

NooJ and Corpógrafo

A New Partnership

Sérgio Matos & Belinda Maia

Linguatca and Faculdade de Letras da Universidade do Porto, PT

June 2008

- Corpógrafo - Linguateca**
- Gestor
 - Pesquisa
 - Centro de Conhecimento
 - BD Terminológicas
 - Gestor de Relações
 - BD Fraseológicas
 - Centro de Comunicação

Informações

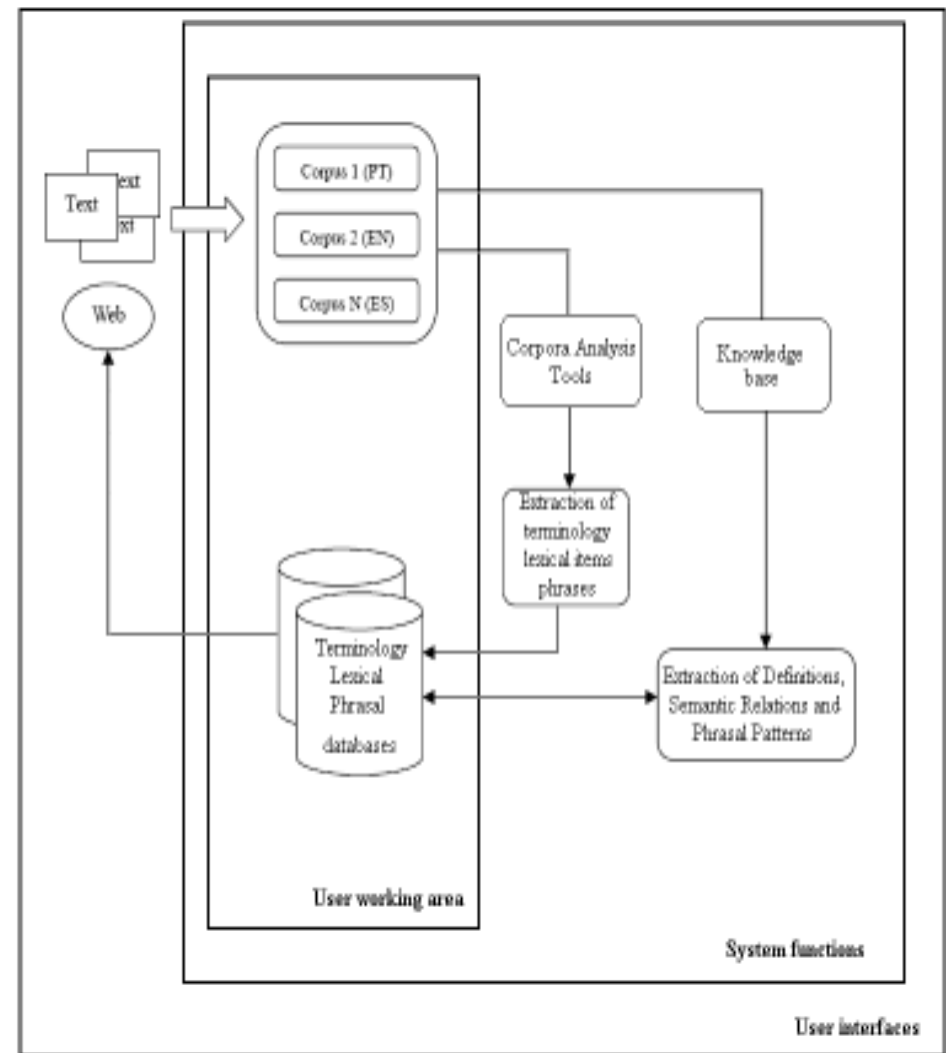
V4

Bem-vindo ao Corpógrafo V4!

www.linguateca.pt/corpografo

Corpógrafo

- ◆ General functionalities
 - ◆ Text manipulation
 - ◆ Importing
 - ◆ Cleaning texts
 - ◆ Corpora analysis
 - ◆ Concordances
 - ◆ N-grams
 - ◆ Terminology
 - ◆ Databases
 - ◆ Extraction of terms
 - ◆ Extraction of relations and definitions



Corpógrafo

- Concordances
 - Using regular expressions (e.g. “(the|a) \w+ is \w+”)

Corpógrafo - Linguateca

- Gestor
- Pesquisa
 - Concordância Frase**
 - Estudo de N-Gramas
 - Concordância Janela
 - Concordância KWIC
 - Concordâncias paralelas
- Centro de Conhecimento
- Centro de Comunicação

Concordâncias

- * [Voltar ao menu](#)
- * [Analisar Frequências](#)
- * [Contexto Janela 1-1](#)
- * [Contexto Janela 2-2](#)

Imprimir!

Informações

 Expressão de pesquisa: **neurons**

Corpus pesquisado: Neurons (29181 átomos : 37 ficheiros)

Nº de concordâncias obtidas: 166 distribuídas por 20 ficheiros

Frequência da totalidade das instâncias da concordância: 0.56 %

||<| <<| >>| ||

#	f	Frase onde ocorre a concordância	Info
1	0	Neurons	
2	2	Cells of the nervous system , called nerve cells or neurons , are specialized to carry " messages " through an electrochemical process .	
3	3	The human brain has about 100 billion neurons .	
4	4	To learn how neurons carry messages , read about the action potential .	
5	5	Neurons come in many different shapes and sizes .	
6	6	Some of the smallest neurons have cell bodies that are only 4 microns wide .	
7	7	Some of the biggest neurons have cell bodies that are 100 microns wide .	
8	9	Neurons are similar to other cells in the body because : Neurons are surrounded by a cell membrane .	
9	9	Neurons are similar to other cells in the body because : Neurons are surrounded by a cell membrane .	
10	10	Neurons have a nucleus that contains genes .	
11	11	Neurons contain cytoplasm , mitochondria and other " organelles " .	
12	12	Neurons carry out basic cellular processes such as protein synthesis and energy production .	
13	13	However , neurons differ from other cells in the body because : Neurons have specialized extensions called dendrites and axons .	
14	13	However , neurons differ from other cells in the body because : Neurons have specialized extensions called dendrites and axons .	
15	15	Neurons communicate with each other through an electrochemical process .	
16	16	Neurons contain some specialized structures (for example , synapses) and chemicals (for example , neurotransmitters) .	
17	18	One way to classify neurons is by the number of extensions that extend from the neuron's cell body (soma) .	

Corpógrafo

- ◆ Concordances
 - ◆ Using regular expressions (e.g. “(the|a) \w+ is \w+”)
- ◆ N-grams
- ◆ Identification of multi-word expressions / lexical bundles

Corpógrafo - Linguateca

- Gestor
- Pesquisa
 - Concordância Frase
 - Estudo de N-Gramas**
 - Concordância Janela
 - Concordância KWIC
 - Concordâncias paralelas
- Centro de Conhecimento
- Centro de Comunicação

N-Gramas

- * [Exportar](#)
- * [Voltar ao menu](#)

Imprimir!

Informações

Corpus: Natura-PUBLICO (1726679 átomos)
 Lista com 981104 N-Gramas de comprimento 3

[|<](#) [<<](#) [>>](#) [|>](#)

#	n-grama	# oc.	freq. %
1	milhões de contos	628	0.036
2	, o que	514	0.029
3	, que se	348	0.020
4	no entanto ,	331	0.019
5	, com a	298	0.017
6	, com o	293	0.016
7	, depois de	292	0.016
8	o presidente da	249	0.014
9	por cento ,	246	0.014
10	milhões de dólares	239	0.013
11	secretário de estado	239	0.013
12	, e a	217	0.012
13	em lisboa ,	216	0.012
14	a partir de	215	0.012
15	por cento ,	215	0.012
16	, e o	213	0.012
17	de acordo com	213	0.012
18	de contos ,	212	0.012
19	, em que	211	0.012
20	, no entanto	192	0.011
21	do ano passado	186	0.010
22	, em lisboa	179	0.010
23	o facto de	179	0.010
24	ano passado ,	177	0.010
25	por cento do	176	0.010
26	, apesar de	174	0.010

Corpógrafo

- ◆ Concordances
 - ◆ Using regular expressions (e.g. “(the|a) \w+ is \w+”)
- ◆ N-grams
 - ◆ Identification of multi-word expressions / lexical bundles
- ◆ Term extraction
 - ◆ Based on n-grams (using filters / stop-words)

Corpógrafo

- Term extraction

The screenshot shows the Corpógrafo web application in a Mozilla Firefox browser. The address bar displays the URL: <http://poloclup.linguatca.pt/cgi-bin/corpografo/user/sagi/sagi.pl>. The page title is "Corpógrafo - neurodemo - Mozilla Firefox".

The main content area is titled "Candidatos a Termo" and displays the following information:

- Corpus: Neurons
- Tamanho: 29181 Tokens
- BD Terminológica: Neurons
- Candidatos encontrados: 1499
- Apresentados candidatos de 1 a 200

Navigation links include: [<<] [<] [>] [>>]

Summary: Candidatos encontrados: 1499, Apresentados candidatos de 1 a 200.

#	candidato a termo	#	OPM	Inserir!
1	cell body	37	1267.94	<input type="checkbox"/>
2	nervous system	29	993.79	<input type="checkbox"/>
3	spinal cord	27	925.25	<input type="checkbox"/>
4	nerve cell	24	822.45	<input type="checkbox"/>
5	action potential	22	753.91	<input type="checkbox"/>
6	synaptic cleft	19	651.10	<input type="checkbox"/>
7	central nervous system	13	445.49	<input type="checkbox"/>
8	plasma membrane	10	342.68	<input type="checkbox"/>
9	glial cell	9	308.41	<input type="checkbox"/>
10	peripheral nervous system	9	308.41	<input type="checkbox"/>
11	electrical activity	9	308.41	<input type="checkbox"/>
12	protein synthesis	9	308.41	<input type="checkbox"/>

Corpógrafo

- ◆ Concordances
 - ◆ Using regular expressions (e.g. “(the|a) \w+ is \w+”)
- ◆ N-grams
 - ◆ Identification of multi-word expressions / lexical bundles
- ◆ Term extraction
 - ◆ Based on n-grams (using filters / stop-words)
- Extraction of definitions / semantic relations
 - Pattern-based

Corpógrafo

➤ Extraction of Definitions

Example patterns:

```
(^|the |^a|^an | the | a | an )__TERMO__ is (a|an|the) \w+ \w+ of the
(^|the |^a|^an | the | a | an )__TERMO__ is (a|an|the) \w+ of
(^|the |^a|^an | the | a | an )__TERMO__ is \w+ with
(^|the |^a|^an | the | a | an )__TERMO__ is made of
(^|the |^a|^an | the | a | an )__TERMO__ is made up of
(^|the |^a|^an | the | a | an )__TERMO__ is the main
(^|the |^a|^an | the | a | an )__TERMO__ is (that|the) part of
(^|the |^a|^an | the | a | an )__TERMO__ is (part|composed) of
```

Corpógrafo - Linguateca

- Gestor
- Pesquisa
- Centro de Conhecimento
 - BD Terminológicas
 - Gestor de Relações
 - BD Fraseológicas
- Centro de Comunicação

Definições do Termo

- * [Listar Definições](#)
- * [Inserir ou Pesquisar](#)
- * [Ficha do Termo](#)

Informações

neuron

Pesquisa efectuada sobre o corpus **Neurons**.
Definições propostas [19]:

The neuron is the functional unit of the nervous system .	Guardar
Padrão de suporte: (^ ^the ^a ^an the a an)neuron is (a an the) \w+ \w+ of the	
Ficheiro de origem: EN_nervoussystem.htm (The Nervous System)	

A neuron is the functioning unit of the nervous system ; specialized to receive , integrate , and transmit information .	Guardar
Padrão de suporte: (^ ^the ^a ^an the a an)neuron is (a an the) \w+ \w+ of the	
Ficheiro de origem: EN_StructureoftheNeuron-Introduction.htm (Structure of the Neuron - Introduction)	

A neuron has many of the same " organelles " , such as mitochondria , cytoplasm and a nucleus , as other cells in the body .	Guardar
Padrão de suporte: (^ ^the ^a ^an the a an)neuron has (many several a an all one two various the)	
Ficheiro de origem: EN_Neurons.htm (Neurons)	

The Myelin Sheath of a neuron consists of fat-containing cells that insulate the axon from electrical activity .	Guardar
Padrão de suporte: (^ ^the ^a ^an ^each the a an each)neuron (results releases prevents holds dominates causes originates represents controls contains conducts defines spreads influences determines produces becomes presents occurs describes regulates acts enables equals brings consists corresponds leads depends takes receives derives plays)	
Ficheiro de origem: EN_StructureoftheNeuron-MyelinSheath.htm (Structure of the Neuron - Myelin Sheath)	

The Soma (Latin , som a : body) or cell body of a neuron contains the nucleus and other structures common to living cells .	
Padrão de suporte: (^ ^the ^a ^an ^each the a an each)neuron (results releases prevents holds dominates	

• Extraction of Semantic Relations

Example patterns for holonymy / meronymy:

```
(^|a |an |the | a | an | the )__TERMO2__ is located in the __TERMO1__
(^|a |an |the | a | an | the )__TERMO1__ is (composed|made up) of .* : the .* ,
(the|a|an) __TERMO2__
(^|a |an |the | a | an | the )__TERMO1__ is (composed|made up) of .* :
(the|a|an) __TERMO2__
(^|a |an |the | a | an | the )__TERMO1__ which is composed of \w+ __TERMO2__
(^|a |an |the | a | an | the )__TERMO2__ is the \w+ of the __TERMO1__
(^|a |an |the | a | an | the )__TERMO2__ , the \w+ part of the __TERMO1__
(^|a |an |the | a | an | the )__TERMO1__ contains the \w+ , __TERMO2__
(^|a |an |the | a | an | the )__TERMO1__ consists of (the|a|an) __TERMO2__
```

Corpógrafo - Linguatca

- Gestor
- Pesquisa
- Centro de Conhecimento
 - BD Terminológicas
 - Gestor de Relações
 - BD Fraseológicas
- Centro de Comunicação

Pesquisar relações

- * Usando restrições lexicais
- * Definindo manualmente
- * Ver todas as relação
- * Ficha do Termo

Informações

neuron é holónimo de (composto por) ...

dendrites ["neuron" já está definido como <i>holónimo de (composto por)</i> "dendrites"]	Guardar!
Thus , the dendrites of a neuron provide a surface for receiving synaptic inputs from other neurons .	
Padrão de suporte: (^ ^the the)dendrites of (the a an) neuron	
The arbor formed by the dendrites of a neuron often has a characteristic shape as determined by the spatial domains into which the dendrites ramify .	
Padrão de suporte: (^ ^the the)dendrites of (the a an) neuron	
The axon is considerably thicker and longer than the dendrites of a neuron .	
Padrão de suporte: (^ ^the the)dendrites of (the a an) neuron	
Thus , the dendrites of a neuron provide a surface for receiving synaptic inputs from other neurons .	
Padrão de suporte: (^ ^the the)dendrites of (the a an) neuron	
The arbor formed by the dendrites of a neuron often has a characteristic shape as determined by the spatial domains into which the dendrites ramify .	
Padrão de suporte: (^ ^the the)dendrites of (the a an) neuron	
The axon is considerably thicker and longer than the dendrites of a neuron .	
Padrão de suporte: (^ ^the the)dendrites of (the a an) neuron	
One neuron usually has several dendrites , and these branch out in a tree-like fashion .	
Padrão de suporte: (^ ^a ^an ^the ^one a an the one)neuron (typically usually normally) has (several many various) dendrites	
A neuron typically has many dendrites and one axon .	
Padrão de suporte: (^ ^a ^an ^the ^one a an the one)neuron (typically usually normally) has (several many various) dendrites	

axon ["neuron" já está definido como <i>holónimo de (composto por)</i> "axon"]	Guardar!
However , each neuron has only one axon (examples : spinal motor neurons , pyramidal neurons , Purkinje cells) .	

1) Dados Gerais

idioma	
tipo	entrada terminológica
estado administrativo	estandardizado
registo	neutro
frequência de utilização	usado com frequência
proveniência	empréstimo interdisciplinar

2) Pesquisadores

Dados relativos aos pesquisadores indisponíveis

3) Autores

- 1: Silvia Helena Cardoso, PhD
- 2: Ulrika Kahl
- 3: Li I. Zhang
- 4: Ulrika Kahl, PhD
- 5: Mu-ming Poo
- 6: M.J. Farabee

4) Fontes

Dados relativos às fontes indisponíveis

5) Morfologia

género	?
número	S
morfologia	NC

6) Definições

The neuron is the functional unit of the nervous system .

Corpus: Neurons

Ficheiro: EN_nervoussystem.htm

A neuron is the functioning unit of the nervous system ; specialized to receive , integrate , and transmit information .

Corpus: Neurons

Ficheiro: EN_StructureoftheNeuron-Introduction.htm

A neuron has many of the same " organelles " , such as mitochondria , cytoplasm and a nucleus , as other cells in the body .









Corpus: Neurons

Ficheiro: EN_Neurons.htm

7) Contextos

Não existem contextos para este termo.

8) Relações Semânticas

neuron	é merónimo de (parte de)	brain 
neuron	é holónimo de (composto por)	axon 
neuron	é holónimo de (composto por)	dendrites 
neuron	é holónimo de (composto por)	myelin sheath 
neuron	é holónimo de (composto por)	cell body 
neuron	é holónimo de (composto por)	terminal button 
neuron	é holónimo de (composto por)	synaptic site 
neuron	é holónimo de (composto por)	dendritic branching 

9) Termos Relacionados

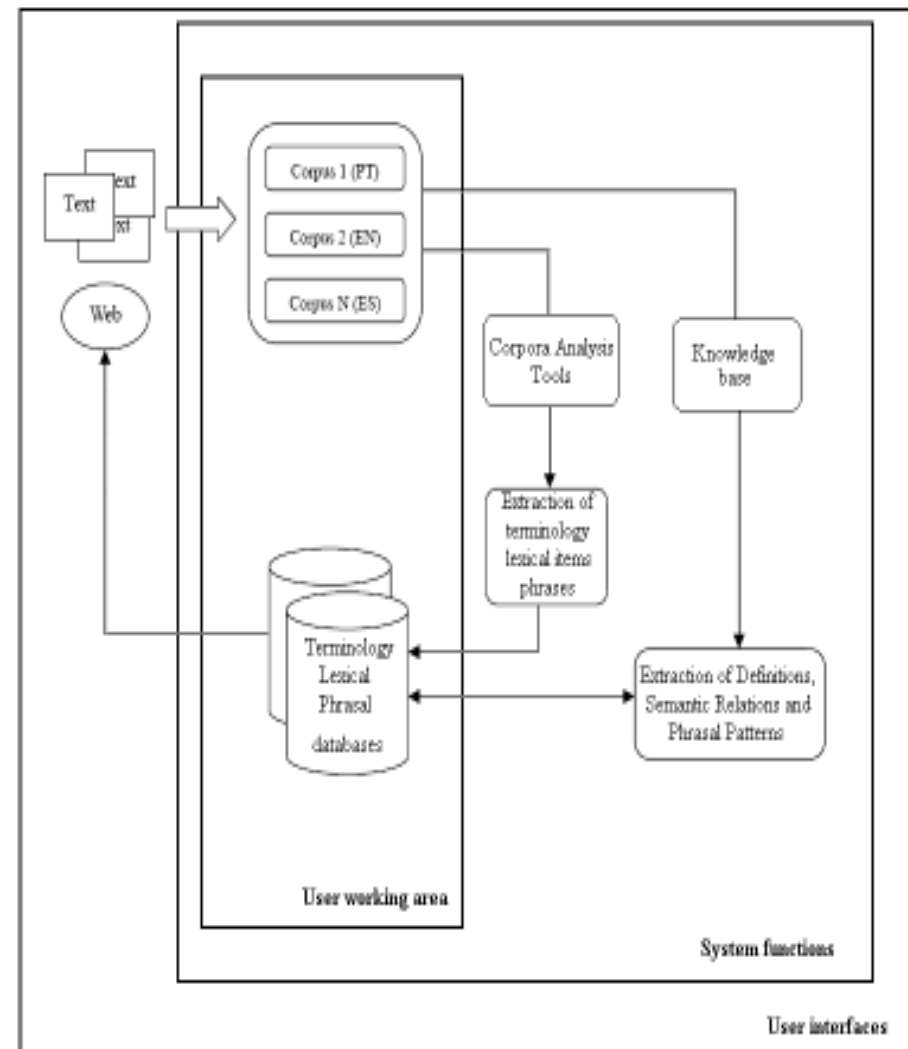
Não existem termos termos relacionados.

10) Equivalentes de Tradução

neurónio	Português (PT)	sinónimo
neurone	Italiano	sinónimo
neuron	Alemão	sinónimo
neurona	Espanhol	sinónimo
neurone	Francês	sinónimo

Corpógrafofo

- New tools
 - Phraseology / Lexicography
 - Identification of phrases and lexical items
 - Relations and translation equivalents
 - Discourse analysis
 - Genre analysis
 - Aligned concordances
 - Parallel corpora
 - Comparable corpora



Integrating NooJ

- ◆ Motivation

- ◆ Integrate POS information for obtaining concordances and extracting lexical items / phrases
- ◆ Syntactic and semantic analysis

- ◆ Why NooJ?

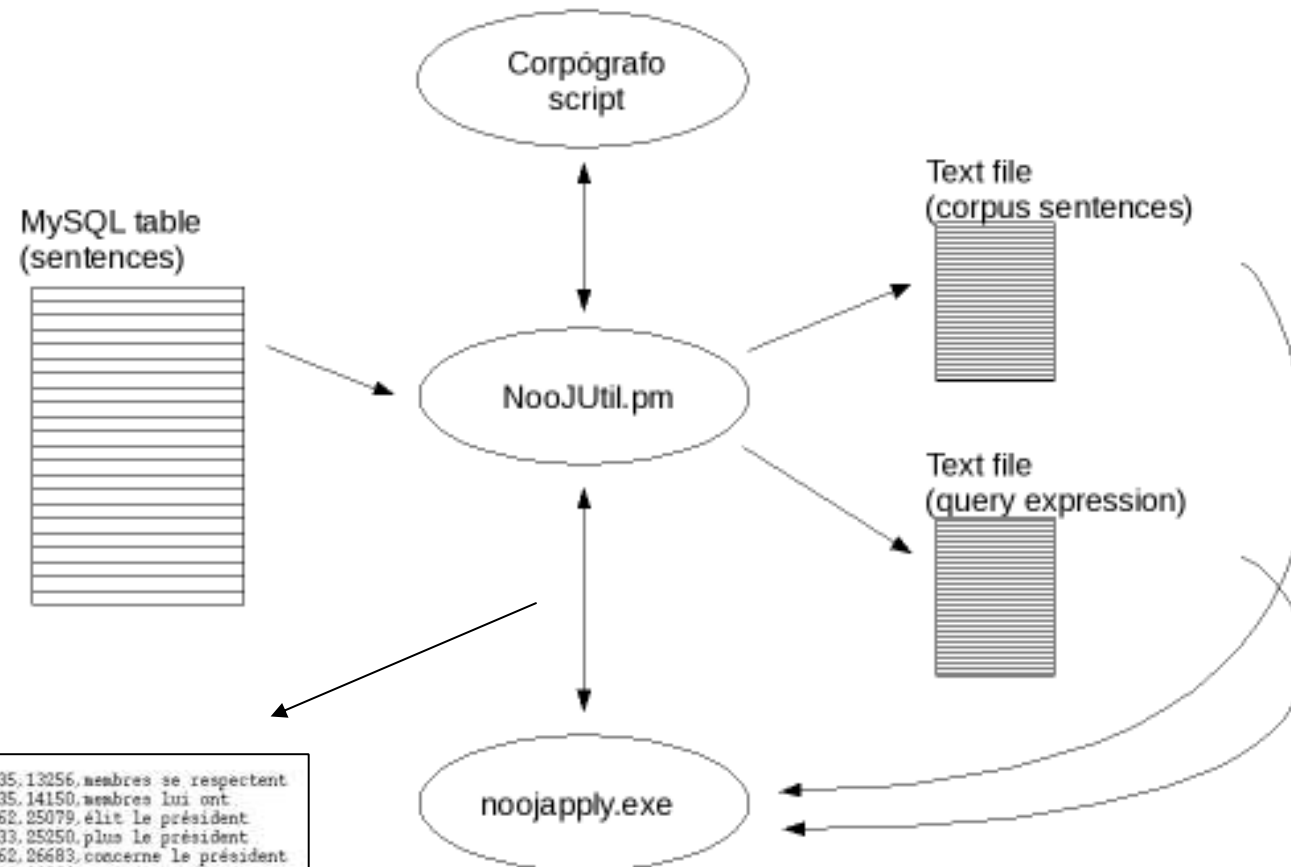
- ◆ Comprehensive corpus processing engine
- ◆ Extensible through the use of lexical and syntactic grammars
- ◆ Linguistic resources available for various languages

Integrating NooJ

- ◆ Practical aspects
 - ◆ Executable in Linux under the Mono platform (<http://www.mono-project.com/>)
 - ◆ Command-line programme (nooapply.exe)
 - ◆ Communication through text files
- ◆ Implementation
 - ◆ Perl module (NooJUtil.pm)
 - ◆ **init**: write corpus sentences to file; select dictionary
 - ◆ **query**: obtain NooJ concordances with nooapply.exe
 - ◆ **concordance**: obtain results and convert to Corpógrafo's concordance format

Integrating NooJ

- NooJUtil module



```
/var/tmp/frases_corpus_sntos21.txt,13235,13256,membres se respectent
/var/tmp/frases_corpus_sntos21.txt,14135,14150,membres lui ont.
/var/tmp/frases_corpus_sntos21.txt,25062,25079,élit le président
/var/tmp/frases_corpus_sntos21.txt,25233,25250,plus le président
/var/tmp/frases_corpus_sntos21.txt,26662,26683,concerne le président
/var/tmp/frases_corpus_sntos21.txt,26919,26939,traités en disposent
/var/tmp/frases_corpus_sntos21.txt,28683,28703,traités en disposent
/var/tmp/frases_corpus_sntos21.txt,29351,29369,qualifiés qui sont
/var/tmp/frases_corpus_sntos21.txt,31953,31973,traités en disposent
/var/tmp/frases_corpus_sntos21.txt,32062,32082,traités le prévoient
/var/tmp/frases_corpus_sntos21.txt,40035,40057,membres qui souhaitent
/var/tmp/frases_corpus_sntos21.txt,41063,41084,membres y participent
/var/tmp/frases_corpus_sntos21.txt,47795,47815,traités en disposent
/var/tmp/frases_corpus_sntos21.txt,70806,70829,membres qui constituent
/var/tmp/frases_corpus_sntos21.txt,72500,72523,membres qui remplissent
/var/tmp/frases_corpus_sntos21.txt,72976,72995,membres lui doivent
```

Integrating NooJ

- Converting results

```
/var/tmp/frases_corpus_smatos21.txt, 13235, 13256, membres se respectent
/var/tmp/frases_corpus_smatos21.txt, 14135, 14150, membres lui ont
/var/tmp/frases_corpus_smatos21.txt, 25062, 25079, élit le président
/var/tmp/frases_corpus_smatos21.txt, 25233, 25250, plus le président
/var/tmp/frases_corpus_smatos21.txt, 26662, 26683, concerne le président
/var/tmp/frases_corpus_smatos21.txt, 26919, 26939, traités en disposent
/var/tmp/frases_corpus_smatos21.txt, 28683, 28703, traités en disposent
/var/tmp/frases_corpus_smatos21.txt, 29351, 29369, qualifiée qui sont
/var/tmp/frases_corpus_smatos21.txt, 31953, 31973, traités en disposent
/var/tmp/frases_corpus_smatos21.txt, 32062, 32082, traités le prévoient
/var/tmp/frases_corpus_smatos21.txt, 40035, 40057, membres qui souhaitent
/var/tmp/frases_corpus_smatos21.txt, 41063, 41084, membres y participent
/var/tmp/frases_corpus_smatos21.txt, 47795, 47815, traités en disposent
/var/tmp/frases_corpus_smatos21.txt, 70806, 70829, membres qui constituent
/var/tmp/frases_corpus_smatos21.txt, 72500, 72523, membres qui remplissent
/var/tmp/frases_corpus_smatos21.txt, 72976, 72995, membres lui doivent
```

Select corresponding sentence

Concordancia.pm

- left and right context
- origin file
- occurrences

Using NooJ Expressions in Corpógrafo

Applications Places System Sérgio Matos Fri 18 Apr, 11:52

Corpógrafo - smatos21 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Corpógrafo - Linguateca

- Gestor
- Pesquisa
 - Concordância Frase
 - Estudo de N-Gramas
 - Concordância Janela
 - Concordância KWIC**
 - Concordâncias paralelas
- Centro de Conhecimento
- Centro de Comunicação

Concordância KWIC

1. Seleccione corpus a pesquisar:
2. Introduza expressão de pesquisa:
3. Contexto de visualização
 - palavras
 - caracteres

Done

Corpógrafo - smatos21 - Mozilla Firefox

http://poloclup.linguateca.pt - Concordã

Poderá também obter concordâncias utilizando **expressões regulares NooJ**.

Alguns símbolos utilizados nas Expressões de Pesquisa:

- ^ início do texto
- \$ final do texto
- + alternativa (ou)
- * repetição
- <A> adjetivo
- <ADV> advérbio
- <DET> determinante
- <E> expressão vazia (ver exemplo)
- <INT> interjeição
- <N> nome / substantivo
- <NB> sequência de dígitos
- <PREP> preposição
- <PRO> pronome
- <V> verbo
- <WF> uma palavra
- MP="..." pesquisar com padrões (ver exemplo)

[Ver exemplos](#) [Voltar](#)

Done

Using NooJ Expressions in Corpógrafo

Applications Places System Sérgio Matos Wed 16 Apr, 14:43

Corpógrafo - smatos21 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Corpógrafo - Linguateca

Gestor
Pesquisa
Concordância Frase
Estudo de N-Gramas
Concordância Janela
Concordância KWIC
Concordâncias paralelas
Centro de Conhecimento
Centro de Comunicação

Expressão de pesquisa: <V><PRO><V+PR+3+p>
Corpus pesquisado: Traité de Lisbonne (103899 átomos : 1 ficheiros)

Nº de concordâncias obtidas: 48 distribuídas por 1 ficheiros
Frequência da totalidade das instâncias da concordância: 0.04 %

||< << >> >|

#	POS (C/P)	esq.	conc.	dir.	nº frase	info
1	67 / 14	et les États	membres se respectent	et s' assistent	61	
2	106 / 21	que les États	membres lui ont	attribuées dans les	69	
3	4 / 2	Il	élit le président	de la Commission	142	
4	49 / 10	cent cinquante ,	plus le président	.	145	
5	139 / 28	en ce qui	concerne le président	de la Commission	162	
6	73 / 15	cas où les	traités en disposent	autrement .	165	
7	69 / 15	cas où les	traités en disposent	autrement .	180	
8	71 / 11	de la majorité	qualifiée qui sont	applicables jusqu' au	186	
9	115 / 24	cas où les	traités en disposent	autrement .	210	
10	75 / 13	Commission lorsque les	traités le prévoient	.	211	
11	10 / 3	Les États	membres qui souhaitent	instaurer entre eux	274	
12	290 / 49	moins neuf États	membres y participent	.	278	
13	130 / 28	cas où les	traités en disposent	autrement .	313	
14	10 / 3	Les États	membres qui constituent	entre eux des	432	
15	10 / 3	Les États	membres qui remplissent	des critères plus	441	
16	102 / 21	les autres États	membres lui doivent	aide et assistance	443	
17	10 / 3	Les États	membres qui remplissent	des critères plus	441	

Done

Corpógrafo - smatos21 - Mozilla Firefox

Using NooJ Expressions in Corpógrafo

Applications Places Desktop Fri Mar 7, 4:37 PM

Corpógrafo - RTraducao - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Corpógrafo - Linguateca

Gestor
Ficheiros
Corpora
Autores
Organizações
Repositório multimédia
Pesquisador
Corpora paralelos
Pesquisa
Centro de Conhecimento
Centro de Comunicação

Pesquisar corpora alinhados
* Voltar
Informações

Pesquisar corpora paralelos

Resultado da pesquisa por: <ADV+MP="mente\$"> <WF>

Resultados encontrados: 180
Apresentando resultados de 1 a 50

[<<] [<] [>] [>>]

#	Texto original	Texto alinhado
1	Em virtude do princípio da cooperação leal , a União e os Estados-Membros respeitam-se e assistem-se mutuamente no cumprimento das missões decorrentes dos Tratados .	Pursuant to the principle of sincere cooperation , the Union and the Member States shall , in full mutual respect , assist each other in carrying out tasks which flow from the Treaties .
2	Em virtude do princípio da subsidiariedade , nos domínios que não sejam da sua competência exclusiva , a União intervém apenas se e na medida em que os objetivos da acção considerada não possam ser suficientemente alcançados pelos Estados-Membros , tanto ao nível central como ao nível regional e local , podendo contudo , devido às dimensões ou aos efeitos da acção considerada , ser mais bem alcançados ao nível da União .	Under the principle of subsidiarity , in areas which do not fall within its exclusive competence , the Union shall act only if and insofar as the objectives of the proposed action cannot be sufficiently achieved by the Member States , either at central level or at regional and local level , but can rather , by reason of the scale or effects of the proposed action , be better achieved at Union level .
3	Os cidadãos estão directamente representados , ao nível da União , no Parlamento Europeu .	Citizens are directly represented at Union level in the European Parliament .
4	Os Estados-Membros estão representados no Conselho Europeu pelo respectivo Chefe de Estado ou de Governo e no Conselho pelos respectivos Governos , eles próprios democraticamente responsáveis , quer perante os respectivos Parlamntos nacionais , quer perante os seus cidadãos .	Member States are represented in the European Council by their Heads of State or Government and in the Council by their governments , themselves democratically accountable either to their national Parliaments , or to their citizens .
5	As instituições , recorrendo aos meios adequados , dão aos cidadãos e às associações representativas a possibilidade de expressarem e partilharem publicamente os seus pontos de vista sobre todos os domínios de acção da União .	The institutions shall , by appropriate means , give citizens and representative associations the opportunity to make known and publicly exchange their views in all areas of Union action . 2 .
6	Artigo 8.o -C Os Parlamntos nacionais contribuem activamente para o bom funcionamento da União : a) Sendo informados pelas instituições da União e notificados dos projectos de actos legislativos da União , de acordo com o Protocolo relativo ao papel dos Parlamntos nacionais na União Europeia; b) Garantindo o respeito pelo princípio da subsidiariedade , de acordo com os procedimentos previstos	Article 8 C National Parliaments contribute actively to the good functioning of the Union : (a) through being informed by the institutions of the Union and having draft legislative acts of the Union forwarded to them in accordance with the Protocol on the role of national Parliaments in the European Union; (b) by seeing to it that the principle of subsidiarity is respected in accordance with the procedures provided for in the Protocol

Done

Corpógrafo - RTraducao - Mozilla Firefox

Using NooJ Expressions in Corpógrafo

Applications Places Desktop Fri Mar 7, 5:27 PM

Corpógrafo - RTraducao - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Base de dados activa: BD Fraseológica

Corpógrafo - Linguateca

- Gestor
- Pesquisa
- Centro de Conhecimento
 - BD Terminológicas
 - Gestor de Relações
 - BD Fraseológicas**
- Centro de Comunicação

Resultado da Pesquisa

- * Voltar...
- * Menu principal da BD
- * Listar e editar Itens Lexicais

Informações

Base de dados activa: BD Fraseológica

Candidatos a Item Lexical

Corpus: Seguros
 Tamanho: 20734 Tokens
 BD Fraseológica: **BD Fraseológica**

Expressão de pesquisa: <ADV> <WF>

Candidatos encontrados: 452
 Apresentados candidatos de 1 a 200

[<< | < | ≥]

Candidatos encontrados: 452
 Apresentados candidatos de 1 a 200

#	candidato a item lexical	#	OPM	Inserir
1	Actualmente a	1	48.22	<input type="checkbox"/>
2	agora confrontadas	1	48.22	<input type="checkbox"/>
3	ainda as	1	48.22	<input type="checkbox"/>
4	ainda definir	1	48.22	<input type="checkbox"/>
5	ainda mais	1	48.22	<input type="checkbox"/>
6	amplamente melhorada	1	48.22	<input type="checkbox"/>
7	apenas destinado	1	48.22	<input type="checkbox"/>
8	Apenas deve	1	48.22	<input type="checkbox"/>
9	apenas dos	1	48.22	<input type="checkbox"/>
10	apenas necessitam	1	48.22	<input type="checkbox"/>
11	apenas passa	1	48.22	<input type="checkbox"/>
12	Apenas sábados	2	96.45	<input type="checkbox"/>
13	apenas se	2	96.45	<input type="checkbox"/>
14	automaticamente ajusta	1	48.22	<input type="checkbox"/>
15	automaticamente em	1	48.22	<input type="checkbox"/>
16	automaticamente na	1	48.22	<input type="checkbox"/>
17	automaticamente para	1	48.22	<input type="checkbox"/>
18	automaticamente pelo	1	48.22	<input type="checkbox"/>
19	automaticamente todas	1	48.22	<input type="checkbox"/>
20	Basicamente são	1	48.22	<input type="checkbox"/>
21	bem como	4	144.68	<input type="checkbox"/>
22	cada ambiente	1	48.22	<input type="checkbox"/>
23	cada apólice	2	96.45	<input type="checkbox"/>

http://localhost - Concordâncias de "bem como" - Mozilla Firefox

Pesquisa de contexto para o candidato a termo: bem como
 Corpus pesquisado: Seguros (20734 átomos)
 Ocorrências encontradas: 4
 Ocorrências por milhão no corpus (OCM): 192.910

Listando resultados...

#	concordância no contexto de uma frase
1	Criação de novos registos no protocolo Está disponível o tratamento via protocolo de valores de balanço por apólice bem como as cargas da apólice .
2	Assim , a solução GIS é constituída por um conjunto de módulos aplicativos que suportam todas as actividades fundamentais de qualquer interveniente da actividade seguradora , bem como permite facilmente interligar e compatibilizar os sistema de informação dos diferentes intervenientes .
3	Através do módulo de configuração é possível definir as múltiplas informações requeridas nos questionários dos processos de sinistro de cada ramo/tipo de seguro bem como as regras de tratamento associadas (provisões de sinistros , controlo de parcelas por natureza , agravamento/bónus do prémio , regras de aplicação de franquias ou participações , etc.)
4	Todo o funcionamento do sistema é baseado numa configuração de produtos , que depois é interpretada pelos componentes de negócio , de forma a disponibilizar o conjunto de funcionalidades e serviços associados ao produto , bem como a sua total integração com o sistema de gestão de back office , implementando assim uma arquitectura em três camadas .

[Fechar]

Done

Using NooJ Expressions in Corpógrafo

Applications Places System Sérgio Matos Fri 18 Apr, 11:37

Corpógrafo - neurodemo - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

Corpógrafo - Linguateca

- Gestor
- Pesquisa
 - Concordância Frase
 - Estudo de N-Gramas
 - Concordância Janela
 - Concordância KWIC
 - Concordâncias paralelas
- Centro de Conhecimento
- Centro de Comunicação

Concordâncias

- * [Voltar ao menu](#)
- * [Analisar Freqüências](#)
- * [Contexto Janela 1-1](#)
- * [Contexto Janela 2-2](#)

[Imprimir!](#)

Informações

Expressão de pesquisa: <DET> (neuron + neurons) <V+PR>
 Corpus pesquisado: Neurons (29181 átomos : 37 ficheiros)

Nº de concordâncias obtidas: 39 distribuídas por 16 ficheiros
 Freqüência da totalidade das instâncias da concordância: 0.13 %

[|<](#) [<<](#) [>>](#) [|>](#)

#	f	Frase onde ocorre a concordância	Info
1	1	The Axon of a neuron is a singular fiber that carries information away from the soma to the synaptic sites of other neurons (dendrites and somas) , muscles , or glands .	
2	1	The Dendrites (Greek , dendro : tree) of a neuron are its many short , branching fibers extending from the cell body or soma .	
3	1	A neuron is the functioning unit of the nervous system ; specialized to receive , integrate , and transmit information .	
4	1	The Myelin Sheath of a neuron consists of fat-containing cells that insulate the axon from electrical activity .	
5	1	The Soma (Latin , som a : body) or cell body of a neuron contains the nucleus and other structures common to living cells .	
6	1	The Terminal Buttons of a neuron are the small knobs at the end of an axon that release chemicals called neurotransmitters .	
7	1	A neuron is a cell specialized to conduct electrochemical impulses called nerve impulses or action potentials .	
8	1	This page will describe the way that neurons work .	
9	3	It is at these junctions that neurons are excited , inhibited , or modulated .	
10	4	Many neurons do not have axons .	
11	5	Information from one neuron flows to another neuron across a synapse .	
12	6	The neuron is the functional unit of the nervous system .	
13	6	The dendritic branching of a neuron is changeable in the mature nervous system , either growing or retracting .	
14	7	The axons of many neurons are wrapped in a myelin sheat , which is composed of the membranes of interstitial cells and is wrapped around the axons to form several concentric layers .	
15	8	While variable in size and shape , all neurons have three parts .	
16	9	Thus , the dendrites of a neuron provide a surface for receiving synaptic inputs from other neurons .	
17	12	Dendrites of some neurons are smooth , tapered processes , such as in motor neurons of the spinal cord .	

Done

Corpógrafo - neurodemo - Mozilla Firefox

Using NooJ Expressions in Corpógrafo

Applications Places System Sérgio Matos Mon 21 Apr, 14:49

Corpógrafo - neurodemo - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Corpógrafo - Linguateca

Expressão de pesquisa: <DET> <neurone> <V+PR>
Corpus pesquisado: Le Neurone (22748 átomos : 12 ficheiros)

Nº de concordâncias obtidas: 19 distribuídas por 6 ficheiros
Frequência da totalidade das instâncias da concordância: 0.08 %

Navigation icons: <| << >> >|

#	f	Frase onde ocorre a concordância	Info
1	1	Les neurones reçoivent simultanément des milliers d'informations activant des récepteurs-canaux et/ou des récepteurs liés aux protéines G.	Info
2	6	Ceci est basé sur l'éloignement entre la cellule endocrinienne et sa cible et le fait que le message emprunte la voie systémique , alors que les neurones sont en contact quasi direct par l'intermédiaire des synapses .	Info
3	9	La sécrétion , très focalisée et dirigée uniquement vers les cellules avec lesquelles le neurone est connecté , se fait au niveau des synapses .	Info
4	10	Les dendrites de certains neurones contiennent un nombre important de canaux sodiques et/ou calciques sensibles au potentiel .	Info
5	10	Le neurone diffère ainsi (à quelques exceptions près) des autres cellules sécrétrices de l'organisme comme les cellules hormonales , qui libèrent leur produit de sécrétion dans la circulation sanguine (sécrétion endocrine) .	Info
6	11	Les neurones ont des formes très diverses et leur morphologie (dimension et orientation des champs dendritiques , taille du soma) joue un rôle fondamental dans leurs capacités d'intégration et de traitement de l'information .	Info
7	12	Les neurones sont amitotiques (incapables de se reproduire) et ont donc une longévité extrême .	Info
8	15	LES NEURONES ONT UN CORPS CELLULAIRE D' OU PARTENT DEUX TYPES DE PROLONGEMENTS : LES DENDRITES ET L'AXONE L'ARBRE SOMATO-DENDRITIQUE REPRÉSENTE LE PÔLE RÉCEPTEUR DU NEURONE	Info
9	19	Le neurone comprend :	Info
10	20	Les neurones effectuent donc des opérations complexes nécessitant la sommation de tous les PPSE pour produire une dépolarisation postsynaptique significative : c'est l'intégration synaptique .	Info
11	21	2 . Le neurone est une cellule dite polarisée (asymétrique) avec d'un côté les dendrites , qui constituent le pôle normalement récepteur .	Info
12	24	Le neurone est la cellule qui élabore , modifie et transmet l'influx nerveux .	Info
13	25	Les neurones sont ainsi dits " épineux " ou " lisses " .	Info
14	28	Lorsqu'ils ont atteint leur position définitive , les neurones acquièrent un cône de croissance , qui s'allonge progressivement pour former l'axone .	Info
15	39	Entre les neurones sont entremêlés de nombreux neurites .	Info

Done

Corpógrafo - neurodemo - Mozilla Firefox

Using NooJ Expressions in Corpógrafo

Applications Places System Sérgio Matos Fri 18 Apr, 12:09

Corpógrafo - neurodemo - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://poloclup.linguateca.pt/cgi-bin/corpografo/user/sagi/sagi.pl

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

Corpógrafo - Linguateca

- Gestor
- Pesquisa
 - Concordância Frase
 - Estudo de N-Gramas
 - Concordância Janela
 - Concordância KWIC
 - Concordâncias paralelas
- Centro de Conhecimento
- Centro de Comunicação

Concordâncias

- * [Voltar ao menu](#)
- * [Analisar Frequências](#)
- * [Contexto Janela 1-1](#)
- * [Contexto Janela 2-2](#)

[Imprimir!](#)

Informações

Expressão de pesquisa: <WF> <V> <WF> (of + for + in) <WF>
 Corpus pesquisado: Neurons (29181 átomos : 37 ficheiros)

Nº de concordâncias obtidas: 323 distribuídas por 31 ficheiros
 Frequência da totalidade das instâncias da concordância: 1.10 %

[|<](#) [<<](#) [>>](#) [|>](#)

#	f	Frase onde ocorre a concordância	Info
1	0	Activation of the transcription factor NF-[kappa]B in Schwann cells is required for peripheral myelin formation	
2	1	The human body is made up of trillions of cells .	
3	1	Nerve cells - or neurons - are the cells in our body that are responsible for bringing information from the surrounding , analyze this input , and finally create a perception .	
4	1	Dendrites branch out in treelike fashion and serve as the main apparatus for receiving signals from other nerve cells .	
5	1	This site will help you learn about the five major sections of a Nerve cell :	
6	1	A neuron is the functioning unit of the nervous system ; specialized to receive , integrate , and transmit information .	
7	1	The gaps (approximately 1 micrometer wide) formed between myelin sheath cells long the axons are called Nodes of Ranvier .	
8	1	Plant cells are eukaryotic and have many of the structures found in animal cells .	
9	1	The brain is made up of more than just nerve cells (neurons) .	
10	1	Activation of the transcription factor NF-B in Schwann cells is required for peripheral myelin formation .	
11	2	Nervous tissue is composed of two main cell types : neurons and glial cells .	
12	2	The axon is the main conducting unit of the neuron , capable of conveying electrical signals along distances that range from as short as 0.1 mm to as long as 2 m .	
13	2	In addition , all eukaryotic cells contain elaborate systems of internal membranes which set up various membrane-enclosed compartments within the cell .	
14	2	The energy released is trapped in the form of ATP for use by all the energy-consuming activities of the cell.	
15	2	The membrane with its mosaic of proteins is responsible for many important functions : keeping certain ions and small molecules out of the cell and letting others in , accumulating nutrients , rejecting harmful substances , catalyzing enzymatic reactions , and conducting an impulse being sensitive to particular neurotransmitters and modulators .	
16	2	Cilia and Flagella Cells contain elaborate arrays of protein fibers that serve such functions as : establishing cell shape providing	

Done

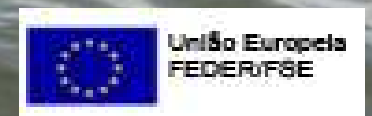
Corpógrafo - neurodemo - Mozilla Firefox

What next?

- ◆ Performance – function calls to NooJ library?
- ◆ Annotations
 - ◆ <TERM>, <LEX>, <DISCRS>, user defined?
 - ◆ Show / edit annotations
- ◆ Grammars
 - ◆ Disambiguation
 - ◆ Syntactic / semantic analysis
- ◆ Lexical / phrasal items
 - ◆ Rules and grammars
- ◆ Discourse analysis
 - ◆ Discourse markers – create dictionary?

Thank you!

This work was done in the scope of the Linguateca, contract nº339/1.3/C/NAC, project jointly funded by the Portuguese Government and the European Union.



Corpógrafo

- Parallel Corpora

The image displays the Corpógrafo application interface, which is used for creating parallel corpora. It shows a file explorer on the left, a central workspace for file alignment, and a results window at the bottom.

File Alignment Window:

Ficheiros a alinhar:

#	Ficheiro origem	Ficheiro destino	Ver / Editar	Eliminar
1	FRAC0306PTC_002.pdf	FRAC0306MC_002.pdf	Ver	Eliminar

Resultado do alinhamento:

Frases alinhadas: 2293
Aproximando frases de 2 a 90

#	FRAC0306PTC_002.pdf	FRAC0306MC_002.pdf
1	EGN 2725 2442 C 385(Jornal Oficial da União Europeia Número de informação sobre Página 2047C 29501 Tratado de Lisboa que altera o Tratado do União Europeia e o Tratado que institui a Comunidade Europeia, assinado em Lisboa em 12 de Dezembro de 2007...	EGN 1125 1425 C 306(Official Journal of the European Union Number No Contents Page 2047C 38501 Treaty of Lisbon amending the Treaty in European Union and the Treaty establishing the European Community, signed at Lisbon, 12 December 2007...
2	1 ALIENACIÖES INTRODUCIDAS NO TRATADO DA UNIÃO EUROPEIA E NO TRATADO QUE INSTITUI A COMUNIDADE EUROPEIA...	1 AMENDMENTS TO THE TREATY IN EUROPEAN UNION AND TO THE TREATY ESTABLISHING THE EUROPEAN COMMUNITY... 11 Article 1...
3		10 Article 2...

Corpógrafo

- “Aligned” concordances

In order to

1) Dados Gerais		
idioma		
tipo	unidade fraseológica	
estado administrativo	Indefinido	
registo	Indefinido	
frequência de utilização	usado com frequência	
proveniência	Indefinido	
2) Pesquisadores		
Dados relativos aos pesquisadores indisponíveis		
3) Autores		
Dados relativos aos autores indisponíveis		
4) Fontes		
1: Indefinida		
5) Morfologia		
género	?	
número	?	
morfologia	PP NC PP	
6) Definições		
Não existem definições para este item lexical.		
7) Contextos		
You will need linguistic skills , in order to cope with the language side of the degree . Corpus: Corpus EN Ficheiro: Manchester_English Language.doc		
In order to maintain quality , Cambridge has in recent years developed a team of Language Teaching Officers , native speakers who co-ordinate such teaching across the Faculty . Corpus: Corpus EN Ficheiro: Cambridge_mml.pdf		
8) Relações Léxico-Semânticas		
Não existem relações léxico-semânticas definidas para este item lexical.		
9) Relações Discursivas		
Propósito		
10) Itens Relacionados		
in order for	Inglês (UK)	sinónimo aproximado
in order that	Inglês (UK)	sinónimo aproximado
11) Equivalentes de Tradução		
para que	Português (PT)	sinónimo aproximado
de forma a	Português (PT)	sinónimo
De modo a	Português (PT)	sinónimo

Corpógrafo - UnifPorto - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://pokoilup.linguatca.pf/cgi-bin/corpgrafo/uso/sag/sag.pl

Corpógrafo - Linguatca

Concordâncias corpora comparáveis

Resultado da pesquisa por: **in order to** e **De modo a**

Apresentando resultados de 1 a 33

[=>] [-] [<=]

#	Texto original	Texto alinhado
1	In order to maintain quality , Cambridge has in recent years developed a team of Language Teaching Officers , native speakers who co-ordinate such teaching across the Faculty .	De modo a contextualizar a orientação principal (Parte II) , o Programa é iniciado com uma referência aos estudos empíricos da linguagem (Parte I) .
2	To provide students with the opportunity to study another modern foreign language (French/Spanish) in order to acquire practical skills and methods in translating English from into another foreign language .	Lógica e Linguagem Logic and Language Objectivos Os objectivos da disciplina são os seguintes : 1. fornecer instrumentos formais básicos que permitam compreender algumas relações fundamentais nas línguas ; 2. relacionar lógica e linguística de modo a tornar mais claros vários aspectos das línguas , em particular o significado ; 3. levar os estudantes a compreender que alguns sistemas lógicos fornecem meios rigorosos para o estudo das línguas , mas não se confundem com estas .
3	Students may take a paper on Language for Literature 1300-present as an alternative to the language paper , so it is not essential to have a particular skill in a foreign language in order to take Part I. Students also have the option of writing an open dissertation on any topic in English literature and its contexts (1300-present) .	Pretende-se que os alunos sejam capazes de estruturar a informação de modo a reconstituírem os mecanismos de organização quer da administração central quer periférica de Portugal Programa 0 .
4	You can also apply , after taking your final school or college examinations , for an unconditional offer of a place : in that case , you will normally be expected to have achieved AAA (or sometimes AAB) at A2 or equivalent in order to have a chance of success .	Lógica e Linguagem Logic and Language 2007/2008 - IS Objectivos Os objectivos da disciplina são os seguintes : 1. fornecer instrumentos formais básicos que permitam compreender algumas relações fundamentais nas línguas ; 2. relacionar lógica e linguística de modo a tornar mais claros vários aspectos das línguas , em particular o significado ; 3. levar os estudantes a compreender que alguns sistemas lógicos fornecem meios rigorosos para o estudo das línguas , mas não se confundem com estas .
5	It is a good idea to make early contact with the Admissions Tutor of the College (s) to which you are considering applying in order to explain your needs .	Os discentes serão iniciados nas técnicas , programas e equipamentos de imprensa , rádio , televisão e online de modo a incrementar a sua proficiência .
6	Skills Politics and International Relations : key skills This	GenPolitics e GenEstratégia Objectivos a) Reflectir sobre os

Done