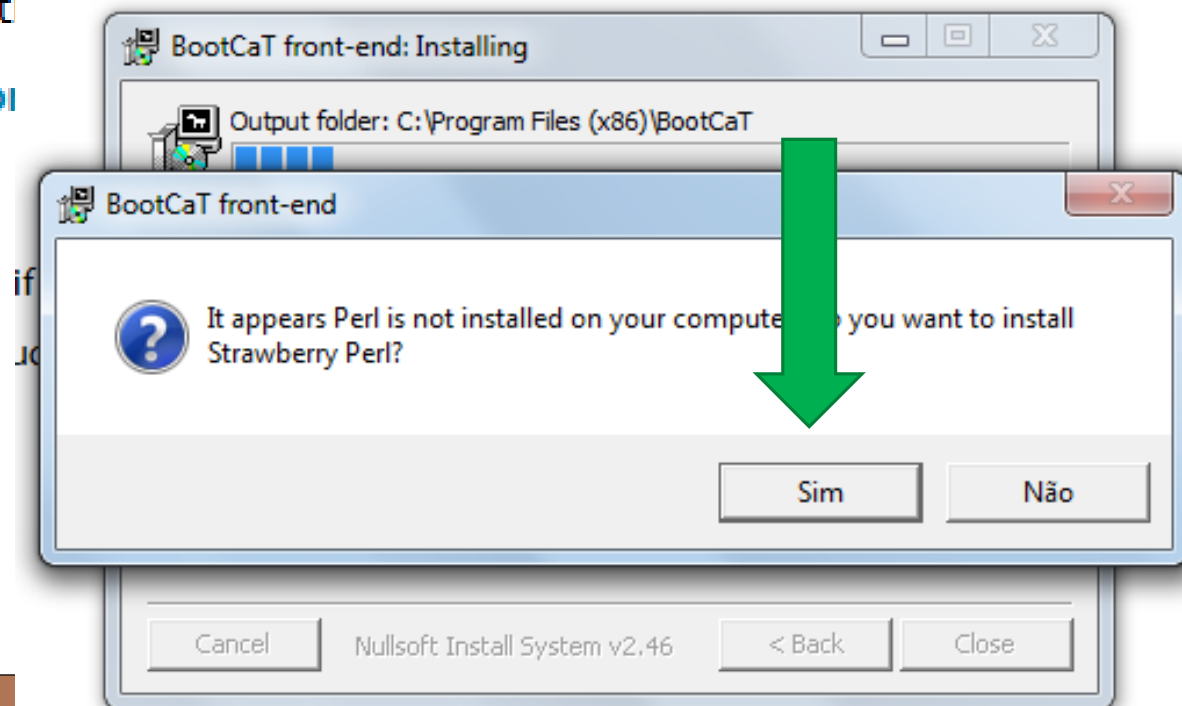


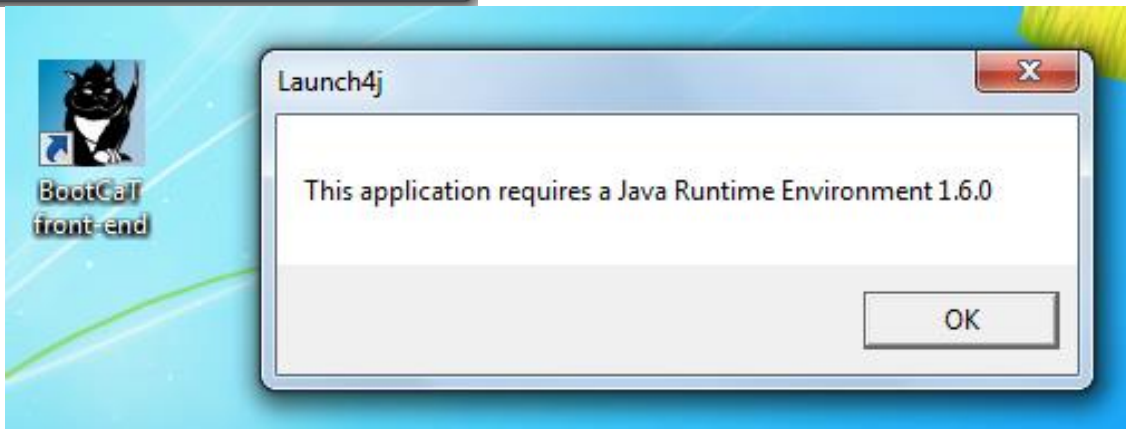
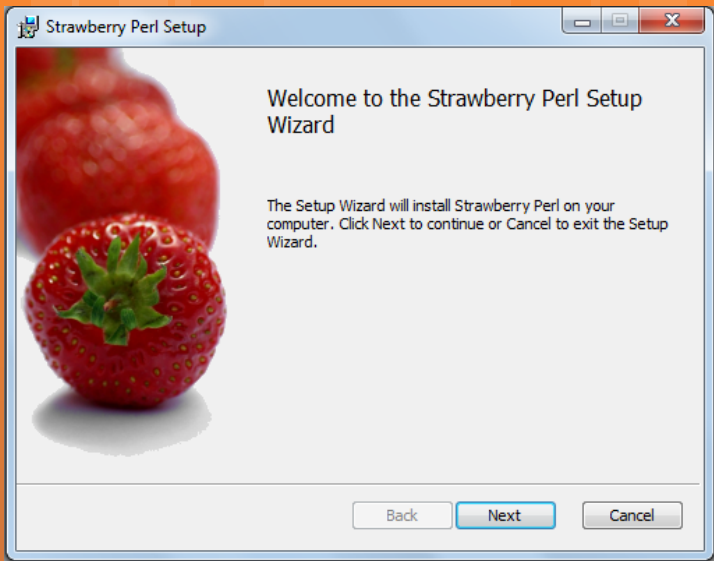
<http://bootcat.sslmit.unibo.it/>  
ma.

## QUICK START

Build your first corpus in no time:

- **download** and install the software
- follow the **online tutorial**





Depois de instalar o Strawberry, talvez você tenha que instalar uma versão atualizada do Java. Não se preocupe; o instalador do BootCat fará isso. Apenas siga as instruções.

### Download gratuito do Java

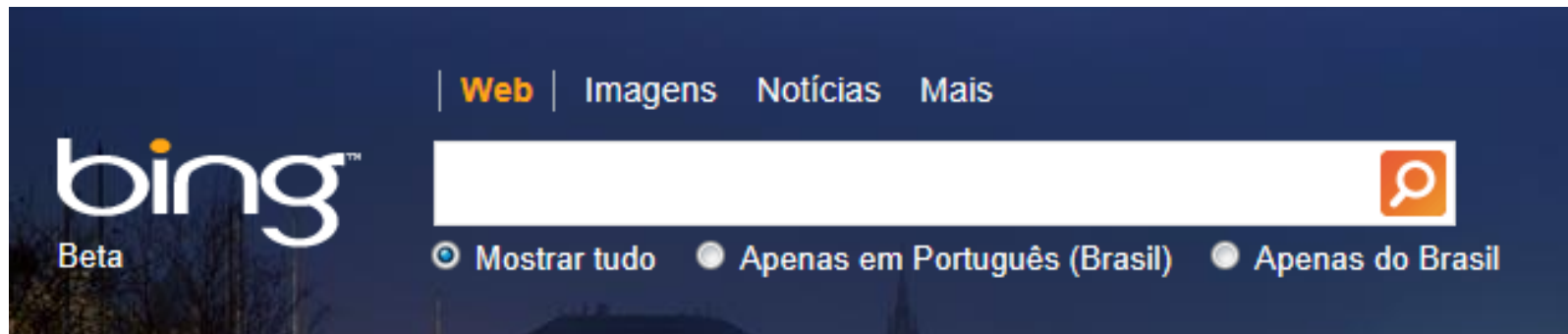
Faça o download do Java para seu computador agora!

Version 6 Update 31

[Download gratuito do Java](#)

[» O que é o Java?](#) [» Tenho o Java?](#) [» Precisa de ajuda?](#)

- Bootcat usa a ferramenta de busca Bing para selecionar URLs da Internet.



- Bing = ferramenta de busca operada pela Microsoft



- Já estamos “quase” prontos.



Clique no ícone para abrir o programa.

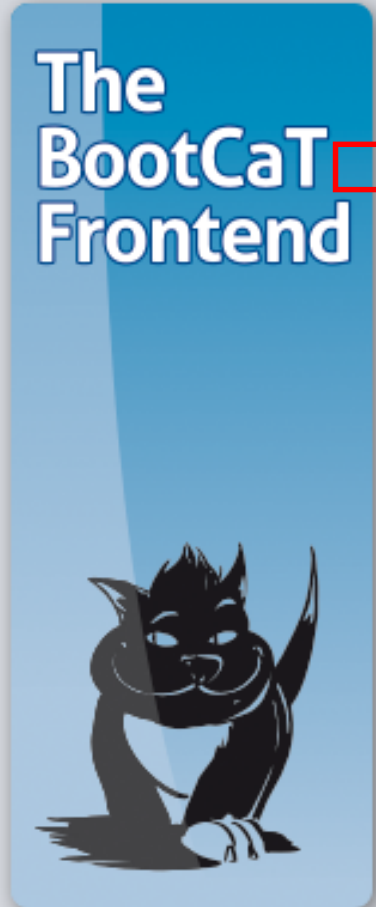




BootCaT frontend



File Edit Help



Project definition

Corpus name

Language

More options

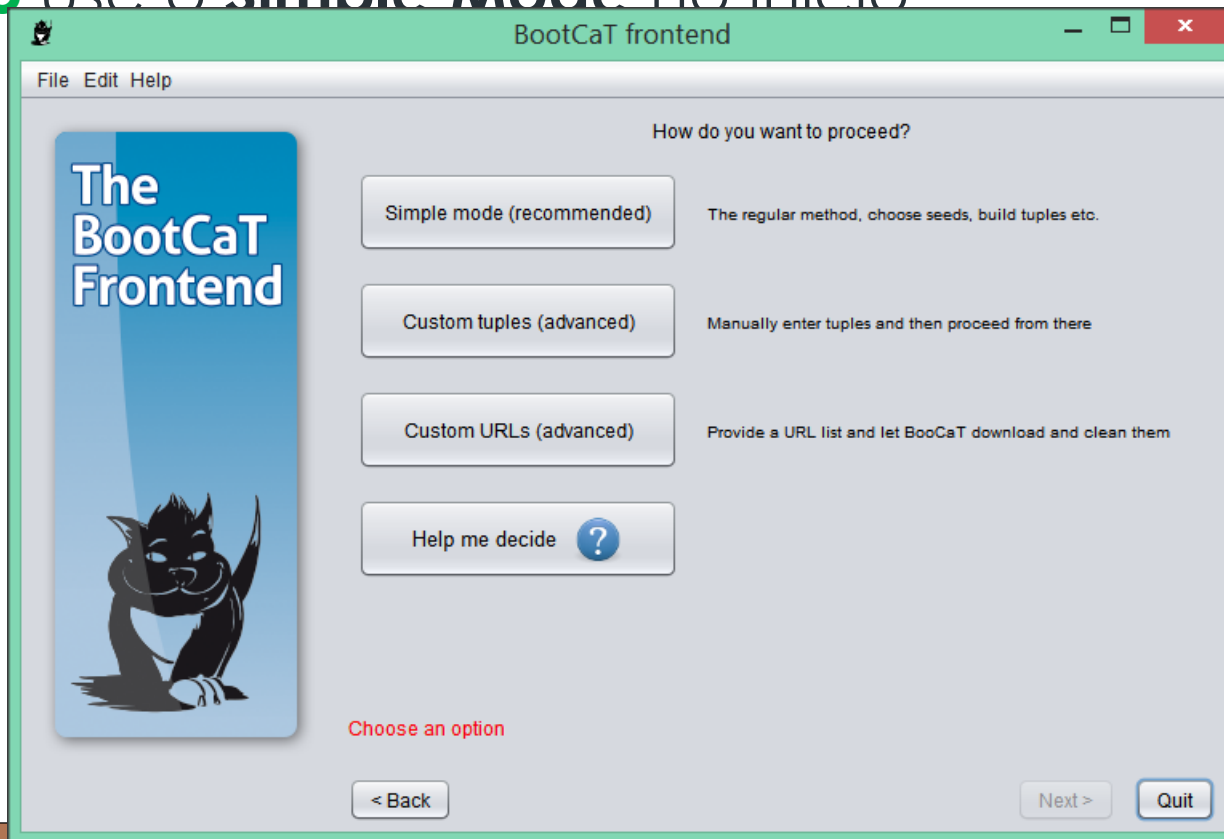
< Back

Next >

Quit

# Para começar a criar o corpus

- Use o **Simple Mode** no início





Insert one seed per line, multi-word seeds go on the same line

# The BootCaT Frontend



hair straightening  
iron  
flat iron  
blow dry  
hair  
curly  
kinky  
nappy  
comb  
relaxer  
straightener

O BootCat usará combinações de *seeds* para buscar textos na web. Use palavras relevantes para o tema do *corpus*. Assim que terminar, clique em *I'm done editing seeds*.

I'm done editing seeds



Type seeds and check the relevant box when you're done

< Back

Next >

Quit



# The BootCaT Frontend



## The tuples that will be used as queries

- "blow dry" "flat iron" iron
- comb relaxer "blow dry"
- "blow dry" curly straightener
- relaxer nappy kinky
- relaxer "hair straightening" hair
- hair "flat iron" nappy
- straightener iron "blow dry"
- "flat iron" "blow dry" nappy
- hair kinky iron
- "flat iron" straightener curly

As *tuples* são combinações de *seeds* que o BootCat usará para buscar textos na web.

Generate tuples

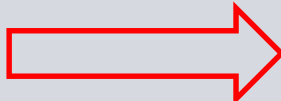
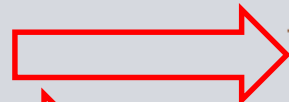
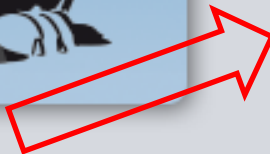
Tuple length

N. of tuples (max: 165)

< Back

Next >

Quit



# The BootCaT Frontend



Your Bing AppId

Remember AppId on this computer

You are about to query a search engine using the tuples you just created. In order to do this, you must provide a valid Bing AppId (i.e. a sort of password).

If you do not have a Bing AppId, click on the button below for instructions on how to obtain one.

Ao clicar, abre-se um tutorial para criar o Bing App ID.



[Click here if you don't have an AppId](#)

You must provide a search engine key

< Back

Next >

Quit

- O BootCat precisa de um AppID para acessar o Bing.

- Veja os passos para criar o AppID no Bing gratuitamente:

- 1. **“Windows Live ID”** (i.e. uma “velha” conta de e-mail **Hotmail** ou uma “nova” do **Outlook**). Se você não possui uma conta Outlook, faça uma aqui: [www.outlook.com](http://www.outlook.com) .

- 2. Entre no site: [datamarket.azure.com](http://datamarket.azure.com)
- Clique em “Sign In.”

Windows Azure Marketplace

Hello, Joacyr Oliveira | English | Brazil | Support | Sign Out

Learn Applications **Data** My Account Publish

Search the Marketplace

HOME > DATA

price

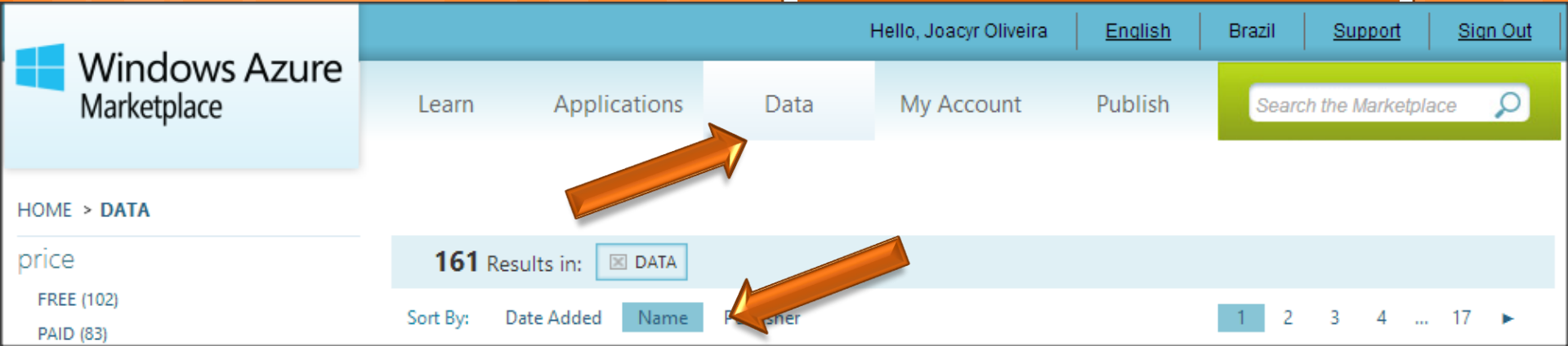
FREE (102)

PAID (83)

161 Results in: DATA

Sort By: Date Added Name Publisher

1 2 3 4 ... 17 ▶



## Bing Search API

published by: Microsoft

The Bing Search API enables developers to embed search results in applications or websites using XML or JSON. Access web, image, news and video results as well as related searches and spelling suggestions.



## Bing Search API

Data



Published by: Microsoft  
Categories: Reference, Developer Services  
Date added: 11/04/2012

[Get support for this offering](#)

[Request higher query volume](#)

### What is the Bing Search API?

The Bing Search API enables developers to embed search results in

5,000  
Transactions/month

\$0,00  
per month

[SIGN UP](#)

10,000  
Transactions/month

\$19,93  
per month

[BUY](#)

20,000  
Transactions/month

\$39,85  
per month

[BUY](#)

50,000

\$99.62





# Sign Up



**BING SEARCH API**

**Published By Microsoft**

[View Publisher Offer Terms](#)



Print

**Your Selection:**

Subscription with 5,000 Transactions

**Price:**

Free, starting on 20/03/2014.

You can cancel your service at any time by visiting <https://datamarket.azure.com/account/datasets>.



Offers are licensed to you by the Publisher under Publisher's Offer Terms. Microsoft provides no rights or licenses for third party Publisher Offers.

I have read and agree to the above publisher's Offer Terms and Privacy Policy.

[← Cancel](#)



**SIGN UP**

# Account Keys

Account keys are used by applications to access your Windows Azure Marketplace dataset subscriptions. Please make sure to secure your account key as you would a credit card number. Do not share your account keys with other users.

Example account key usage: Excel, PowerPivot or third-party applications may require that you provide your account key in order to access your Windows Azure Marketplace dataset subscriptions.

Description	Value	
	eFVSs[REDACTED]iW9M=	Edit



ADD

• Você não precisa criar um novo AppID sempre que instalar o BootCaT, pois o mesmo código pode ser usado em diversos computadores.

• Use os comandos CTRL-C (copiar) e CTRL-V (colar) na hora de inserir o seu código no BootCaT.

File Edit Help

# The BootCaT Frontend



Your Bing Appld

 Remember Appld on this computer

You are about to query a search engine using the tuples you just created. In order to do this, you must provide a valid Bing Appld (i.e. a sort of password).

If you do not have a Bing Appld, click on the button below for instructions on how to obtain one.

[Click here if you don't have an Appld](#)

You must provide a search engine key

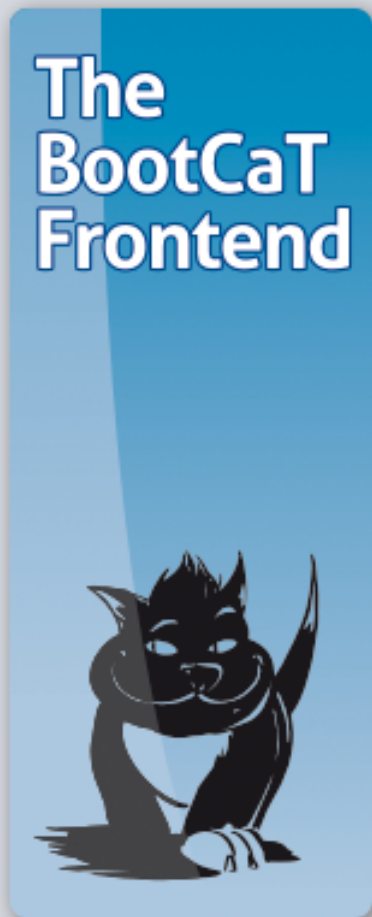
&lt; Back

Next &gt;

Quit



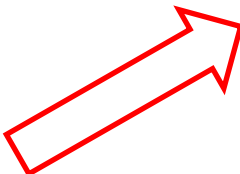
File Edit Help



Limit search to the following Internet domain (e.g. .edu):

Exclude the following Internet domains (e.g. .com, wikipedia.org):

- answers.yahoo.com
- amazon.com
- ebay.com
- youtube.com



Adult filter (filter sexually explicit material)

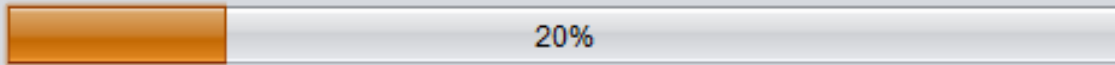
Moderate ▾

Maximum number of URLs to return for each tuple

10 ▴ ▾

URLs returned by search engine

- CURRENT\_QUERY "hair straightening" iron "blow dry"
- <http://www.webmd.com/beauty/hair-styling/keratin-hair-straightening-treatments>
- <http://www.ehow.com/fashion/hair-care/hair-straightening/>
- <http://www.buzzle.com/articles/best-hair-straightening-spray.html>
- <http://www.besthairstraighteningproducts.com/index.html>
- [http://www.how-to-div.org/AEWITFiX02mB011/How-I-Straighten-My-Hair-\(Blow-Dry-&-Flat-Iron\)-Relaxed-Hair.html](http://www.how-to-div.org/AEWITFiX02mB011/How-I-Straighten-My-Hair-(Blow-Dry-&-Flat-Iron)-Relaxed-Hair.html)



Collect URLs

Retrieving URLs, please wait

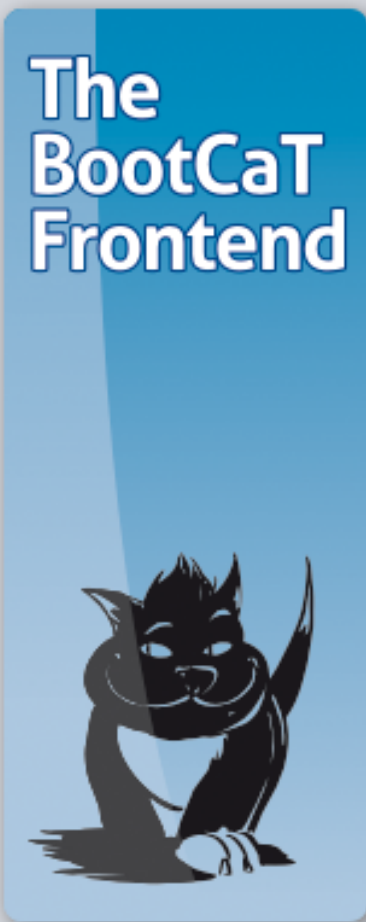
< Back

Next >

Quit



Here you can verify and remove individual URLs from the list



- <http://beauty.about.com/od/curlyhair/a/curlyhairstyle.htm>
- <http://beauty.about.com/od/stylingyourhair/tp/dryerflatiron.htm>
- <http://bestflatironfornaturalhair.net>
- <http://blackhair.about.com/b/2013/05/15/pressing-comb-or-flat-iron.htm>
- <http://blackhair.about.com/od/basiccare/f/Should-I-Blow-Dry-My-Hair-Before-I-Flat-Iron-It.htm>
- <http://blackhair.about.com/od/toolsaccessories/tp/flatirontips.htm>
- <http://blog.sellstraightener.com/chi-flat-irons/curly-hair-how-to-comb-a-nice-hair-2.html>
- <http://ezinearticles.com/?Straightening-Process-For-Naturally-Kinky-Hair&id=4399970>
- [http://forum.blackhairmedia.com/nappy-after-relaxer\\_topic7619.html](http://forum.blackhairmedia.com/nappy-after-relaxer_topic7619.html)
- [http://forum.blackhairmedia.com/nappy-looking-hair-after-relaxer\\_topic10542.html](http://forum.blackhairmedia.com/nappy-looking-hair-after-relaxer_topic10542.html)
- <http://hair-styling.wonderhowto.com/how-to/blow-dry-flat-iron-protect-long-hair-332271>
- <http://nappy-hair-jhfdt.blogspot.com/2010/06/hair-extensions-and-relaxers.html>
- [http://nappy-hair1.blogspot.com/2009/07/nappy-hairhow-do-i-straighten-it\\_30.html](http://nappy-hair1.blogspot.com/2009/07/nappy-hairhow-do-i-straighten-it_30.html)
- <http://nappyhairproducts.com/?p=509>
- <http://nappyhairproducts.com/?tag=natural-hair-to-relaxer>
- <http://nappyme.wordpress.com/about-me/relaxer-disaster-my-first-perm>
- <http://naturalicioushair.com/2011/11/26/natural-hair-straightening-products-necessary-to-straighten.html>
- <http://newhairstyles2014.net/curly/curly-hair-straightening-with-blow-dryer.html>
- <http://revivehairspa.com/salon-services>

Selected URLs: 91

Open URL in browser

Retrieved URLs: 93

< Back

Next >

Quit

# WHAT'S BETTER FOR YOUR HAIR? FLAT IRON VS. HOT COMB

22 COMMENTS

April 19, 2012 - By Kimberly Shorter



<http://blazeadams247.blogspot.com>

My journey to the promised land of natural hair has been a path paved with many trials and errors, shampoos, conditioners, creams and oils. As I am learning to embrace and celebrate my natural hair in all its coil-y glory, I still like to wear my hair in a sleek, straight bob every now and then.

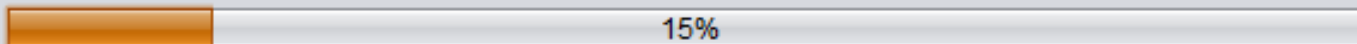


File Edit Help

# The BootCaT Frontend



Downloading <http://beauty.about.com/od/curlyhair/a/curlyhairstyle.htm>  
Downloading <http://beauty.about.com/od/stylingyourhair/tp/dryerflatiron.htm>  
Downloading <http://bestflatironfornaturalhair.net>  
Downloading <http://blackhair.about.com/b/2013/05/15/pressing-comb-or-flat-iron.htm>  
Downloading <http://blackhair.about.com/od/basiccare/f/Should-I-Blow-Dry-My-Hair-Before-I-Flat-Iron-It.htm>  
Downloading <http://blackhair.about.com/od/toolsaccessories/tp/flatirontips.htm>  
Downloading <http://blog.sellstraightener.com/chi-flat-irons/curly-hair-how-to-comb-a-nice-hair-2.html>  
Downloading <http://ezinearticles.com/?Straightening-Process-For-Naturally-Kinky-Hair&id=4399970>  
Downloading [http://forum.blackhairmedia.com/nappy-after-relaxer\\_topic7619.html](http://forum.blackhairmedia.com/nappy-after-relaxer_topic7619.html)  
Downloading [http://forum.blackhairmedia.com/nappy-looking-hair-after-relaxer\\_topic10542.html](http://forum.blackhairmedia.com/nappy-looking-hair-after-relaxer_topic10542.html)  
Downloading <http://hair-styling.wonderhowto.com/how-to/blow-dry-flat-iron-protect-long-hair-332271>  
Downloading <http://nappyhairproducts.com/?p=509>  
Downloading <http://nappyhairproducts.com/?tag=natural-hair-to-relaxer>  
Downloading <http://nappyme.wordpress.com/about-me/relaxer-disaster-my-first-perm>



Build corpus

Open corpus folder

Building corpus, please wait

< Back

Quit

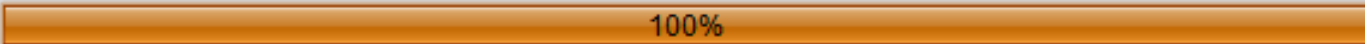


File Edit Help

# The BootCaT Frontend



Downloading <http://beauty.about.com/od/curlyhair/a/curlyhairstyle.htm>  
Downloading <http://beauty.about.com/od/stylingyourhair/tp/dryerflatiron.htm>  
Downloading <http://bestflatironfornaturalhair.net>  
Downloading <http://blackhair.about.com/b/2013/05/15/pressing-comb-or-flat-iron.htm>  
Downloading <http://blackhair.about.com/od/basiccare/f/Should-I-Blow-Dry-My-Hair-Before-I-Flat-Iron-It.htm>  
Downloading <http://blackhair.about.com/od/toolsaccessories/tp/flatirontips.htm>  
Downloading <http://blog.sellstraightener.com/chi-flat-irons/curly-hair-how-to-comb-a-nice-hair-2.html>  
Downloading <http://ezinearticles.com/?Straightening-Process-For-Naturally-Kinky-Hair&id=4399970>  
Downloading [http://forum.blackhairmedia.com/nappy-after-relaxer\\_topic7619.html](http://forum.blackhairmedia.com/nappy-after-relaxer_topic7619.html)  
Downloading [http://forum.blackhairmedia.com/nappy-looking-hair-after-relaxer\\_topic10542.html](http://forum.blackhairmedia.com/nappy-looking-hair-after-relaxer_topic10542.html)  
Downloading <http://hair-styling.wonderhowto.com/how-to/blow-dry-flat-iron-protect-long-hair-332271>  
Downloading <http://nappyhairproducts.com/?p=509>  
Downloading <http://nappyhairproducts.com/?tag=natural-hair-to-relaxer>  
Downloading <http://nappyme.wordpress.com/about-me/relaxer-disaster-my-first-perm>  
Downloading <http://naturalicioushair.com/2011/11/26/natural-hair-straightening-products-necessary-to-straightening-y>  
Downloading <http://newhairstyles2014.net/curly/curly-hair-straightening-with-blow-dryer.html>  
Downloading <http://revivehairsa.com/salon-services>  
Downloading <http://straighteningnaturalhair.org/how-to-straighten-curly-hair>  
Downloading <http://sultra.com/styling-tips-diva-straight.asp>  
Downloading <http://voices.yahoo.com/best-hair-straightening-tips-dry-frizzy-hair-with-5880458.html>



Build corpus

Open corpus folder

Corpus saved in C:\Users\Joa\Documents\BootCaT Corpora\Hair Straightening

< Back

Finish



Navigation icons: back, forward, up, down, and a folder icon. Address bar: << BootCaT Corpora >> Hair Straightening

Documents

- Any Video Converter
- Apowersoft Screen Recorder
- Arquivos do Outlook
- Blocos de Anotações do One
- BootCaT Corpora
  - Cats
  - Hair Straightening
  - Rent
  - Splitter

Name

- corpus.txt
- seeds.txt
- tuples.txt
- url\_list\_cleaned.txt
- url\_list\_collected.txt
- url\_list\_edited.txt

any hair texture, these ceramic and tourmaline flat irons come in blue or pink, making them perfect gifts for teens and adults alike. Buy it online . Buy Direct Remington S-9951 Frizz Therapy, \$49.99 Created for frizzy hair, this new ceramic flatiron works like a regular flatiron with a catch. The plates of the flatiron have a porous ceramic coating that actually emits a silicon-based conditioner. According to Allure Magazine, tests on the flatiron showed it decreased frizziness 65 percent more effectively than an average flatiron. Hair was kept straight for up to 15 hours. What's more, the fluid in the iron lasts online. It's a pretty good deal. Compare prices on Amazon.com. More than 4,700 reviews on Amazon.com (the most reviewed and best loved flat iron sometimes, too), this flat iron is popular for frizzy hair, it has a swiveling cord so you won't get tangled. It has multiple settings in the control and it's a great bang. You can also curl and flip hair with this iron. Compare prices online . Ceramic Tools CT2555 1" Ceramic Flat Iron, \$30 So you don't want to pay for the extra-special professional flatirons? No problem. A favorite of reviewers on Amazon.com, this iron costs less than \$35. **CURRENT** URL <http://bestflatironfornaturalhair.net>\*Click on any of the columns to sort the data to help make your decision easier. # How we did it? We researched the various available flat hair irons which have been popular in 2013 2014 to find the best flat iron for natural hair. We selected those hair straightener brands and models where at least

## Find

Find what: CURRENT

Find Next

Direction

 Up  Down

Cancel

 Match case

corpus.txt - Notepad

File Edit Format View Help

actually emits a silicon-based conditioner. According to Allure Magazine, tests on the flatiron showed it decreased frizziness 65 percent more effectively than an average flatiron. Hair was kept straight for up to 15 hours. What's more, the fluid in the iron lasts 4 years. At under \$100, it's a great deal. Compare prices online . HSI The Styling Iron has 4,700 reviews on Amazon.com and a consistent 4.5 star rating. This flat iron is popular out there. At under \$100, it's a great deal. Compare prices online . Ceramic Tools CT2555 1" Ceramic Flat Iron, \$30 So you don't want to pay for the extra-special because it comes with a swiveling cord so you won't get tangled and the results last a long time. You can also curl and flip hair with this iron. Compare prices online . Ceramic Tools CT2555 1" Ceramic Flat Iron, \$30 So you don't want to pay for the extra-special professional flatirons? No problem. A favorite of reviewers on Amazon.com, this iron costs less than \$35.

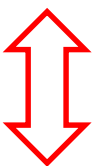
**CURRENT** URL <http://bestflatironfornaturalhair.net>\*Click on any of the columns to sort the data to help make your decision easier. # How we did it? We researched the various available flat hair irons which have been popular in 2013 2014 to find the best flat iron for natural

Find

Find what: CURRENT Find Next Cancel

Direction  
 Up  Down

Match case



# Para separar os textos

- Usar o Text File Splitter.  
(Gratuito)
- Ou usar a ferramenta *Split Text* do WordSmith Tools.

- Para salvar uma cópia dessa apresentação, entre no site do Projeto COMET e clique em Publicações. Use o QR Code abaixo para acesso imediato em seu celular ou tablet.



The image shows a screenshot of a web browser displaying the website for the COMET project. The browser's address bar shows the URL `comet.fflch.usp.br/projeto`. The website has a yellow and orange header with the COMET logo, which features a computer mouse and a bar chart. Below the logo is a navigation menu with buttons for 'Projeto', 'Equipe', 'CorTec', 'CorTrad', 'Publicações', and 'Corpora Online'. A large QR code is positioned on the right side of the page. The main content area is titled 'Projeto' and includes a 'Ver' button and an 'Editar' button. Below this, the text reads 'O PROJETO' and describes the project as a multilingual corpus for teaching and translation, developed at the University of São Paulo. It lists three sub-corpora: CorTec, CorTrad, and CoMAprend.

comet.fflch.usp.br/projeto

Conteúdo Estrutura Aparência Configurações Ajuda Buscar 0 / 1 7303866 Sair

COMET

Projeto Equipe CorTec CorTrad Publicações Corpora Online

Início

## Projeto

Ver Editar

### O PROJETO

O CoMET – *Corpus* Multilíngue para Ensino e Tradução, desenvolvido junto ao Departamento de Letras Modernas Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, é um *corpus* eletrônico que tem por objetivo ser pesquisas linguísticas, principalmente nas áreas de Tradução, Terminologia e Ensino e Aprendizagem de Línguas. O CoMI três subcorpora:

- *Corpus* Técnico-Científico: **CorTec**
- *Corpus* de Tradução: **CorTrad**
- *Corpus* Multilíngue de Aprendizizes: **CoMAprend**

Meus agradecimentos a

*Joacyr Oliveira*

[joacyr@outlook.com](mailto:joacyr@outlook.com)

# Exercícios

1. Escolham um domínio para construir um corpus.
2. Seleccionem 2 ou 3 textos desse domínio.
3. Façam uma WordList desses textos e identifiquem as primeiras palavras de conteúdo.
4. Usem essas palavras para encontrar mais textos, totalizando 10 textos – se for coleta manual -, e 30 textos se for pelo BootCat.
5. Criem um cabeçalho de acordo com suas necessidades de pesquisa.

# Exercícios

6. Façam uma WordList dos seu corpus.
7. Façam uma KeyWord List usando a WL do corpus de referência fornecido.
8. Identifiquem as colocações das 3 ou 5 primeiras palavras-chave.
9. Façam um corpus comparável na sua língua-alvo.
10. Identifiquem os equivalentes (e colocações) das palavras que vocês investigaram.



Obrigada

*Stella*

seotagni@usp.br