



CURSO DE MESTRADO EM ENGENHARIA INFORMÁTICA

**ANÁLISE E DESENVOLVIMENTO DE FERRAMENTAS  
COMPUTACIONAS PARA O PROCESSAMENTO  
DE LÍNGUA PORTUGUESA**

**Vol. II**

José Eduardo de Carvalho Rosado Saianda

**Orientador:** Prof. Doutor Paulo Miguel Torres Duarte Quaresma

Évora, Dezembro de 2011





CURSO DE MESTRADO EM ENGENHARIA INFORMÁTICA

**ANÁLISE E DESENVOLVIMENTO DE FERRAMENTAS  
COMPUTACIONAS PARA O PROCESSAMENTO  
DE LÍNGUA PORTUGUESA**

**Vol. II**

José Eduardo de Carvalho Rosado Saianda

**Orientador:** Prof. Doutor Paulo Miguel Torres Duarte Quaresma

Évora, Dezembro de 2011



## ÍNDICE

Índice .....	iii
Anexo I – Ajuda à Redacção .....	9
Exemplo 2.1 - Definição de entradas no dicionário do Jspell .....	3
Exemplo 2.2 - utilização do Jspell a partir da linha de comando .....	3
Exemplo 2.3 - Obtenção do adjectivo cordiais a partir de cordial .....	4
Exemplo 2.4 - Conjugação do verbo cordear na segunda pessoa do plural do presente do indicativo.....	4
Exemplo 2.6 - Utilização do analisador morfológico do WebJspell .....	5
Anexo II – Alinhadores .....	7
Exemplo 2.7 - Alinhamento de textos utilizando o alinhador online CEPRIL.....	9
Exemplo 2.8 - Alinhamento de dois textos utilizando o VisualLIHLA.....	10
Exemplo 2.9 - Utilização do VisualTCA no alinhamento de textos .....	13
Anexo III – Analisadores .....	17
Exemplo 2.10 – Utilização do concordanciador de um milhão de palavras .....	19
Exemplo 2.11 - Exemplificação de utilização da ferramenta GojolParser .....	25
Exemplo 2.12 - Utilização das funções ‘atomiza’, ‘tokeniza’ e ‘tokenize’.....	27
Exemplo 2.13 - Utilização da função ‘cqptokens’ .....	28
Exemplo 2.14 - Utilização da função ‘cqptokens’ com opções.....	28
Exemplo 2.15 - Retorno das frases de um texto .....	28
Exemplo 2.16 - Retorno de um texto com uma frase por linha.....	28
Exemplo 2.17 - Aplicação de uma etiqueta XML a cada frase .....	28
Exemplo 2.18 - Utilização de etiquetas no módulo fsentences.....	29
Exemplo 2.19 - Utilização do lematizador verbal LX-Suite.....	29
Exemplo 2.20 – Utilização do conjugador verbal LX-Suite.....	30
Exemplo 2.21 - Utilização do flexionador nominal LX-Suite .....	32
Exemplo 2.22 - Utilização do anotador categorial LX-Suite .....	33
Exemplo 2.23 - Utilização do reconhecedor de nomes próprios do LX-Suite.....	33
Exemplo 2.24 – Utilização do navegador corpus do LX-Suite .....	34
Exemplo 2.25 - Utilização do buscador TreeBank.....	35
API do módulo Tokenizer da biblioteca Freeling .....	35
Exemplo 2.26 – Ficheiro de regras de tokenização.....	36
Exemplo 2.27 - Definição de uma macro .....	39

API do módulo splitter da biblioteca FreeLing.....	40
Exemplo 2.28 - Ficheiro de opções do splitter.....	40
Exemplo 2.29 - Secção <General>.....	41
Exemplo 2.30 - Secção <Markers >.....	41
Exemplo 2.31 - Secção <SentenceEnd>.....	41
Exemplo 2.32 - Secção <SentenceStart>.....	42
API do módulo analisador morfológico da biblioteca FreeLing.....	42
API do módulo detector de números da biblioteca FreeLing.....	43
API do módulo detector de marcas de pontuação da biblioteca FreeLing.....	44
Exemplo 2.33 - Ficheiro de tag's associadas a marcas de pontuação.....	44
Exemplo 2.34 - Linha com tag especial.....	44
API do módulo detector de datas da biblioteca FreeLing.....	45
API do módulo de pesquisa de dicionário da biblioteca FreeLing.....	45
Exemplo 2.35 - Exemplo de linhas de um ficheiro plano de dicionário.....	45
Exemplo 2.36 - Definição de contracção no ficheiro do dicionário.....	46
Exemplo 2.37 - Ficheiro de dicionário indexado.....	46
Exemplo 2.38 - Eliminação de afixo de uma palavra.....	49
Exemplo 2.39 - Adição de afixo a uma palavra.....	49
Exemplo 2.40 - Condição da tag condicional.....	49
Exemplo 2.41 - Afectação de lema.....	49
Exemplo 2.42 - Regras de prefixo.....	49
Exemplo 2.43 - Regra de sufixo.....	50
API do módulo identificador multpalavra da biblioteca FreeLing.....	51
Exemplo 2.44 - Ficheiro de definição multpalavra.....	51
API do módulo identificador de entidades nomeadas da biblioteca FreeLing.....	52
Exemplo 2.45 - Declaração da secção <FunctionWords>.....	52
Exemplo 2.46 - Declaração da secção <SpecialPunct>.....	53
Exemplo 2.47 - Declaração da secção <NE_Tag>.....	53
Exemplo 2.48 - Declaração da secção <Ignore>.....	53
Exemplo 2.49 - Declaração da secção <Names>.....	54
Exemplo 2.50 - Declaração das secções <RE_NounAdj> e <RE_Closed>.....	54
Exemplo 2.51 - Declaração da secção <TitleLimit>.....	54
Exemplo 2.52 - Declaração da secção <SplitMultiwords>.....	55
API do módulo bioner da biblioteca FreeLing.....	55
Exemplo 2.53 - Declaração da secção <RGF>.....	55

Exemplo 2.54 - Declaração da secção <AdaBoostModel> .....	55
Exemplo 2.55 - Declaração da secção <Lexicon> .....	56
Exemplo 2.56 - Declaração da secção <Classes> .....	56
Exemplo 2.57 - Declaração da secção <InitialProb> .....	56
Exemplo 2.58 - Declaração da secção <TransitionProb> .....	57
Exemplo 2.59 - Declaração da secção <TitleLimit> .....	57
API do módulo identificador de quantidades da biblioteca FreeLing.....	57
Exemplo 2.60 - Declaração da secção <Currency>.....	58
Exemplo 2.61 - Declaração da secção <Measure>.....	58
Exemplo 2.62 - Declaração da secção <MeasureNames>.....	59
API do módulo afectador de probabilidades e reconhecedor de palavras desconhecidas da biblioteca FreeLing.....	59
Exemplo 2.63 - Declaração da secção <FormTagFreq>.....	60
Exemplo 2.64 - Declaração da secção <ClassTagFreq> .....	60
Exemplo 2.65 - Declaração da secção <SingleTagFreq> .....	60
Exemplo 2.66 - Declaração da secção <Theeta>.....	60
Exemplo 2.67 - Declaração da secção <Suffixes>.....	61
API do módulo etiquetador de sentidos da biblioteca FreeLing.....	61
API do módulo desambiguador de sentidos da biblioteca FreeLing.....	61
API do módulo etiquetador do discurso oral - hmm_tagger - da biblioteca FreeLing.....	62
Exemplo 2.68 - Declaração da secção <Tag> .....	62
Exemplo 2.69 - Declaração da secção <Bigram> com a probabilidade de transição entre o estado inicial de uma frase e a primeira palavra que contém a tag 'AQ' .....	62
Exemplo 2.70 – Declaração da secção <Bigram> com a probabilidade de transição entre duas tag's consecutivas.....	63
Exemplo 2.71 - Declaração da secção <Trigram> com a probabilidade de uma palavra com a tag 'NC' aparecer após uma sequência '0.AQ' .....	63
Exemplo 2.72 - Declaração da secção <Trigram> com a probabilidade de uma palavra com a tag 'SP' aparecer após duas palavras com as tag's 'DA' e 'NC' .....	63
Exemplo 2.73 - Declaração da secção <Initial> com a probabilidade de uma frase iniciar com um determinante. ....	63
Exemplo 2.74 - Declaração da secção <Initial> com a probabilidade de uma frase iniciar com uma tag desconhecida. ....	64
Exemplo 2.75 - Declaração da secção <Word>.....	64
Exemplo 2.76 - Declaração da secção <Smoothing>.....	64

Exemplo 2.77 - Declaração da secção <Forbidden> .....	64
API do módulo etiquetador do discurso oral - relax_tagger - da biblioteca FreeLing .....	65
Exemplo 2.78 - Declaração da secção SET .....	65
Exemplo 2.79 - Declaração de restrições. ....	66
API do módulo classificador de entidades nomeadas da biblioteca FreeLing .....	66
API do módulo parser gráfico- da biblioteca FreeLing.....	66
API do módulo parser de dependências- da biblioteca FreeLing.....	67
Exemplo 2.80 - Definição de regra .....	67
Exemplo 2.81 - Definição de regra .....	67
Exemplo 2.82 - Definição de regra .....	68
Exemplo 2.83 - Definição de regra .....	68
API do módulo de resolução de correferências da biblioteca FreeLing.....	68
Exemplo 2.84 - Declaração da secção <ABModel> .....	68
Exemplo 2.85 – Declaração da secção <SemDB>.....	69
API do módulo da base de dados semântica da biblioteca FreeLing.....	69
Exemplo 2.86 – Declaração do ficheiro que contém o dicionário de sentidos.....	69
Exemplo 2.87 - Anotação de um texto com o PoS Tree-Tagger .....	70
Exemplo 2.88 - Lematização de uma palavra com o PtStemmer.....	70
Exemplo 2.89 - Classificação de entidades mencionadas .....	71
anexo IV – Conjugadores verbais .....	73
Exemplo 2.90 - Conjugação do verbo ser utilizando o conjugue .....	75
Exemplo 2.91 - Conjugação do verbo ser através do serviço disponível online em <a href="http://lxcenter.di.fc.ul.pt/pt/LXServicesConjugatorPT.html">http://lxcenter.di.fc.ul.pt/pt/LXServicesConjugatorPT.html</a> .....	76
Anexo V – Extractores de N-Gramas.....	79
Exemplo 2.92 - Listagem de bigramas num ficheiro introduzido como input .....	81
Exemplo 2.93 - Listagem de trigramas num ficheiro introduzido como input .....	82
Exemplo 2.94 - Listagem das combinações de um ou mais tokens numa ou mais posições fixas .....	83
Exemplo 2.95 - Funcionamento do programa statistic.pl .....	84
Exemplo 2.96 – Combinação de contagens de frequências a partir de um ficheiro de bigramas gerado pelo programa ‘count.pl’ .....	85
Exemplo 2.97 - Ordenação alfabética de um ficheiro de bigramas criado pelo programa ‘count.pl’ .....	87
Exemplo 2.98 – Execução do programa ‘huge-count.pl’ sobre o ficheiro ‘test.txt’ introduzido.....	87



Exemplo 2.99 - Eliminação de bigramas cujas frequências se situam entre um dado limite inferior e um dado limite superior.....	101
Exemplo 2.100 - Junção de ficheiros de bigramas existentes no directório /test .....	102
Exemplo 2.101 - Ordenação alfabética de um ficheiro de bigramas .....	108
Exemplo 2.102 - Partição de um ficheiro de bigramas em ficheiros mais pequenos .....	112
Exemplo 2.103 - Lista das co-ocorrências de um ficheiro de bigramas.....	117
Exemplo 2.104 - Cálculo do coeficiente de correlação relativamente a duas estatísticas distintas para o mesmo conjunto de dados.....	117
Anexo VI – Ferramentas especializadas .....	165
Exemplo 2.105 - Gramática com quatro regras baseadas em dependências.....	167
Exemplo 2.106 - Representação simples baseada em dependências de uma frase .....	167
Exemplo 2.107 - Análise de discurso.....	168
Exemplo 2.108 - Etiquetagem de um texto.....	170
Exemplo 2.109 - Análise estatística de corpora .....	173
Exemplo 2.110 - Funcionamento da ferramenta lingua.sh existente no package LinguaToolkit .....	174
Exemplo 2.111 - Análise de um texto dado como input com o Multilingual Dependency Parser .....	175
Exemplo 2.112 - Ilustração da utilização do extractor de termos multilíngua galeXtra .....	175
Exemplo 2.113 – Utilização da ferramenta O Constructor .....	176
Exemplo 2.114 – Utilização do SciPo para a escrita de um resumo .....	176
Exemplo 2.115 - Identificação da língua na qual um dado documento se encontra escrito .....	177
Exemplo 2.116 - Funcionamento do concordanciador do projecto TextQuim.....	177
Exemplo 2.117 – Funcionamento do listador de palavras do package TextQuim .....	178
Exemplo 2.118 - Funcionamento da ferramenta n-gramas.....	180
Exemplo 2.119 - Funcionamento da ferramenta concordanciador alinhado .....	182
Anexo VII – Tradução automática .....	183
Exemplo 2.120 - Tradução de um texto de Português para Espanhol utilizado o Apertium .....	185
Exemplo 2.121 - Tradução de um texto de Português para Espanhol utilizado o Galician- Portuguese translator .....	186
Anexo VIII – Processo de instalação do CMS Joomla!.....	187
Anexo IX – Ficheiro de estilização do menu principal do site .....	193
ANEXO X – Ficheiro de estilização da template .....	197



## **ANEXO I – AJUDA À REDACÇÃO**



## EXEMPLO 2.1 - DEFINIÇÃO DE ENTRADAS NO DICIONÁRIO DO JSPELL

Considerar/CAT=v,T=inf,TR=\_/XYPLcv/  
 o/CAT=art,G=m,N=s/p/  
 seguinte/CAT=a\_nc,G=2,N=s/pm/ comentário...  
 exemplo/CAT=nc,G=m,N=s/p/

Como o exemplo ilustra cada linha é constituída por quatro partes distintas:

1. lema;
2. lista de propriedades morfológicas;
3. conjunto de identificadores de regras de derivação ou flexão;
4. comentários.

Analisemos a última linha do nosso exemplo:

- CAT=nc - a categoria é *nome comum*;
- G=m - o género é *masculino*;
- N=s - o número é *singular*;
- p - indica que a palavra pode formar plural (*exemplos*) utilizando a regra 'p'mostrada abaixo.

## EXEMPLO 2.2 - UTILIZAÇÃO DO JSPELL A PARTIR DA LINHA DE COMANDO

```
$ jspell -d português -a
@(#) International Jspell Version 1.00b1, 11/07/2001
coreana
* coreana 0 : lex(coreano, [CAT=adj,N=s,G=m], [], [G=f], [])
```

Neste exemplo, após a primeira linha que nos indica o número da versão do programa, temos mais duas linhas:

- uma em que indicamos a palavra que pretendemos analisar;
- outra onde o jspell nos fornece informação acerca da palavra pretendida (o asterisco que aparece no início desta linha indica que a palavra existe no dicionário, que é gerada a partir do lema(adjectivo) 'coreano' modificando o género para o feminino).

### **EXEMPLO 2.3 - OBTENÇÃO DO ADJECTIVO CORDIAIS A PARTIR DE CORDIAL**

cordiais

\* cordiais 0 : lex(cordial, [CAT=adj,N=s,G=\_], [], [N=p], [])

Este exemplo mostra que ‘cordiais’ pode ser obtido a partir do adjectivo ‘cordial’ (sem género), alterando o número para o plural.

### **EXEMPLO 2.4 - CONJUGAÇÃO DO VERBO CORDEAR NA SEGUNDA PESSOA DO PLURAL DO PRESENTE DO INDICATIVO**

cordeais

⤴ cordeais 0 : lex(cordear, [CAT=v,T=inf,Tr=t], [], [P=2,N=p,T=p], [])

⤴

Neste caso, ‘cordeais’ corresponde à conjugação do verbo ‘cordear’ na segunda pessoa do plural, do presente do indicativo.

### **EXEMPLO 2.5 - RESPOSTA DO JSPELL A CASOS DE PALAVRAS QUE NÃO EXISTEM NO DICIONÁRIO**

crodeais

& crodeais 0:

CORDEIAS : lex(cordear, [CAT=v,T=inf,Tr=t], [], [P=2,N=p,T=p], [])

RODEIAS : lex(rodear, [CAT=v,T=inf,Tr=\_], [], [P=2,N=p,T=p], [])

Este exemplo permite-nos verificar o que acontece quando pedimos ao Jspell informações acerca de palavras que não existem no dicionário (é colocado um ‘& ’ antes da palavra desconhecida, são procuradas palavras aproximada àquela que procuramos e é-nos dada informação acerca dessas palavras).

## EXEMPLO 2.6 - UTILIZAÇÃO DO ANALISADOR MORFOLÓGICO DO WEBSPELL







## **ANEXO II – ALINHADORES**



## EXEMPLO 2.7 - ALINHAMENTO DE TEXTOS UTILIZANDO O ALINHADOR ONLINE CEPRIL

\*\*\* Link: 1 - 1 \*\*\*

Turin European Council 29 March 1996 Presidency Conclusions .EOS

EUROPÄISCHER RAT TURIN 29. MÄRZ 1996 SCHLUSSFOLGERUNGEN DES VORSITZES .EOS

.EOP

\*\*\* Link: 1 - 1 \*\*\*

The European Council began its proceedings by exchanging ideas with Mr Klaus Hänsch, President of the European Parliament, on the main subjects for discussion at this meeting.

.EOS

Der Europäische Rat hat zunächst einen Gedankenaustausch mit dem Präsidenten des Europäischen Parlaments, Herrn Klaus Hänsch, über die wichtigsten auf dieser Tagung zur Erörterung anstehenden Themen geführt. .EOS

.EOP

### Interpretação dos resultados

- \*\*\* início de um trecho alinhado;
- *Link*:
  - 1 - 1 Um trecho do primeiro texto corresponde a exatamente um trecho do segundo texto.
  - 0 - 1: Um trecho do segundo texto não possui equivalente no primeiro texto.
  - 1 - 0: Um trecho do primeiro texto não possui equivalente no segundo texto.
  - 2 - 1: Dois trechos do primeiro texto são equivalentes a um trecho do segundo texto.
  - 1 - 2: Um trecho do primeiro texto é equivalente a dois trechos do segundo texto.
  - 2 - 2: Dois trechos do primeiro texto são equivalentes a dois trechos do segundo texto.
- *The European Council ...* Início do trecho do primeiro texto.
- *Der Europäische Rat ...* Início do trecho do segundo texto.
- *.EOS* Marcador de final de unidade menor (no caso, frase).
- *.EOP* Marcador de final de unidade maior (no caso, parágrafo).

**EXEMPLO 2.8 - ALINHAMENTO DE DOIS TEXTOS UTILIZANDO O VISUALIHLA**Textos originais

Cuidado , Popeye : espinafre é um perigo .	Tiembla Popeye : la espinaca es un peligro .
Há pessoas que evitam ingerir carne e laticínios e , para assegurar o necessário suprimento de ferro e cálcio , substituem esses alimentos por vegetais - em geral os verde - escuros como espinafre , couve - manteiga e brócolis .	Existen personas que evitan ingerir carne y lácteos y , para asegurarse el necesario aporte de hierro y calcio , reemplazan estos alimentos por vegetales - en general los de color verde oscuro , como las espinacas , coles y brócolis .
Contudo , segundo Jocelem Mastrodi Salgado , da Escola Superior de Agricultura Luiz de Queiroz ( Esalq ) , nem sempre esses minerais são aproveitados pelo organismo quando se consome verduras : algumas delas contêm substâncias tóxicas que impedem a absorção dos nutrientes .	Con todo , según Jocelem Mastrodi Salgado , de la Escuela Superior de Agricultura Luiz de Queiroz ( Esalq ) , no siempre esos minerales son aprovechados por el organismo cuando se consumen verduras : algunas de éstas contienen sustancias tóxicas que impiden la absorción de los nutrientes .
Pior : podem causar efeitos tóxicos , " o que ocorre com a grande ingestão de espinafre " , segundo Salgado .	Y lo que es peor : pueden causar efectos tóxicos : " es lo que ocurre con una gran ingestión de espinaca " , dice Salgado .
A pesquisadora revela que os ácidos fítico e oxálico , este último presente em grande quantidade no espinafre , têm a capacidade de ligar - se ao cálcio e ao ferro , " fazendo com que fiquem indisponíveis para serem absorvidos pelo nosso corpo " e acabem por ser eliminados nas fezes .	La investigadora revela que los ácidos fítico y oxálico , este último presente en gran cantidad en la espinaca , tienen la capacidad de combinarse con el calcio y el hierro , " haciendo que se tornen indisponibles para su absorción por nuestro cuerpo " , y acaben por siendo eliminados en las heces .
E mais : comer espinafre numa refeição " pode reduzir a biodisponibilidade do cálcio " contido em outros alimentos .	Es más : comer espinaca en una comida " puede reducir la biodisponibilidad del calcio " contenido en otros alimentos .
Quanto aos efeitos tóxicos , ela lembra que em 1951 , nos Estados Unidos , " a ingestão de leite batido com espinafre ( o objetivo era enriquecer a bebida com ferro ) causou a morte de crianças recém - nascidas " .	Con relación a los efectos tóxicos , Salgado recuerda que , en 1951 , en Estados Unidos , " la ingestión de leche batida con espinaca ( el objetivo era enriquecer la bebida con hierro ) causó la muerte de niños recién nacidos " .
O fato , no entanto , acabou encoberto na época .	No obstante , ese hecho acabó siendo encubierto en la época .
Uma pesquisa orientada por Salgado mostrou que cobaias em cuja dieta acrescentaram folhas de espinafre secas e moídas morreram na primeira semana com hemorragias e toxidade renal .	Una investigación orientada por Salgado mostró que cobayos en cuya dieta se incorporaron hojas de espinaca secas y molidas , murieron durante la primera semana con hemorragias e intoxicación

	renal .
Folhas cozidas tiveram o mesmo efeito .	Las hojas cocidas tuvieron el mismo efecto .
O estudo também avaliou a couve - manteiga e a couve - flor e não constatou nenhum efeito tóxico sobre os animais .	El estudio también analizó coles y coliflor y no constató ningún efecto tóxico sobre los animales .
Em vista disso , Salgado aconselha a substituição do espinafre : " Couve , brócolis , folha de mostarda , agrião , folhas de cenoura , beterraba e couve - flor , e leguminosas como feijões , ervilha , lentilha e soja são as melhores opções para quem quer consumir fontes alternativas de cálcio e ferro " .	En vista de ello , Salgado aconseja la sustitución de la espinaca : " Coles , brócolis , hojas de mostaza , berro , hojas de zanahoria , remolacha y coliflor , y leguminosas como frijoles , arvejas , lentejas y soja son las mejores opciones para quienes quieren consumir fuentes alternativas de calcio y hierro " .

### Textos analisados

<s snum=1>Cuidado:1 ,:3 Popeye:2 ::9 espinafre:0 é:0 um:0 perigo:0 .:0 </s>	<s snum=1>Tiembla:1 Popeye:3 ::2 la:0 espinaca:0 es:0 un:0 peligro:0 .:4 </s>
<s snum=2>Há:1 pessoas:2 que:3 evitam:4 ingerir:5 carne:6 e:7_8_9 laticínios:7_8_9 e:7_8_9 ,:10 para:11 assegurar:12_13 o:12_13 necessário:14 suprimento:15 de:16 ferro:17_18_19 e:17_18_19 cálcio:17_18_19 ,:20 substituem:21_22 esses:21_22 alimentos:23 por:24 vegetais:25 -:26 em:27 geral:28 os:29_30_31 verde:32 -:34 escuros:0 como:35 espinafre:36_37 ,:38 couve:39_40 -:42 manteiga:0 e:0 brócolis:41 .:0 </s>	<s snum=2>Existen:1 personas:2 que:3 evitan:4 ingerir:5 carne:6 y:7_8_9 lácteos:7_8_9 y:7_8_9 ,:10 para:11 asegurarse:12_13 el:12_13 necesario:14 aporte:15 de:16 hierro:17_18_19 y:17_18_19 calcio:17_18_19 ,:20 reemplazan:21_22 estos:21_22 alimentos:23 por:24 vegetales:25 -:26 en:27 general:28 los:29 de:29 color:29 verde:30 oscuro:0 ,:31 como:33 las:34 espinacas:34 ,:35 coles:36 y:36 brócolis:40 .:37 </s>
<s snum=3>Contudo:1_2 ,:3 segundo:4 Jocelem:5 Mastrodi:6 Salgado:7 ,:8 da:10_11 Escola:10_11 Superior:12 de:9 Agricultura:14 Luiz:15 de:13 Queiroz:17 (:18 Esalq:19 ):20 ,:21 nem:22_23_24 sempre:22_23_24 esses:22_23_24 minerais:25 são:26 aproveitados:27 pelo:28_29 organismo:30 quando:31 se:32 consume:33 verduras:34 ::35 algumas:36 delas:37_38_39 contêm:37_38_39 substâncias:40 tóxicas:41 que:42 impedem:43_44 a:43_44 absorção:45 dos:46_47 nutrientes:48 .:49 </s>	<s snum=3>Con:1 todo:1 ,:2 según:3 Jocelem:4 Mastrodi:5 Salgado:6 ,:7 de:11 la:8_9 Escuela:8_9 Superior:10 de:14 Agricultura:12 Luiz:13 de:0 Queiroz:15 (:16 Esalq:17 ):18 ,:19 no:20_21_22 siempre:20_21_22 esos:20_21_22 minerales:23 son:24 aprovechados:25 por:26 el:26 organismo:27 cuando:28 se:29 consumen:30 verduras:31 ::32 algunas:33 de:34_35 éstas:34_35 contienen:34_35 sustancias:36 tóxicas:37 que:38 impiden:39_40 la:39_40 absorción:41 de:42 los:42 nutrientes:43 .:44 </s>
<s snum=4>Pior:1_2_3_4_5 ::6 podem:7 causar:8 efeitos:9 tóxicos:10 ,:24 ":12 o:13_14 que:15 ocorre:16 com:17_18	<s snum=4>Y:1 lo:1 que:1 es:1 peor:1 ::2 pueden:3 causar:4 efectos:5 tóxicos:6 ::19 ":8 es:9 lo:9 que:10 ocurre:11 con:12_13

a:17_18 grande:19 ingestão:20 de:21 espinafre:22 ":23 ,:11 segundo:25 Salgado:26 .:27 </s>	una:12_13 gran:14 ingestión:15 de:16 espinaca:17 ":18 ,:7 dice:20 Salgado:21 .:22 </s>
<s snum=5>A:1_2 pesquisadora:1_2 revela:3 que:4 os:5 ácidos:6 fítico:7 e:8 oxálico:9 ,:10 este:11 último:12 presente:13 em:14 grande:15 quantidade:16_17_18_19 no:16_17_18_19 espinafre:16_17_18_19 ,:20 têm:21_22 a:21_22 capacidade:23 de:24 ligar:25_26_27_28_29_30_31 -:32 se:36 ao:37 cálcio:37 e:37 ao:37 ferro:37 ,:46 ":33 fazendo:34 com:34 que:35 fiquem:0 indisponíveis:38 para:39 serem:40_41_42_43_44 absorvidos:40_41_42_43_44 pelo:40_41_42_43_44 nosso:40_41_42_43_44 corpo:40_41_42_43_44 ":45 e:47 acabem:48 por:49 ser:50 eliminados:51 nas:52_53 fezes:54 .:55 </s>	<s snum=5>La:1_2 investigadora:1_2 revela:3 que:4 los:5 ácidos:6 fítico:7 y:8 oxálico:9 ,:10 este:11 último:12 presente:13 en:14 gran:15 cantidad:16_17_18 en:16_17_18 la:16_17_18 espinaca:16_17_18 ,:19 tienen:20_21 la:20_21 capacidad:22 de:23 combinarse:24 con:24 el:24 calcio:24 y:24 el:24 hierro:24 ,:25 ":33 haciendo:34_35 que:36 se:26 tornen:27_28_29_30_31 indisponibles:38 para:39 su:40_41_42_43_44 absorción:40_41_42_43_44 por:40_41_42_43_44 nuestro:40_41_42_43_44 cuerpo:40_41_42_43_44 ":45 ,:32 y:46 acaben:47 por:48 siendo:49 eliminados:50 en:51 las:51 heces:52 .:53 </s>
<s snum=6>E:1_2 mais:1_2 ::3 comer:4 espinafre:5_6_7_8 numa:5_6_7_8 refeição:5_6_7_8 ":9 pode:10 reduzir:11 a:12 biodisponibilidade:13 do:14_15 cálcio:14_15 ":16 contido:17 em:18_19 outros:18_19 alimentos:20 .:21 </s>	<s snum=6>Es:1_2 más:1_2 ::3 comer:4 espinaca:5_6_7 en:5_6_7 una:5_6_7 comida:5_6_7 ":8 puede:9 reducir:10 la:11 biodisponibilidad:12 del:13_14 calcio:13_14 ":15 contenido:16 en:17_18 otros:17_18 alimentos:19 .:20 </s>
<s snum=7>Quanto:1_2 aos:1_2 efeitos:5 tóxicos:6 ,:7 ela:8_9 lembra:8_9 que:10 em:12 1951:13 ,:11 nos:15 Estados:16 Unidos:17 ,:14 ":19 a:3 ingestão:21 de:22 leite:23 batido:24 com:25_26 espinafre:25_26 (:27 o:28 objetivo:29 era:30 enriquecer:31 a:32 bebida:33 com:34_35 ferro:34_35 ):36 causou:37_38_39 a:37_38_39 morte:37_38_39 de:40 crianças:41_42 recém:41_42 -:44 nascidas:43 ":45 .:18 </s>	<s snum=7>Con:1_2 relación:1_2 a:17 los:0 efectos:3 tóxicos:4 ,:5 Salgado:6_7 recuerda:6_7 que:8 ,:11 en:9 1951:10 ,:15 en:12 Estados:13 Unidos:14 ,:43 ":16 la:0 ingestión:18 de:19 leche:20 batida:21 con:22_23 espinaca:22_23 (:24 el:25 objetivo:26 era:27 enriquecer:28 la:29 bebida:30 con:31_32 hierro:31_32 ):33 causó:34_35_36 la:34_35_36 muerte:34_35_36 de:37 niños:38_39 recién:38_39 nacidos:41 ":40 .:42 </s>
<s snum=8>O:2 fato:2 ,:3 no:1 entanto:4_5_6_7 ,:12 acabou:9_10 encoberto:8 na:0 época:11 .:0 </s>	<s snum=8>No:4 obstante:1_2 ,:3 ese:5 hecho:5 acabó:5 siendo:5 encubierto:8 en:7 la:7 época:10 .:6 </s>
<s snum=9>Uma:1_2 pesquisa:1_2 orientada:3 por:4 Salgado:5 mostrou:6 que:7 cobaias:8_9 em:8_9 cuja:10 dieta:11 acrescentaram:12_13_14 folhas:12_13_14 de:15 espinafre:16 secas:17 e:18_19 moídas:18_19 morreram:18_19 na:18_19 primeira:24 semana:25 com:26 hemorragias:27 e:28 toxidade:29 renal:30	<s snum=9>Una:1_2 investigación:1_2 orientada:3 por:4 Salgado:5 mostró:6 que:7 cobayos:8_9 en:8_9 cuya:10 dieta:11 se:12_13 incorporaron:12_13 hojas:12_13 de:14 espinaca:15 secas:16 y:17_18_19_20 molidas:17_18_19_20 ,:0 murieron:0 durante:0 la:0 primera:21 semana:22 con:23 hemorragias:24 e:25 intoxicación:26

.:31 </s>	renal:27 .:28 </s>
<s snum=10>Folhas:1_2 cozidas:3 tiveram:4_5_6 o:4_5_6 mesmo:4_5_6 efeito:7 .:8 </s>	<s snum=10>Las:1 hojas:1 cocidas:2 tuvieron:3_4_5 el:3_4_5 mismo:3_4_5 efecto:6 .:7 </s>
<s snum=11>O:1 estudo:2 também:3_4_5_6_7_8 avaliou:3_4_5_6_7_8 a:3_4_5_6_7_8 couve:3_4_5_6_7_8 -:17 manteiga:11 e:11 a:11 couve:11 -:0 flor:0 e:0 não:9 constatou:10 nenhum:0 efeito:12 tóxico:13 sobre:14 os:15 animais:16 .:0 </s>	<s snum=11>El:1 estudio:2 también:3_4_5_6 analizó:3_4_5_6 coles:3_4_5_6 y:3_4_5_6 coliflor:3_4_5_6 y:3_4_5_6 no:15 constató:16 ningún:8_9_10_11 efecto:18 tóxico:19 sobre:20 los:21 animales:22 .:7 </s>
<s snum=12>Em:1 vista:2 disso:3_4 ,:5 Salgado:6 aconselha:7 a:8 substituição:9 do:10_11_12 espinafre:10_11_12 .:13 ":14 Couve:15 ,:16 brócolis:17 ,:18 folha:19 de:20 mostarda:21 ,:22 agrião:0 ,:24 folhas:25 de:26 cenoura:27 ,:28 beterraba:29_30_31 e:29_30_31 couve:29_30_31 -:32 flor:33 ,:37 e:36 leguminosas:34 como:35 feijões:38 ,:39 ervilha:40_41 ,:57 lentilha:43_44 e:43_44 soja:42 são:0 as:0 melhores:45 opções:46 para:47 quem:48_49 quer:48_49 consumir:50 fontes:51 alternativas:52 de:53 cálcio:0 e:0 ferro:23 ":58 .:0 </s>	<s snum=12>En:1 vista:2 de:3 ello:3 ,:4 Salgado:5 aconseja:6 la:7 sustitución:8 de:9_10 la:9_10 espinaca:9_10 .:11 ":12 Coles:13 ,:14 brócolis:15 ,:16 hojas:17 de:18 mostaza:19 ,:20 berro:56 ,:22 hojas:23 de:24 zanahoria:25 ,:26 remolacha:27_28_29 y:27_28_29 coliflor:27_28_29 ,:30 y:31 leguminosas:34 como:35 frijoles:33 ,:32 arvejas:36 ,:37 lentejas:38 y:38 soja:42 son:40_41 las:40_41 mejores:45 opciones:46 para:47 quienes:48_49 quieren:48_49 consumir:50 fuentes:51 alternativas:52 de:53 calcio:0 y:0 hierro:0 ":39 .:57 </s>

## EXEMPLO 2.9 - UTILIZAÇÃO DO VISUALTCA NO ALINHAMENTO DE TEXTOS

### Textos originais

<p>Cuidado, Popeye: espinafre é um perigo. Há pessoas que evitam ingerir carne e laticínios e, para assegurar o necessário suprimento de ferro e cálcio, substituem esses alimentos por vegetais - em geral os verde-escuros como espinafre, couve-manteiga e brócolis. Contudo, segundo Joclem Mastrodi Salgado, da Escola Superior de Agricultura Luiz de Queiroz (Esalq), nem sempre esses minerais são aproveitados pelo organismo quando se</p>	<p>Tiembla Popeye: la espinaca es un peligro. Existen personas que evitan ingerir carne y lácteos y, para asegurarse el necesario aporte de hierro y calcio, reemplazan estos alimentos por vegetales -en general los de color verde oscuro, como las espinacas, coles y brócolis. Con todo, según Joclem Mastrodi Salgado, de la Escuela Superior de Agricultura Luiz de Queiroz (Esalq), no siempre esos minerales son aprovechados por el organismo cuando se consumen</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<p>consume verduras: algumas delas contêm substâncias tóxicas que impedem a absorção dos nutrientes. Pior: podem causar efeitos tóxicos, "o que ocorre com a grande ingestão de espinafre", segundo Salgado.</p> <p>A pesquisadora revela que os ácidos fítico e oxálico, este último presente em grande quantidade no espinafre, têm a capacidade de ligar-se ao cálcio e ao ferro, "fazendo com que fiquem indisponíveis para serem absorvidos pelo nosso corpo" e acabem por ser eliminados nas fezes. E mais: comer espinafre numa refeição "pode reduzir a biodisponibilidade do cálcio" contido em outros alimentos. Quanto aos efeitos tóxicos, ela lembra que em 1951, nos Estados Unidos, "a ingestão de leite batido com espinafre (o objetivo era enriquecer a bebida com ferro) causou a morte de crianças recém-nascidas".</p> <p>O fato, no entanto, acabou encoberto na época. Uma pesquisa orientada por Salgado mostrou que cobaias em cuja dieta acrescentaram folhas de espinafre secas e moídas morreram na primeira semana com hemorragias e toxidade renal. Folhas cozidas tiveram o mesmo efeito. O estudo também avaliou a couve-manteiga e a couve-flor e não constatou nenhum efeito tóxico sobre os animais. Em vista disso, Salgado aconselha a substituição do espinafre: "Couve, brócolis, folha de mostarda, agrião, folhas de cenoura, beterraba e couve-flor, e leguminosas como feijões, ervilha, lentilha e soja são as melhores opções para quem quer consumir fontes alternativas de cálcio e ferro".</p>	<p>verduras: algunas de éstas contienen sustancias tóxicas que impiden la absorción de los nutrientes.</p> <p>Y lo que es peor: pueden causar efectos tóxicos: "es lo que ocurre con una gran ingestión de espinaca", dice Salgado. La investigadora revela que los ácidos fítico y oxálico, este último presente en gran cantidad en la espinaca, tienen la capacidad de combinarse con el calcio y el hierro, "haciendo que se tornen indisponibles para su absorción por nuestro cuerpo", y acaben por siendo eliminados en las heces.</p> <p>Es más: comer espinaca en una comida "puede reducir la biodisponibilidad del calcio" contenido en otros alimentos. Con relación a los efectos tóxicos, Salgado recuerda que, en 1951, en Estados Unidos, "la ingestión de leche batida con espinaca (el objetivo era enriquecer la bebida con hierro) causó la muerte de niños recién nacidos". No obstante, ese hecho acabó siendo encubierto en la época.</p> <p>Una investigación orientada por Salgado mostró que cobayos en cuya dieta se incorporaron hojas de espinaca secas y molidas, murieron durante la primera semana con hemorragias e intoxicación renal. Las hojas cocidas tuvieron el mismo efecto. El estudio también analizó coles y coliflor y no constató ningún efecto tóxico sobre los animales. En vista de ello, Salgado aconseja la sustitución de la espinaca: "Coles, brócolis, hojas de mostaza, berro, hojas de zanahoria, remolacha y coliflor, y leguminosas como fríjoles, arvejas, lentejas y soja son las mejores opciones para quienes quieren consumir fuentes alternativas de calcio y hierro".</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Textos analisados

Cuidado, Popeye: espinafre é um perigo.	Tiembra Popeye: la espinaca es un peligro.
<p>Há pessoas que evitam ingerir carne e laticínios e, para assegurar o necessário suprimento de ferro e cálcio, substituem esses alimentos por vegetais - em geral os verde-escuros como espinafre, couve-manteiga e brócolis. Contudo, segundo Joicelem Mastrodi Salgado, da Escola Superior de Agricultura Luiz de Queiroz (Esalq), nem sempre esses minerais são aproveitados pelo organismo quando se consome verduras: algumas delas contêm substâncias tóxicas que impedem a absorção dos nutrientes. Pior: podem causar efeitos tóxicos, "o que ocorre com a grande ingestão de espinafre", segundo Salgado.</p>	<p>Existen personas que evitan ingerir carne y lácteos y, para asegurarse el necesario aporte de hierro y calcio, reemplazan estos alimentos por vegetales -en general los de color verde oscuro, como las espinacas, coles y brócolis. Con todo, según Joicelem Mastrodi Salgado, de la Escuela Superior de Agricultura Luiz de Queiroz (Esalq), no siempre esos minerales son aprovechados por el organismo cuando se consumen verduras: algunas de éstas contienen sustancias tóxicas que impiden la absorción de los nutrientes.</p>
<p>A pesquisadora revela que os ácidos fítico e oxálico, este último presente em grande quantidade no espinafre, têm a capacidade de ligar-se ao cálcio e ao ferro, "fazendo com que fiquem indisponíveis para serem absorvidos pelo nosso corpo" e acabem por ser eliminados nas fezes. E mais: comer espinafre numa refeição "pode reduzir a biodisponibilidade do cálcio" contido em outros alimentos. Quanto aos efeitos tóxicos, ela lembra que em 1951, nos Estados Unidos, "a ingestão de leite batido com espinafre (o objetivo era enriquecer a bebida com ferro) causou a morte de crianças recém-nascidas".</p>	<p>Y lo que es peor: pueden causar efectos tóxicos: "es lo que ocurre con una gran ingestión de espinaca", dice Salgado. La investigadora revela que los ácidos fítico y oxálico, este último presente en gran cantidad en la espinaca, tienen la capacidad de combinarse con el calcio y el hierro, "haciendo que se tornen indisponibles para su absorción por nuestro cuerpo", y acaben por siendo eliminados en las heces.</p> <p>Es más: comer espinaca en una comida "puede reducir la biodisponibilidad del calcio" contenido en otros alimentos. Con relación a los efectos tóxicos, Salgado recuerda que, en 1951, en Estados Unidos, "la ingestión de leche batida con espinaca (el objetivo era enriquecer la bebida con hierro) causó la muerte de niños recién nacidos". No obstante, ese hecho acabó siendo encubierto en la época.</p>
<p>O fato, no entanto, acabou encoberto na época. Uma pesquisa orientada por Salgado mostrou que cobaias em cuja dieta acrescentaram folhas de espinafre secas e moídas morreram na primeira semana com</p>	<p>Una investigación orientada por Salgado mostró que cobayos en cuya dieta se incorporaron hojas de espinaca secas y molidas, murieron durante la primera semana con hemorragias e intoxicación</p>

<p>hemorragias e toxidade renal. Folhas cozidas tiveram o mesmo efeito. O estudo também avaliou a couve-manteiga e a couve-flor e não constatou nenhum efeito tóxico sobre os animais. Em vista disso, Salgado aconselha a substituição do espinafre: "Couve, brócolis, folha de mostarda, agrião, folhas de cenoura, beterraba e couve-flor, e leguminosas como feijões, ervilha, lentilha e soja são as melhores opções para quem quer consumir fontes alternativas de cálcio e ferro".</p>	<p>renal. Las hojas cocidas tuvieron el mismo efecto. El estudio también analizó coles y coliflor y no constató ningún efecto tóxico sobre los animales. En vista de ello, Salgado aconseja la sustitución de la espinaca: "Coles, brócolis, hojas de mostaza, berro, hojas de zanahoria, remolacha y coliflor, y leguminosas como frijoles, arvejas, lentejas y soja son las mejores opciones para quienes quieren consumir fuentes alternativas de calcio y hierro".</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## **ANEXO III – ANALISADORES**







32 tador, mas uma forma de modelar o seu desenho por computador. Ele nos dá a  
possibilidade de verifica

33 rios via Web e Seagate Crystal Reports Web em seu computador..

34 r, Netscape Navigator, ou Mosaic instalado em seu computador.

35 tenha os seguintes aplicativos instalados em seu computador.

36 rios via Web e Seagate Crystal Reports Web em seu computador..

37 um dos seguintes servidores Web instalados em seu computador:

38 Mosaic instalado em seu computador.

39 Como o protótipo local está contido no seu computador, desde o início do  
ciclo de desenvolvim

40 onfirme a capacidade do sistema de conectar o seu computador à fonte de  
dados. Lance mão de um utili

41 onfirme a capacidade do sistema de conectar o seu computador à fonte de  
dados. Lance mão de um utili

42 selecionou e da velocidade do computador. Depois de instalados os  
pacotes,

43 quais portas de impressora estão disponíveis no computador. Selecione

44 Informe o nome do computador em que a impressora está  
diretamente li

45 no campo "Computador remoto" e o nome da fila no computador remoto que é

46 computador. Aparecerá uma tela contendo o boot e o

47 rnativamente, se você tem MS-DOS instalado em seu computador, você

48 CRHL. Dependendo da configuração do computador e do tipo de  
instalação

49 Nome da máquina: O nome de seu computador, um nome qualquer.

50 cionados são carregados através da rede até o seu computador, e são

51 fornecidos a seu computador através de um sistema servidor  
de NFS.

52 desejar instalar o Conectiva Red Hat Linux em um computador que

53 uso. Se você comprar um computador com software pré-instalado,  
é bem

54 partições FAT32. Se o seu computador tiver uma partição FAT32,  
pode-se

55 System (ou BIOS) do próprio computador.

56 selecionou e da velocidade do computador. Depois de instalados os pacotes,

57 o projeto estiver terminado, basta instalá-lo no computador servidor de arquivos para uso imediato.

58 computador. Não fique com medo de experimentar e b

59 izar com uma boa aparência em qualquer sistema de computador. (Em termos técnicos, isso significa

60 Informe o nome do computador em que a impressora está ligada no camp

61 sistema de inicialização de seu computador não seja alterado. Com um disquete

62 ido, que é automaticamente carregado pelo BIOS do computador. Caso o LILO

63 configuração do disco do computador. Ela não foi desenvolvida para ser usad

64 ndições de trocar os pentes de memória RAM de seu computador para

65 Reinicialize o computador.

66 Supondo um computador caseiro que não terá centenas de usuári

67 e da velocidade do computador.

68 totalidade dos dados armazenados no computador serão apagados. Esta decisão

69 tentar detectar adaptadores SCSI em seu computador. Em alguns casos, o programa

70 conectado ao seu computador. Ao escolher Sim, a próxima caixa de di

71 configuração do disco do computador. Ela não foi desenvolvida para ser usad

72 totalidade dos dados armazenados no computador serão apagados. Esta decisão

73 e da velocidade do computador.

74 um computador portátil, ele provavelmente será um mou

75 Note que se seu computador possui algum outro sistema operacional



76 disponíveis no computador. Selecione Ok e tecl. Agora vá até  
a

77 Reinicie seu computador

78 computador e também no servidor onde você  
hospedar

79 seu computador

80 instaladas no seu computador

81 computador

82 FrontPage 2000 estiverem instaladas no seu computador

83 o perl.exe no seu computador ( informações de como

84 seu computador ;

85 determina uma fonte que não esteja disponível no computador do

86 m pode trafegar em alta velocidade entre qualquer computador conectado a essa  
rede. E por essa razão

87 ernet, usualmente significa ter uma "conta" em um computador "servidor" que  
esteja conectado à Inter

88 serviços de acesso à Internet. Essa "conta" nesse computador ligado à Internet é  
usualmente acessada

89 "lopes" representa uma identificação da pessoa no computador em que tem uma  
conta e "artnet.com.br"

90 tem uma conta e "artnet.com.br" o endereço desse computador na Internet. O  
".br" no final do endere

91 ternet. O ".br" no final do endereço indica que o computador é da rede brasileira.

92 m o FTP fazem transferência de arquivos entre seu computador local e outro  
remoto. O FTP é um dos re

93 essa conta e transferir os arquivos dela para seu computador local, ou vice-versa.  
Contudo, uma poss

94 , ou seja, todo tipo de informação em arquivos de computador. Normalmente no  
FTP anônimo o login e a

95 stas listas funcionam da seguinte forma: em algum computador na Internet,  
instala-se um programa cha

96 ciais para Gopher que podem ser instalados no seu computador caso você  
disponha de conexões TCP/IP e

97 o Sistema Remate-Ouro desenvolveu um programa de computador exclusivo  
para o mercado de leilões. De

98 tamente automatizados possuem terminais ligados a computador de grande porte, localizado no Cesec. M

99 nefícios, a substituição de papéis e rotinas pelo computador liberou funcionários para o relacioname

100 , via telex, informações imediatas fornecidas por computador, sobre os prazos de produção, recebimen

101 Como pioneiro na introdução do computador no sistem bancário brasileiro, o Brades

102 âneo e transmite instruções em áudio e vídeo, via computador.

103 s produtos e serviços que o Bradesco oferece, via computador.

104 iços, oferecendo dados atualizados diretamente do computador, de acordo com as necessidades de cada

105 eletrônica de extratos, com a emissão direta pelo computador, propiciando maior rapidez na expedição

106 dia e noite e 89 microestações de satélite ao seu computador central, na Cidade de Deus. De norte a

107 e as empresas cedentes transmitam e recebam dados computador a computador, já atinge 50% da Carteira

108 cedentes transmitam e recebam dados computador a computador, já atinge 50% da Carteira.

109 (gerenciador de linhas de comunicação, ligando ao computador central as Agências, PAB, "ATM" e dema

110 o Sistema Remate-Ouro desenvolveu um programa de computador exclusivo para o mercado de leilões. De

111 los, das palavras, das revisões, do auxílio com o computador, foram parceiros na construção de conhe

**EXEMPLO 2.11 - EXEMPLIFICAÇÃO DE UTILIZAÇÃO DA FERRAMENTA GOJOLPARSER**

Dado como *input* o texto “O primeiro fabricante mundial de «ratos» para computador, a empresa suíça Logitech, apresentou esta semana numa feira especializada.” o GojolParser produz o seguinte *output*:

1	_O	->3	DTAT	.DET->N
2	primeiro	->3	ADJC	.DET->NNCO
3	fabricante	->11	NNCO	.PLUSN->N
4	mundial	->3	ADJC	.ADJ->N
5	de	->3	PREP	P->N
6	ratos	->5	NNCO	.PREP->PREP
7	para	->6	PREP	P->NNCO
8	computador	->7	NNCO	.PREP->PREP
9	,		COMA	
10	a	->11	DTAT	.DET->N
11	empresa	->15	NNCO	.SUBJ->VRBV
12	suíça	->11	ADJC	.ADJ->NNCO
13	Logitech	->11	NNPR	.GNTV->N
14	,		COMA	
15	apresentou	root	VRBV	
16	esta	->17	DTAT	.DET->NNCO
17	semana	->15	NNCO	N->VRBV
18	numa	->15	PREP	P->VRBV
19	feira	->18	NNCO	.PREP->PREP
20	especializada	->19	VRBP	.DETVB->NNCO
21	.		STOP	

Tree output for the same :

(S  
  (.SUBJ  
    (N  
      (.PLUSN  
        (N  
          (N  
            (N  
              (.DET  
              (DTAT \_O))  
            (N  
              (.DET  
              (ADJC primeiro))  
              (NNCO fabricante)))  
            (.ADJ  
            (ADJC mundial)))  
          (P  
            (PREP de)  
            (.PREP  
            (N  
              (NNCO ratos)  
              (P  
                (PREP para)  
                (.PREP  
                (NNCO computador)))))))))  
    (COMA ,)  
    (N  
      (N  
        (.DET  
        (DTAT a))  
      (N  
        (NNCO empresa)  
      (.ADJ  
      (ADJC suíça))))

```

(.GNTV
  (NNPR Logitech))))))
(COMA ,)
(VRBV apresentou)
(N
  (.DET
    (DTAT esta))
  (NNCO semana))
(P
  (PREP numa)
  (.PREP
    (N
      (NNCO feira)
      (.DETVB
        (VP
          (VRBP especializada)
          (STOP .))))))

```

### EXEMPLO 2.12 - UTILIZAÇÃO DAS FUNÇÕES 'ATOMIZA', 'TOKENIZA' E 'TOKENIZE'

```

my @atomos = atomiza($texto); # tb chamada 'tokenize'
my $atomos_um_por_linha = tokeniza($texto);
my $atomos_separados_por_virgula =
tokeniza({fs=>','},$texto);

```

### **EXEMPLO 2.13 - UTILIZAÇÃO DA FUNÇÃO 'CQPTOKENS'**

```
cqptokens( { irs => "\n\n" }, "file" );
```

### **EXEMPLO 2.14 - UTILIZAÇÃO DA FUNÇÃO 'CQPTOKENS' COM OPÇÕES**

```
cqptokens( { enc => ":utf8"}, "file" ); # enc => charset  
# outenc => charset
```

### **EXEMPLO 2.15 - RETORNO DAS FRASES DE UM TEXTO**

```
@frases = frases($texto);
```

### **EXEMPLO 2.16 - RETORNO DE UM TEXTO COM UMA FRASE POR LINHA**

```
$frases = separa_frases($texto);
```

### **EXEMPLO 2.17 - APLICAÇÃO DE UMA ETIQUETA XML A CADA FRASE**

```
xmlsentences({st=> "tag"}, text)
```

**EXEMPLO 2.18 - UTILIZAÇÃO DE ETIQUETAS NO MÓDULO FSENTENCES**

nomes das etiquetas (s => 's', p=>'p', t=>'text')

- t: 0 - nenhuma  
 1 - numeracao  
 f - ficheiro [DEFAULT]
- p: 0 - nenhuma  
 1 - numeracao 1 nivel [DEFAULT]  
 2 - numercao 2 niveis (N.N)
- s: 0 - nenhuma  
 1 - numeração 1 nível [DEFAULT]  
 2 - numeração 2 níveis (N.N)  
 3 - numeração 3 níveis (N.N.N)

**EXEMPLO 2.19 - UTILIZAÇÃO DO LEMATIZADOR VERBAL LX-SUITE**

Ao introduzir como *input* a forma verbal 'sido' o programa devolve o seguinte *output*:

**sido**

ser

partícipio passado | masculino | singular

cjg >> trl >>

**outros lemas possíveis**

sedir

indicativo | presente | 1ª pessoa | singular

cjg >> trl >>

sidar

indicativo | presente | 1ª pessoa | singular

cjg >> trl >>

sider

indicativo | presente | 1ª pessoa | singular

cjg >> trl >>

sidir

indicativo | presente | 1ª pessoa | singular

cjg >> trl >>

**EXEMPLO 2.20 – UTILIZAÇÃO DO CONJUGADOR VERBAL LX-SUITE**

Ao introduzir como *input* o verbo 'ser' o programa devolve o seguinte *output*:

VERBO SER	
INDICATIVO	
<i>Presente</i>	<i>Pretérito Perfeito Composto</i>
eu sou tu és ele/ ela/ você é nós somos vós sois eles/ elas/ vocês são	eu hei/ tenho sido tu hás/ tens sido ele/ ela/ você há/ tem sido nós havemos/ temos sido vós haveis/ tendes sido eles/ elas/ vocês hão/ têm sido
<i>Pretérito Imperfeito</i>	<i>Pretérito Mais-que-Perfeito Composto</i>
eu era tu eras ele/ ela/ você era nós éramos vós éreis eles/ elas/ vocês eram	eu havia/ tinha sido tu havias/ tinhas sido ele/ ela/ você havia/ tinha sido nós havíamos/ tínhamos sido vós havíeis/ tínheis sido eles/ elas/ vocês haviam/ tinham sido
<i>Pretérito Perfeito Simples</i>	
eu fui tu foste ele/ ela/ você foi nós fomos vós fostes eles/ elas/ vocês foram	
<i>Pretérito Mais-Que-Perfeito Simples</i>	<i>Pretérito Mais-que-Perfeito Anterior</i>
eu fora tu foras ele/ ela/ você fora nós fôramos vós fôreis eles/ elas/ vocês foram	eu houvera/ tivera sido tu houveras/ tiveras sido ele/ ela/ você houvera/ tivera sido nós houvêramos/ tivêramos sido vós houvêreis/ tivêreis sido eles/ elas/ vocês houveram/ tiveram sido



<i>Futuro do Presente Simples</i>	<i>Futuro do Presente Composto</i>
eu serei tu serás ele/ ela/ você será nós seremos vós sereis eles/ elas/ vocês serão	eu haverei/ terei sido tu haverás/ terás sido ele/ ela/ você haverá/ terá sido nós haveremos/ teremos sido vós havereis/ tereis sido eles/ elas/ vocês haverão/ terão sido
<b>CONJUNTIVO</b>	
<i>Presente</i>	<i>Pretérito Perfeito</i>
eu seja tu sejas ele/ ela/ você seja nós sejamos vós sejais eles/ elas/ vocês sejam	eu haja/ tenha sido tu hajas/ tenhas sido ele/ ela/ você haja/ tenha sido nós hajamos/ tenhamos sido vós hajais/ tendais sido eles/ elas/ vocês hajam/ tenham sido
<i>Pretérito Imperfeito</i>	<i>Pretérito Mais-Que-Perfeito</i>
eu fosse tu fosses ele/ ela/ você fosse nós fôssemos vós fôsseis eles/ elas/ vocês fossem	eu houvesse/ tivesse sido tu houvesse/ tivesses sido ele/ ela/ você houvesse/ tivesse sido nós houvéssomos/ tivéssemos sido vós houvésseis/ tivésseis sido eles/ elas/ vocês houvessem/ tivessem sido
<i>Futuro Simples</i>	<i>Futuro Composto</i>
eu for tu fores ele/ ela/ você for nós formos vós fordes eles/ elas/ vocês forem	eu houver/ tiver sido tu houveres/ tiveres sido ele/ ela/ você houver/tiver sido nós houvermos/ tivermos sido vós houverdes/ tiverdes sido eles/ elas/ vocês houverem/ tiverem sido

IMPERATIVO	
<i>Afirmativo</i>	<i>Negativo</i>
sê tu	não sejas tu
seja você	não seja você
sejamos nós	não sejamos nós
sede vós	não sejais vós
sejam vocês	não sejam vocês
INFINITIVO	
<i>Infinitivo Pessoal Presente</i>	<i>Infinitivo Pessoal Pretérito</i>
eu ser	eu haver/ ter sido
tu seres	tu haveres/ teres sido
ele/ ela/ você ser	ele/ ela/ você haver/ ter sido
nós sermos	nós havermos/ termos sido
vós serdes	vós haverdes/ terdes sido
eles/ elas/ vocês serem	eles/ elas/ vocês haverem/ terem sido
<i>Infinitivo Impessoal Presente</i>	<i>Infinitivo Impessoal Pretérito</i>
ser	haver/ ter sido
GERÚNDIO	
<i>Presente</i>	<i>Pretérito</i>
sendo	havendo/ tendo sido
PARTICÍPIO PASSADO	
sido (inv.)	

### EXEMPLO 2.21 - UTILIZAÇÃO DO FLEXIONADOR NOMINAL LX-SUITE

Ao introduzir a palavra 'imperatrizes', escolhendo o gênero 'Masculino' e número 'Singular' obtemos o seguinte resultado:

<b>Palavra introduzida:</b>	<b>imperatrizes (masc, plur)</b>
<b>Lema:</b>	<b>imperatriz (fem, sing)</b>
<b>Palavra pedida:</b>	<b>imperador (masc, sing)</b>

## EXEMPLO 2.22 - UTILIZAÇÃO DO ANOTADOR CATEGORIAL LX-SUITE

Ao introduzirmos o texto "A uma certa distância todos se parecem... Bonitos, cheirosos e educados. Ao se aproximar, o mau hálito e o sapato gasto influenciam nossa mente a conspirar e mal tratar aleatoriamente, puro preconceito!

Imaginar vivermos num mundo de distintos personagens onde ser diferente talvez seja o normal.", o Anotador Categoral apresenta o seguinte resultado:

```
<p><s>      \*"/CN#??      A/PREP      uma/UM#fs      certa/CERTO/ADJ#fs
          distância/DISTÂNCIA/CN#fs      todos/QNT#mp      se/CL#gn3
parecem/PARECER/V#pi-  3p ...*//PNT </s>
<s>      Bonitos/PNM      ,*//PNT      cheirosos/CHEIROSO/ADJ#mp      e/CJ
educados/EDUCAR,EDUCADO/PPA#mp .*//PNT </s>
<s>      a_/LPREP1      o/LPREP2      se/CL#gn3      aproximar/APROXIMAR/V#inf-nInf      ,*//PNT
o/DA#ms      mau/MAU/ADJ#ms      hálito/HÁLITO/CN#ms      e/CJ      o/DA#ms
sapato/SAPATO/CN#ms      gasto/GASTAR,GASTO/PPA#ms      influenciam/?/V#?
nossa/POSS#fs      mente/MENTE/CN#fs      a/PREP      conspirar/CONSPIRAR/V#INF-
nInf      e/CJ      mal/ADV      tratar/TRATAR/V#inf-nInf      aleatoriamente/ADV      ,*//PNT
puro/PURO/ADJ#ms      preconceito/PRECONCEITO/CN#ms !*//PNT </s>
<s>      Imaginar/PNM      vivermos/VIVER/V#INF-1p      em_/PREP      um/UM#ms
mundo/MUNDO/CN#ms      de/PREP      distintos/DISTINTO/CN#mp
personagens/PERSONAGEM/CN#fp      onde/REL      ser/SER/V#inf-nInf
diferente/DIFERENTE/ADJ#gs      talvez/ADV      seja/SER/V#pc-3s      o/DA#ms
normal/NORMAL/ADJ#ms ./PNT !*//PNT </s></p>
```

## EXEMPLO 2.23 - UTILIZAÇÃO DO RECONHECEDOR DE NOMES PRÓPRIOS DO LX-SUITE

Ao introduzirmos como *input* o texto "Hoje, o João Silva foi ao football.", seleccionando a opção 'baseado em nomes' a ferramenta apresenta o seguinte resultado:

Hoje, o **João Silva** foi ao football

apresentando com uma cor diferente o nome próprio.

**EXEMPLO 2.24 – UTILIZAÇÃO DO NAVEGADOR CORPUS DO LX-SUITE**

1	um animal não é um	<a href="#">computador</a>	. Ele anda mediante a
2	' é um terminal de	<a href="#">computador</a>	ligado a_ a Internet ;
3	sistema através de um simples	<a href="#">computador</a>	, está a ser muito
4	audimetrias através de um simples	<a href="#">computador</a>	, foi dito que em
5	RTP fez - ' qualquer	<a href="#">computador</a>	faz ' - é pôr
6	microscópio , dentro de um	<a href="#">computador</a>	ou sobre uma bancada de
7	, e a modelação em	<a href="#">computador</a>	. Mas os caçadores de
8	, enquanto usam o seu	<a href="#">computador</a>	peçoal . \ " E-Mail Editorial
9	, quiosques e terminais de	<a href="#">computador</a>	. Revolução em marcha \ "
10	poder sentar -me a_ o	<a href="#">computador</a>	! Tenho um jogo novo
11	de gostar de jogos de	<a href="#">computador</a>	e de televisão e ,
12	e a_ os jogos de	<a href="#">computador</a>	. Em_ o próximo volume
13	aquilo já está em_ o	<a href="#">computador</a>	estilo tens um bloco para
14	de_ a impressora em_ o	<a href="#">computador</a>	não não mas esse já
15	os em_ a em_ o	<a href="#">computador</a>	de_ eles portanto ele cria
16	outras vezes faço em_ o	<a href="#">computador</a>	depende hum hum porque em_
17	que fica depois em_ o	<a href="#">computador</a>	e lá em_ o de_
18	um ou atrás de_ um	<a href="#">computador</a>	ou a_ o lado de_
19	quase andar a disputar um	<a href="#">computador</a>	levanta -te tu para me
20	rodas ligado a um pequenino	<a href="#">computador</a>	que manipula apenas com três
21	nem sequer um terminal de	<a href="#">computador</a>	é um endereço virtual é
22	tocarem meticulosamente nota como um	<a href="#">computador</a>	pois sem um mínimo sentido
23	introduzida em_ o em_ o	<a href="#">computador</a>	e que depois fazem há
24	está ela agora levou o	<a href="#">computador</a>	e está a trabalhar em
25	hora de almoço levou o	<a href="#">computador</a>	lá de lá de_ o
26	disse foi olha levas esse	<a href="#">computador</a>	instalas o que tens aí
27	outro lado hum hum qualquer	<a href="#">computador</a>	de_ o centro dá desde
28	desde que eu arranje um	<a href="#">computador</a>	livre tenha placa de som
29	é relativamente fácil encontrar um	<a href="#">computador</a>	livre lá em_ o centro
30	estava despejada de monitor e	<a href="#">computador</a>	ela esteve lá e copiou
31	de maneira o o o	<a href="#">computador</a>	só acelerou a partir de_

## EXEMPLO 2.25 - UTILIZAÇÃO DO BUSCADOR TREEBANK

Quando inroduzimos como *input* a expressão 'S < VP' obtemos o seguinte *output*:

1	A_ a porta de_ a casa , tudo em_ a mesma .	Id:b11
2	As vítimas não fugiam a_ a normalidade .	Id:b18
3	Com tiros de caçadeira , um desconhecido sem motivo aparente abateu uma família inteira .	Id:b3
4	' Falámos de caça , de_ os sócios e de assuntos de homens .	Id:b30
5	' As crianças não querem dormir e as pessoas estão inquietas.	Id:b34
6	Maria Vitória tem razão .	Id:b42
7	Washington acompanhou os movimentos de Saddam desde a primeira hora .	Id:b104
8	Tudo bem .	Id:b56
9	Há uma sinceridade primária que mata uma grande mentira sofisticada .	Id:b65
10	Não deveriam falar .	Id:b68

## API DO MÓDULO TOKENIZER DA BIBLIOTECA FREELING

```
class tokenizer {
public:
    /// Constructor, receives the name of the file with tokenization rules
    tokenizer(const std::string &);
    /// tokenize string with default options
    std::list<word> tokenize(const std::string &);
    /// tokenize string with default options, accumulating byte-offset of words
    std::list<word> tokenize(const std::string &, unsigned long &);
};
```

**EXEMPLO 2.26 – FICHEIRO DE REGRAS DE TOKENIZAÇÃO**

## macros to be used in RegExps rules

<Macros>

ALPHA [^\<>[\\,":;!¿¡«»')^@~|}{\_\\+=&\$#\*+%\\s\d\ -}

ALPHANUM [^\<>[\\,":;!¿¡«»')^@~|}{\_\\+=&\$#\*+%\\s\ -}

NOALPHANUM [\\<>[\\,":;!¿¡«»')^@~|}{\_\\+=&\$#\*+%\\s\ -}

PARTNUM [^\<>[":;!¿¡«»')^@~|}{\s}

OTHERS [\\<>[\\,":;!¿¡«»')^@~|}{\_\\+=&\$#\*+%\\ -}

</Macros>

## Tokenization rules. They are applied in the order of definition.

## The first matching the \*beginning\* of the line is applied

## and a token built. The process is repeated until the line

## has been completely processed.

## -The first field in the rule is the rule name. If it starts

## with a "\*", the RegExp will only produce a token if the

## match is found in abbreviation list <Abbreviations> below.

## -The second field in the rule is the substring to form the token/s with

## It may be 0 (the match of the whole expression) or any number

## from 1 to the number of substrings (up to 9). A token will be

## created for each substring from 1 to the specified value.

## -The third field is the regexp to match against the line

##

<RegExps>

INDEX\_SEQUENCE 0 (\\{4,}|-{2,}|\\\*{2,}|\_{2,}|/{2,})

INITIALS1 1 ([A-Z](\\.[A-Z])+)\\.\.\.

INITIALS2 0 ([A-Z]\\.)+

NAMES\_CODES 0 ({PARTNUM}\*[0-9]{PARTNUM}\*{ALPHANUM})

THREE\_DOTS 0 (\\.\.\.)

QUOTES 0 (`|<<|>>|")

MAILS 0 {ALPHANUM}+([\\.\_]{ALPHANUM}+)\*@{ALPHANUM}+([\\.\_]{ALPHANUM}+)\*

URLS 0 ((mailto:|(news|http|https|ftp|ftps)://)[\\w\\.\ -]|^(www(\\.[\\w\\ -]+)))

KEEP\_COMPOUNDS 0 {ALPHA}+(['\_-]{ALPHA}+)+

\*ABREVIATIONS1 0 ({ALPHA}+\..)+(?!\\.\.)

\*ABREVIATIONS2 0 ({ALPHA}+\..)+(?!\\.\.)

WORD 0 {ALPHANUM}+

OTHERS\_C 0 {OTHERS}

</RegExps>

## Abbreviations. The dot is not tokenized separately

## in the cases listed below.

<Abbreviations>	# cf.	# e.g.	# i.e.	# vs.
# Av.	# Lrg.	# Pr.	# Trav.	# Estr.
# Lt.	# Rot.	# Jan.	# Fev.	# Mar.
# Abr.	# Mai.	# Jun.	# Jul.	# Ago.
# Set.	# Out.	# Nov.	# Dez.	# Seg.
# Qua.	# Qui.	# Sex.	# Sáb.	# Dom.
# Cap.	# Fl.	# Num.	# Pgs.	# Eds.
# Séc.	# Ex.	# N <sup>o</sup> .	# No.	# Pg.
# Págs.	# Tab.	# Fig.	# Pág.	# Sec.
# Vol.	# Sr.	# Snr.	# Snr <sup>a</sup> .	# Snra.
# Sr <sup>a</sup> .	# Sra.	# Srs.	# Pres.	# Sec <sup>o</sup> .
# Sec <sup>a</sup> .	# Dir.	# Dir <sup>a</sup> .	# Dira.	# Ab.
# Fr.	# Pe.	# Prof.	# Prof <sup>a</sup> .	# Profa.
# Dr.	# Dr <sup>a</sup> .	# Dra.	# Arq.	# Arq <sup>o</sup> .
# Arqu.	# Arq <sup>a</sup> .	# Arqa.	# Eng.	# Eng <sup>o</sup> .
# Engo.	# Eng <sup>a</sup> .	# Enga.	# Em.	# Em <sup>a</sup> .
# Ema.	# Ex <sup>o</sup> .	# Exo.	# Exm <sup>o</sup> .	# Exmo.
# Ex <sup>a</sup> .	# Exa.	# Exm <sup>a</sup> .	# Exma.	# Exs.
# Exos.	# Exmos.	# Exas.	# Exmas.	# Exc <sup>a</sup> .
# Exca.	# Eci <sup>a</sup> .	# Ecia.	# Excia.	# Excas.
# Excias.	# Rev.	# Rev <sup>a</sup> .	# Reva.	a.c.
aa.rr.	abrev.	adj.	adm.	admón.
afma.	afmas.	afmo.	afmos.	ag.
am.	ap.	apdo.	art.	arts.
arz.	arzbpo.	assn.	atte.	av.
avda.	bros.	bv.	cap.	caps.
cg.	cgo.	cia.	cit.	cl.

cm.	co.	col.	corp.	cos.
cta.	cte.	ctra.	cts.	cía.
d.c.	dcha.	dept.	depto.	dg.
dl.	dm.	doc.	docs.	dpt.
dpto.	dr.	dra.	dras.	dres.
dto.	dom.	dupdo.	ed.	eds.
ee.uu.	ej.	emma.	emmas.	emmo.
emmos.	entlo.	entpo.	esp.	etc.
ex.	excm.	excma.	excmas.	excmo.
excmos.	fasc.	fdo.	fig.	figs.
fol.	fra.	gb.	gral.	ha.
hnos.	hros.	hz.	ib.	ibid.
ibíd.	id.	ilm.	ilma.	ilmas.
ilmo.	ilmos.	iltre.	inc.	intr.
izq.	izqda.	izqdo.	jr.	kc.
kcal.	kg.	khz.	kl.	km.
kw.	lda.	ldo.	lib.	lim.
loc.	ltd.	ltda.	lám.	ma.
mg.	mhz.	min.	mons.	mr.
mrs.	ms.	mss.	mtro.	máx.
mín.	ntra.	ntro.	núm.	ob.
obpo.	op.	pd.	ph.	pje.
pl.	plc.	pm.	pp.	ppal.
pral.	prof.	prov.	pról.	ps.
pta.	ptas.	pte.	pts.	pza.
pág.	págs.	párr.	q.b.	rda.
rdo.	ref.	reg.	rel.	rev.
revda.	revdo.	rma.	rmo.	rte.
s.	sáb.	sdad.	sec.	secret.
seg.	sg.	sig.	smo.	sp.
sr.	sra.	stras.	sres.	srs.
srta.	ss.mm.	sta.	sto.	sust.
tech.	tel.	telf.	teléf.	ten.



tfono.	tlf.	tít.	ud.	uds.
vda.	vdo.	vid.	vol.	vols.
vra.	vro.	vta.	íd.	ít.
mm.	mms.	ms.	pulg.	yda.
mi.	Ha.	ac.	ml.	dl.
hl.	ac-pie.	oz.	qt.	gal.
pk.	bu.	cr.	crt.	tz.
pt.	mpa.	pa.	psi.	lb.
mmhg.	cmhg.	mhg.	mol.	mg.
gr.	grs.	kg.	kgs.	mgr.
oz.	lb.	ton.	tm.	milgal.
lt.	lps.	gps.	gpm.	gph.
gpd.	mgd.	gal.	gpcd.	mph.
lbf.	yb.	zb.	eb.	pb.
tb.	gb.	mb.	kb.	wb.
cd.	rad.	sr.	hz.	lm.
lx.	nq.	gy.	sv.	</Abbreviations>

### EXEMPLO 2.27 - DEFINIÇÃO DE UMA MACRO

ALPHA [A-Za-z];

## API DO MÓDULO SPLITTER DA BIBLIOTECA FREELING

```
class splitter {
public:
    /// Constructor. Receives a file with the desired options
    splitter(const std::string &);
    /// Add list of words to the buffer, and return complete sentences
    /// that can be build.
    /// The boolean states if a buffer flush has to be forced (true) or
    /// some words may remain in the buffer (false) if the splitter
    /// wants to wait to see what is coming next.
    std::list<sentence> split(const std::list<word> &, bool);
};
```

### EXEMPLO 2.28 - FICHEIRO DE OPÇÕES DO SPLITTER

```
<General>
AllowBetweenMarkers 1
MaxLines 0
</General>
<Markers>
" "
( )
{ }
/* */
</Markers>
<SentenceEnd>
. 0
? 0
! 0
</SentenceEnd>
<SentenceStart>
</SentenceStart>
```

**EXEMPLO 2.29 - SECÇÃO <GENERAL>**

```
<General>
```

```
AllowBetweenMarkers 0
```

```
MaxLines 0
```

```
</General>
```

Se a opção *AllowBetweenMarkers* estiver desligada (0), nunca será introduzida uma divisão frásica que esteja dentro de parêntesis.

Se, pelo contrário, esta opção estiver ligada (1), será permitido ter uma frase dentro de parêntesis.

A opção *MaxLines* permite declarar quantas linhas de texto são lidas antes de forçar o *split* de uma frase entre parêntesis. Um valor de zero significa que nunca haverá *split*.

É de notar que esta opção só tem efeito se a opção *AllowBetweenMarkers* tiver o valor um.

**EXEMPLO 2.30 - SECÇÃO < MARKERS >**

```
<Markers>
```

```
" "
```

```
()
```

```
{ }
```

```
/* */
```

```
</Markers>
```

**EXEMPLO 2.31 - SECÇÃO <SENTENCEEND>**

```
<SentenceEnd>
```

```
. 0
```

```
? 1
```

```
! 0
```

```
</SentenceEnd>
```

### EXEMPLO 2.32 - SECÇÃO <SENTENCESTART>

```
<SentenceStart>
ç
i
</SentenceStart>
```

### API DO MÓDULO ANALISADOR MORFOLÓGICO DA BIBLIOTECA FREELING

```
class maco {
public:
    /// Constructor. Receives a set of options.
    maco(const maco_options &);
    /// analyze and enrich given sentences.
    void analyze(std::list<sentence> &);
};

class maco_options {
public:
    /// Language analyzed
    std::string Lang;
    /// Submodules to activate
    bool AffixAnalysis, MultiwordsDetection,
    NumbersDetection, PunctuationDetection,
    DatesDetection, QuantitiesDetection,
    DictionarySearch, ProbabilityAssignment;
    /// kind of NER wanted (NER_BASIC, NER_BIO, NER_NONE)
    int NERecognition;
    /// Names of data files to provide to each submodule.
    std::string LocutionsFile, QuantitiesFile, AffixFile,
    ProbabilityFile, DictionaryFile,
    NPdataFile, PunctuationFile;
```

---

```

/// Extra parameters for Number Detection module
std::string Decimal, Thousand;
/// Extra parameters for Probability Assignment module
double ProbabilityThreshold;
/// constructor
maco_options(const std::string &);
/// Option setting methods provided to ease perl interface generation.
/// Since option data members are public and can be accessed directly
/// from C++, the following methods are not necessary, but may become
/// convenient sometimes.
/// The order of the parameters is the same they are defined above.
void set_active_modules(bool,bool,bool,bool,bool,bool,bool,bool,int,bool);
void set_data_files(const std::string &,const std::string &,
const std::string &,const std::string &,
const std::string &,const std::string &,
const std::string &,const std::string &);
void set_numerical_points(const std::string &,const std::string &);
void set_threshold(double);
}

```

## **API DO MÓDULO DETECTOR DE NÚMEROS DA BIBLIOTECA FREELING**

```

class numbers {
public:
/// Constructor: receives the language code, and the decimal
/// and thousand point symbols
numbers(const std::string &, const std::string &, const std::string &);
/// Detect number expressions in given sentence
void annotate(sentence &);
};

```

## **API DO MÓDULO DETECTOR DE MARCAS DE PONTUAÇÃO DA BIBLIOTECA FREELING**

```
class punts {  
public:  
    /// Constructor: receives data file name  
    punts(const std::string &);  
    /// Detect punctuation in given sentence  
    void annotate(sentence &);  
};
```

### **EXEMPLO 2.33 - FICHEIRO DE TAG'S ASSOCIADAS A MARCAS DE PONTUAÇÃO**

```
! Fat  
, Fc  
: Fd  
... Fs
```

### **EXEMPLO 2.34 - LINHA COM TAG ESPECIAL**

```
<Other> Fz
```

**API DO MÓDULO DETECTOR DE DATAS DA BIBLIOTECA FREELING**

```
class dates {
public:
/// Constructor: receives the language code
dates(const std::string &);
/// Detect date/time expressions in given sentence
void annotate(sentence &);
};
```

**API DO MÓDULO DE PESQUISA DE DICIONÁRIO DA BIBLIOTECA FREELING**

```
class dictionary {
public:
/// Constructor
dictionary(const std::string &, const std::string &,
bool, const std::string &);
/// Get analysis for a given form, and add them
/// to given analysis list
void search_form(const std::string &, std::list<analysis> &);
/// Analyze words in given sentence
void annotate(sentence &);
}
```

**EXEMPLO 2.35 - EXEMPLO DE LINHAS DE UM FICHEIRO PLANO DE DICIONÁRIO**

```
casa casa NCFS000 casar VMIP3S0 casar VMM02S0
backs back NNS back VBZ
```

### EXEMPLO 2.36 - DEFINIÇÃO DE CONTRACÇÃO NO FICHEIRO DO DICIONÁRIO

del de+el SPS+DA

Esta linha expressa que sempre que a forma *del* seja encontrada, será substituída com duas palavras: *de* e *el*.

Cada uma das duas formas é pesquisada no dicionário e associada a qualquer *tag* que emparelha a sua *tag* expressa no terceiro campo.

### EXEMPLO 2.37 - FICHEIRO DE DICIONÁRIO INDEXADO

```
////////////////////////////////////  
// FreeLing - Open Source Language Analyzers  
//  
// Copyright (C) 2004 TALP Research Center  
//          Universitat Politecnica de Catalunya  
//  
// This library is free software; you can redistribute it and/or  
// modify it under the terms of the GNU General Public  
// License as published by the Free Software Foundation; either  
// version 2.1 of the License, or (at your option) any later version.  
//  
// This library is distributed in the hope that it will be useful,  
// but WITHOUT ANY WARRANTY; without even the implied warranty of  
// MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU  
// General Public License for more details.  
// You should have received a copy of the GNU General Public  
// License along with this library; if not, write to the Free Software  
// Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA  
//  
// contact: Lluís Padro (padro@lsi.upc.es)  
//          TALP Research Center
```



```
//      despatx C6.212 - Campus Nord UPC
//      08034 Barcelona. SPAIN
//
////////////////////////////////////

////////////////////////////////////
//
// Indexate a raw-text dictionary in a BerkeleyDB database
//
// Data read from stdin. argv[1] specifies the name of the DB to be created.
// Expected format for raw text file (read from stdin):
//
// key1 data1
// key2 data2
// ...
//
////////////////////////////////////

#include <iostream>
#include <string>
#include <db_cxx.h>
#include <cstdlib>

using namespace std;

int main(int argc, char *argv[])
{
    int res, nlin;
    string s, s_form, s_analysis;
    Dbt key, data;

    // create database
    Db mydbase(NULL,DB_CXX_NO_EXCEPTIONS);
    if ((res=mydbase.open(NULL,argv[1],NULL,DB_HASH,DB_CREATE,0644))) {
        mydbase.err(res,"Error %d while creating database
```

```
'%s"',res,argv[1]);
    exit(res);
}

// read stdin and insert records in database
for (nlin=0; getline(cin,s); nlin++) {
    // split line in key+data
    s_form=s.substr(0,s.find(" "));
    s_analysis=s.substr(s.find(" ")+1);
    // build db record
    key.set_data((void *)s_form.c_str());
    key.set_size(s_form.length());
    data.set_data((void*)s_analysis.c_str());
    data.set_size(s_analysis.length());

    // insert record in DB.
    res = mydbase.put(NULL, &key, &data, DB_NOOVERWRITE);
    if (res == DB_KEYEXIST) {
        mydbase.err(res,"Unexpected duplicate key %s at line %d",
                    s_form.c_str(),nlin);

        exit(res);
    }
}

if ((res=mydbase.close(0)) {
    mydbase.err(res,"Error %d while closing database",res);
    exit(res);
}
}
```

### **EXEMPLO 2.38 - ELIMINAÇÃO DE AFIXO DE UMA PALAVRA**

crucecita - cecita = cru

### **EXEMPLO 2.39 - ADIÇÃO DE AFIXO A UMA PALAVRA**

cru + z = cruz

### **EXEMPLO 2.40 - CONDIÇÃO DA TAG CONDICIONAL**

cruz é NCFS

### **EXEMPLO 2.41 - AFECTAÇÃO DE LEMA**

R+A, A+L, R+mente

### **EXEMPLO 2.42 - REGRAS DE PREFIXO**

anti \* ^NC AQ0CN0 0 0 1 A+L 0 –

Esta regra diz que ‘anti’ deveria ser removido do início da palavra, nada (\*) deveria ser adicionado e a raiz resultante deveria ser encontrada com uma tag ‘NC’.

Se todas as condições anteriores forem satisfeitas, a palavra deverá receber a tag 'AQ0CN0' e o afixo (anti) mais o lema como o lema da palavra prefixada.

Por exemplo, a palavra *antimísiles* emparelha com esta regra:

- *mísiles* seria encontrado no dicionário com lema *mísil* e tag 'NCMP000', seguidamente à palavra será associado o lema *antimísil* (A+L = anti + mísil) e a tag 'AQ0CN0'.

### EXEMPLO 2.43 - REGRA DE SUFIXO

*cecita* z|za ^NCF5 NCF500A 0 0 1 L 0 -

*les* \* ^V \* 0 1 0 L 1 \$\$+les:\$\$+PP

A primeira regra de sufixo (*cecita*), diz que uma regra de sufixo será aplicada a palavras desconhecidas, para ver se um nome feminino singular válido é obtido quando se substitui o sufixo *cecita* com *z* ou *za*.

Este é o caso de *crucecita* (diminutivo de *cruz*).

Se uma tal forma básica for encontrada, a palavra original é analisada como forma de diminutivo sufixado.

Não é necessária re-tokenização.

A segunda regra (*les*) aplica-se a todas as palavras e tenta verificar se uma forma verbal válida é obtida quando se remove o sufixo *les*. É o caso de palavras tais como *viles*.

Neste caso, a informação de re-tokenização diz que se eventualmente a tag *verb* for selecionada para esta palavra, ela poderá ser re-tokenizada em duas palavras:

- a forma verbal base – referida como \$\$, *vi* no exemplo;
- a palavra *les*

As *tag's* para estas novas palavras são expressas depois de ':':

A forma base deve manter a sua *tag PoS* (isto é o que significa o segundo '\$\$') e a segunda palavra pode ter qualquer tag que inicie com 'PP' e que esteja no dicionário.

Deste modo, para a palavra *viles* obteríamos a sua análise de adjetivo, mais o seu verbo, mais o pronome clítico a partir da seguinte regra de sufixo:

*viles vil AQ0CP0 ver VMIS1S0.*

A segunda análise leva em conta a informação da re-tokenização. Assim, se o *tagger* PoS seleccionar a análise VMI (e a opção **TaggerRetokenize** estiver activa) a palavra será re-tokenizada em:

- vi ver VMIS1S0
- les ellos PP3CPD00

## API DO MÓDULO IDENTIFICADOR MULTIPALAVRA DA BIBLIOTECA FREELING

```
class automat {
public:
  /// Constructor
  automat();
  /// Detect patterns in given sentence
  void annotate(sentence &);
};

class locutions: public automat {
public:
  /// Constructor, receives the name of the file
  /// containing the multiwords to recognize.
  locutions(const std::string &);
};
```

### EXEMPLO 2.44 - FICHEIRO DE DEFINIÇÃO MULTIPALAVRA

```
a_boas_horas a_boas_horas RG
a_causa_de a_causa_de SPS00
<acidente>_de_trabalho acidente_de_trabalho $1:NC
```

A *tag* pode ser especificada directamente, ou como uma referência à *tag* de componentes multipalavra.

Neste exemplo, a última especificação de multipalavra constrói uma multipalavra com qualquer das formas ‘acidente de trabalho’ ou ‘acidentes de trabalho’.

A *tag* de multipalavra será a da primeira forma (\$1) que inicia com NC.

Aquela afectará à multipalavra a *tag* correcta (singular/plural), dependendo do facto de a forma ser *acidente* ou *acidentes*.

## **API DO MÓDULO IDENTIFICADOR DE ENTIDADES NOMEADAS DA BIBLIOTECA FREELING**

```
class np: public ner, public automat {
public:
    /// Constructor, receives a configuration file.
    np(const std::string &);
    /// ("annotate" is inherited from "automat")
    void annotate(sentence &);
};
```

### **EXEMPLO 2.45 - DECLARAÇÃO DA SECÇÃO <FUNCTIONWORDS>**

```
<FunctionWords>
o
a
os
as
de
para
</FunctionWords>
```

### **EXEMPLO 2.46 - DECLARAÇÃO DA SECÇÃO <SPECIALPUNCT>**

```
<SpecialPunct>  
Fpa  
Fp  
Fd  
Fg  
</SpecialPunct>
```

### **EXEMPLO 2.47 - DECLARAÇÃO DA SECÇÃO <NE\_TAG>**

```
<NE_Tag>  
NP00000  
</NE_Tag>
```

### **EXEMPLO 2.48 - DECLARAÇÃO DA SECÇÃO <IGNORE>**

```
<Ignore>  
i 0  
RB 1  
english 1  
dutch 1  
spanish 1  
</Ignore>
```

### **EXEMPLO 2.49 - DECLARAÇÃO DA SECÇÃO <NAMES>**

```
<Names>  
miren  
pel  
zapatero  
china  
</Names>
```

### **EXEMPLO 2.50 - DECLARAÇÃO DAS SECÇÕES <RE\_NOUNADJ> E <RE\_CLOSED>**

```
<RE_NounAdj>  
^(NN$|NNS|JJ)  
</RE_NounAdj>  
<RE_Closed>  
^(D|IN|C)  
</RE_Closed>
```

### **EXEMPLO 2.51 - DECLARAÇÃO DA SECÇÃO <TITLELIMIT>**

```
<TitleLimit>  
3  
</TitleLimit>
```



**EXEMPLO 2.52 - DECLARAÇÃO DA SECÇÃO <SPLITMULTIWORDS>**

```
<SplitMultiwords>  
no  
</SplitMultiwords>
```

**API DO MÓDULO BIONER DA BIBLIOTECA FREELING**

```
class bioner: public ner {  
public:  
    /// Constructor, receives the name of the configuration file.  
    bioner ( const std::string & );  
    /// Recognize NEs in given sentence  
    void annotate ( sentence & );  
};
```

**EXEMPLO 2.53 - DECLARAÇÃO DA SECÇÃO <RGF>**

```
< RGF >  
ner.rgf  
</ RGF >
```

**EXEMPLO 2.54 - DECLARAÇÃO DA SECÇÃO <ADABOOSTMODEL>**

```
<AdaBoostModel>  
ner.abm  
</AdaBoostModel>
```

### **EXEMPLO 2.55 - DECLARAÇÃO DA SECÇÃO <LEXICON>**

```
<Lexicon>  
ner.lex  
</Lexicon>
```

### **EXEMPLO 2.56 - DECLARAÇÃO DA SECÇÃO <CLASSES>**

```
<Classes>  
0 B 1 I 2 O  
</Classes>
```

### **EXEMPLO 2.57 - DECLARAÇÃO DA SECÇÃO <INITIALPROB>**

```
<InitialProb>  
B 0.200072  
I 0.0  
O 0.799928  
</InitialProb>
```

**EXEMPLO 2.58 - DECLARAÇÃO DA SECÇÃO <TRANSITIONPROB>**

```
<TransitionProb>
B B 0.00829346
B I 0.395481
B O 0.596225
I B 0.0053865
I I 0.479818
I O 0.514795
O B 0.0758838
O I 0.0
O O 0.924116
</TransitionProb>
```

**EXEMPLO 2.59 - DECLARAÇÃO DA SECÇÃO <TITLELIMIT>**

```
<TitleLimit>
3
</TitleLimit>
```

**API DO MÓDULO IDENTIFICADOR DE QUANTIDADES DA BIBLIOTECA FREELING**

```
class quantities {
public:
    /// Constructor: receives the language code, and the data file.
    quantities(const std::string &, const std::string &);
    /// Detect magnitude expression in given sentence
    void annotate(sentence &);
};
```

### **EXEMPLO 2.60 - DECLARAÇÃO DA SECÇÃO <CURRENCY>**

```
<Currency>  
CUR  
</Currency>
```

### **EXEMPLO 2.61 - DECLARAÇÃO DA SECÇÃO <MEASURE>**

```
<Measure>  
CUR USD  
CUR FRF  
LN mm  
SP ft/s  
</Measure>
```

Este exemplo mostra a declaração de quatro variáveis:

- USD é do tipo CUR (moeda);
- FRF é do tipo CUR (moeda);
- mm é do tipo LN (comprimento);
- ft/s é do tipo SP (velocidade).

**EXEMPLO 2.62 - DECLARAÇÃO DA SECÇÃO <MEASURENAMES>**

```

<MeasureNames>
french_<franc> FRF $2:N
<franc> FRF $1:N
<dollar> USD $1:N
american_<dollar> USD $2:N
us_<dollar> USD $2:N
<milimeter> mm $1:N
<foot>_per_second ft/s $1:N
<foot>_Fh_second ft/s $1:N
<foot>_Fh_s ft/s $1:N
<foot>_second ft/s $1:N
</MeasureNames>

```

Esta secção reconhecerá *strings* tais como:

- 234\_french\_francs CUR\_FRF:234 Zm;
- one\_dollar CUR\_USD:1 Zm;
- two\_hundred\_fifty\_feet\_per\_second SP\_ft/s:250 Zu.

**API DO MÓDULO AFECTADOR DE PROBABILIDADES E RECONHECEDOR DE PALAVRAS****DESCONHECIDAS DA BIBLIOTECA FREELING**

```

class probabilities {
public:
/// Constructor: receives the language code, the name of the file
// containing probabilities, and a threshold.
probabilities(const std::string &, const std::string &, double);
/// Assign probabilities to all analysis of each word in sentence
void annotate(sentence &);
/// Assign probabilities to all analysis of given word
void annotate_word(word &);
};

```

### **EXEMPLO 2.63 - DECLARAÇÃO DA SECÇÃO <FORMTAGFREQ>**

```
<FormTagFreq>  
japonesas AQ-NC AQ 1 NC 0  
  
</FormTagFreq>
```

### **EXEMPLO 2.64 - DECLARAÇÃO DA SECÇÃO <CLASSTAGFREQ>**

```
<ClassTagFreq>  
AQ-NC AQ 2361 NC 2077  
  
</ClassTagFreq>
```

### **EXEMPLO 2.65 - DECLARAÇÃO DA SECÇÃO <SINGLETAGFREQ>**

```
<SingleTagFreq>  
AQ 7462  
  
</SingleTagFreq>
```

### **EXEMPLO 2.66 - DECLARAÇÃO DA SECÇÃO <THEETA>**

```
<Theeta>  
0.00834  
  
</Theeta>
```

**EXEMPLO 2.67 - DECLARAÇÃO DA SECÇÃO <SUFFIXES>**

```
<Suffixes>
orada 133 AQ0FSP 17 VMP00SF 8 NCFS000 108
</Suffixes>
```

**API DO MÓDULO ETIQUETADOR DE SENTIDOS DA BIBLIOTECA FREELING**

```
class senses {
public:
/// Constructor: receives the name of the dictionary file and a boolean.
senses(const std::string &, bool);
/// sense annotate selected analysis for each word in given sentences
void analyze(std::list<sentence> &);
};
```

**API DO MÓDULO DESAMBIGUADOR DE SENTIDOS DA BIBLIOTECA FREELING**

```
class disambiguator {
public:
/// Constructor. Receives a relation file for UKB, a sense dictionary,
/// and two UKB parameters: epsilon and max iteration number.
disambiguator(const std::string &, const std::string &, double, int);
/// word sense disambiguation for each word in given sentences
void analyze(std::list<sentence> &);
};
```

## API DO MÓDULO ETIQUETADOR DO DISCURSO ORAL - HMM\_TAGGER - DA BIBLIOTECA

### FREELING

```
class hmm_tagger: public POS_tagger {
public:
    /// Constructor
    hmm_tagger(const std::string &, const std::string &, bool, unsigned int);
    /// disambiguate given sentences
    void analyze(std::list<sentence> &);
};
```

### EXEMPLO 2.68 - DECLARAÇÃO DA SECÇÃO <TAG>

```
<Tag>
0 0.03747
AQ 0.00227
NC 0.18894
x 1.07312e-06
</Tag>
```

### EXEMPLO 2.69 - DECLARAÇÃO DA SECÇÃO <BIGRAM> COM A PROBABILIDADE DE TRANSIÇÃO ENTRE O ESTADO INICIAL DE UMA FRASE E A PRIMEIRA PALAVRA QUE CONTÉM A TAG 'AQ'.

```
<Bigram>
0.AQ 0.01403
</Bigram>
```



**EXEMPLO 2.70 – DECLARAÇÃO DA SECÇÃO <BIGRAM> COM A PROBABILIDADE DE TRANSIÇÃO ENTRE DUAS TAG’S CONSECUTIVAS.**

```
<Bigram>  
AQ.NC 0.16963  
</Bigram>
```

**EXEMPLO 2.71 - DECLARAÇÃO DA SECÇÃO <TRIGRAM> COM A PROBABILIDADE DE UMA PALAVRA COM A TAG ‘NC’ APARECER APÓS UMA SEQUÊNCIA ‘0.AQ’.**

```
<Trigram>  
0.AQ.NC 0.204081  
</Trigram>
```

**EXEMPLO 2.72 - DECLARAÇÃO DA SECÇÃO <TRIGRAM> COM A PROBABILIDADE DE UMA PALAVRA COM A TAG ‘SP’ APARECER APÓS DUAS PALAVRAS COM AS TAG’S ‘DA’ E ‘NC’.**

```
<Trigram>  
DA.NC.SP 0.33312  
</Trigram>
```

**EXEMPLO 2.73 - DECLARAÇÃO DA SECÇÃO <INITIAL> COM A PROBABILIDADE DE UMA FRASE INICIAR COM UM DETERMINANTE.**

```
<Initial>  
0.DA -1.744857  
</Initial>
```

**EXEMPLO 2.74 - DECLARAÇÃO DA SECÇÃO <INITIAL> COM A PROBABILIDADE DE UMA FRASE INICIAR COM UMA TAG DESCONHECIDA.**

```
<Initial>  
0.x -10.462703  
</Initial>
```

**EXEMPLO 2.75 - DECLARAÇÃO DA SECÇÃO <WORD>**

```
<Word>  
afortunado -13.69500  
sutil -13.57721  
<UNOBSERVED_WORD> -13.82853  
</Word>
```

**EXEMPLO 2.76 - DECLARAÇÃO DA SECÇÃO <SMOOTHING>**

```
<Smoothing>  
c1 0.120970620869314  
c2 0.364310868831106  
c3 0.51471851029958  
</Smoothing>
```

**EXEMPLO 2.77 - DECLARAÇÃO DA SECÇÃO <FORBIDDEN>**

```
<Forbidden>  
*.PT.NC  
0.DT.VMI  
SP.PP.NC  
</Forbidden>
```

**API DO MÓDULO ETIQUETADOR DO DISCURSO ORAL - RELAX\_TAGGER - DA BIBLIOTECA****FREELING**

```

class relax_tagger : public POS_tagger {
public:
    /// Constructor, given the constraint file and config parameters
    relax_tagger(const std::string &, int, double, double, bool, unsigned int);
    /// disambiguate sentences
    void analyze(std::list<sentence> &);
};

```

**EXEMPLO 2.78 - DECLARAÇÃO DA SECÇÃO SET**

```

DetMasc = DA0MS0 DA0MP0 DD0MS0 DD0MP0 DI0MS0 DI0MP0 DP1MSP DP1MPP
DP2MSP DP2MPP DT0MS0 DT0MP0 DE0MS0 DE0MP0 AQ0MS0 AQ0MP0;
VerbPron = <dar_cuenta> <atrever> <arrepentir> <equivocar> <inmutar>
<morir> <ir> <manifestar> <precipitar> <referir> <venir>;
Animal = [00008019] [00862484] [00862617] [00862750] [00862871] [00863425]
[00863992] [00864099] [00864394] [00865075] [00865379] [00865569]
[00865638] [00867302] [00867448] [00867773] [00867864] [00868028]
[00868297] [00868486] [00868585] [00868729] [00911889] [00985200]
[00990770] [01420347] [01586897] [01661105] [01661246] [01664986]
[01813568] [01883430] [01947400] [07400072] [07501137];

```

### **EXEMPLO 2.79 - DECLARAÇÃO DE RESTRIÇÕES.**

-8.143 DA\* (1 VMI\*) – esta restrição significa que a probabilidade de uma palavra ser um determinante definido (DA\*), se a próxima palavra for uma forma pessoal de um verbo (VMI\*) é muito pequena.

### **API DO MÓDULO CLASSIFICADOR DE ENTIDADES NOMEADAS DA BIBLIOTECA FREELING**

```
class nec {
public:
  /// Constructor
  nec(const std::string &, const std::string &);
  /// Classify NEs in given sentence
  void analyze(std::list<sentence> &) const;
}.
```

### **API DO MÓDULO PARSER GRÁFICO- DA BIBLIOTECA FREELING**

```
class chart_parser {
public:
  /// Constructor
  chart_parser(const std::string&);
  /// Get the start symbol of the grammar
  std::string get_start_symbol(void) const;
  /// Parse sentences in list
  void analyze(std::list<sentence> &);
};
```

## API DO MÓDULO PARSER DE DEPENDÊNCIAS- DA BIBLIOTECA FREELING

```
class dep_txala : public dependency_parser {
public:
    /// constructor
    dep_txala(const std::string &, const std::string &);
    /// Enrich all sentences in given list with a dependency tree.
    void analyze(std::list<sentence> &);
};
```

### EXEMPLO 2.80 - DEFINIÇÃO DE REGRA

```
verb-phr subj d.label=np* d.side=left
```

Esta regra significa que se um nó *verb-phr* tiver um filho à sua esquerda, com uma etiqueta iniciando com *np*, esta dependência deverá ser etiquetada como *subj*.

### EXEMPLO 2.81 - DEFINIÇÃO DE REGRA

```
verb-phr obj d.label=np* d:sn.tonto=Edible p.lemma=eat|gulp
```

Esta regra significa que se um nó *verb-phr* tiver *eat* ou *gulp* como lema, e um descendente com uma etiqueta que inicia com *np*, contendo um filho com a etiqueta *sn* que tem a propriedade *Edible* em EWN Top ontology, esta dependência deverá possuir a etiqueta *obj*.

**EXEMPLO 2.82 - DEFINIÇÃO DE REGRA**

```
verb-phr iobj d.label=pp* d.lemma=to|for Es.label=np*
```

Esta regra significa que se um *verb-phr* tiver um descendente, cuja etiqueta inicia com *pp* (frase preposicional) e lema *to* ou *for*, e houver um filho do mesmo pai, que seja uma frase nominal (*np\**), esta dependência deverá ser etiquetada com *iobj*.

**EXEMPLO 2.83 - DEFINIÇÃO DE REGRA**

```
verb-phr dobj d.label=pp* d.lemma=to|for As.label!=np*
```

Esta regra significa que se um *verb-phr* tiver um descendente com uma etiqueta que inicia com *pp* (frase preposicional) e lema *to* ou *for* e todos os outros filhos do mesmo pai não forem frases nominais (*np\**), esta dependência deverá ser etiquetada com *dobj*.

**API DO MÓDULO DE RESOLUÇÃO DE CORREFERÊNCIAS DA BIBLIOTECA FREELING**

```
class coref {  
public:  
    /// Constructor  
    coref(const std::string &, const int);  
    /// Classify in coreference chains noun phrases in given document  
    void analyze(document &) const;  
};
```

**EXEMPLO 2.84 - DECLARAÇÃO DA SECÇÃO <ABMODEL>**

```
<ABModel>  
coref.abm  
</ABModel>
```

**EXEMPLO 2.85 – DECLARAÇÃO DA SECÇÃO <SEMDB>**

```
<SEMDB>
SenseFile ../senses16.db
WNFile ../../common/wn16.db
</SEMDB>
```

**API DO MÓDULO DA BASE DE DADOS SEMÂNTICA DA BIBLIOTECA FREELING**

```
class semanticDB {
public:
  /// Constructor
  semanticDB(const std::string &, const std::string &);
  /// get list of words for a sense+pos
  std::list<std::string> get_sense_words(const std::string &, const std::string &);
  /// get list of senses for a lemma+pos
  std::list<std::string> get_word_senses(const std::string &, const std::string &);
  /// get sense info for a sensecode+pos
  sense_info get_sense_info(const std::string &, const std::string &);
};
```

**EXEMPLO 2.86 – DECLARAÇÃO DO FICHEIRO QUE CONTÉM O DICIONÁRIO DE SENTIDOS**

```
W:cebolla:N 05760066 08734429 08734702
S:07389783:N chaval chico joven mozo muchacho
```

**EXEMPLO 2.87 - ANOTAÇÃO DE UM TEXTO COM O POS TREE-TAGGER**

Ao introduzir como *input* o seguinte texto ‘O TreeTagger é fácil de utilizar.’, obtemos o seguinte *output*:

palavra	tag PoS	lema
O	DET	o
TreeTagger	NOM	TreeTagger
é	V	ser
fácil	ADJ	fácil
de	PREP	de
utilizar	V	utilizar
.	SENT	.

DET – determinante

NOM – nome comum / próprio

V – verbo

ADJ – adjetivo

PREP – preposição

SENT – separador de orações

**EXEMPLO 2.88 - LEMATIZAÇÃO DE UMA PALAVRA COM O PTSTEMMER**

Ao introduzirmos a palavra ‘optimamente’ o PtStemmer devolve-nos o seguinte resultado:

Stem: optim



### **EXEMPLO 2.89 - CLASSIFICAÇÃO DE ENTIDADES MENCIONADAS**

Considere-se a frase: 'Eu moro na Rua de Angola.'

A SASKIA classificou previamente a entidade *Angola* como sendo *LOCAL/HUMANO/PAIS*.

A aplicação de uma regra gramatical dedicada à captura de ruas vai mudar a classificação da entidade nomeada *Angola* para *LOCAL/HUMANO/RUA*, devido à presença da expressão *Rua da* antes da entidade nomeada.



## **ANEXO IV – CONJUGADORES VERBAIS**



**EXEMPLO 2.90 - CONJUGAÇÃO DO VERBO SER UTILIZANDO O CONJUGUE**

<p># paradigma irregular</p> <p>Formas Nominais:</p> <p>infinitivo: ser</p> <p>gerúndio: sendo</p> <p>particípio: sido</p>	
<p>Presente do Indicativo</p> <p>eu sou</p>	<p>Imperfeito do Indicativo</p> <p>eu era</p> <p>tu eras</p> <p>ele era</p>
<p>Perfeito do Indicativo</p> <p>eu fui</p> <p>tu foste</p> <p>ele foi</p> <p>nós fomos</p> <p>vós fostes</p> <p>eles foram</p>	<p>Mais-que-perfeito do Indicativo</p> <p>eu fora</p> <p>tu foras</p> <p>ele fora</p> <p>nós fôramos</p>
<p>Futuro do Pretérito do Indicativo</p> <p>eu seria</p> <p>tu serias</p> <p>ele seria</p> <p>nós seríamos</p>	<p>Futuro do Presente do Indicativo</p> <p>eu serei</p> <p>tu serás</p>
<p>Presente do Subjuntivo</p> <p>que eu seja</p> <p>que tu sejas</p> <p>que ele seja</p> <p>que nós sejamos</p> <p>que vós sejais</p> <p>que eles sejam</p>	<p>Imperfeito do Subjuntivo</p> <p>se eu fosse</p> <p>se tu fosses</p> <p>se ele fosse</p> <p>se nós fôssemos</p>

<b>Futuro do Subjuntivo</b> quando eu for quando tu fores quando ele for quando nós formos quando vós fordes quando eles forem	
<b>Imperativo Afirmativo</b>	
<b>Imperativo Negativo</b> não sejas tu não seja ele não sejamos nós não sejais vós não sejam eles	<b>Infinitivo Pessoal</b> por ser eu por seres tu por ser ele por sermos nós por serdes vós por serem eles

**EXEMPLO 2.91 - CONJUGAÇÃO DO VERBO SER ATRAVÉS DO SERVIÇO DISPONÍVEL ONLINE EM [HTTP://LXCENTER.DI.FC.UL.PT/PT/LXSERVICESCONJUGATORPT.HTML](http://lxcenter.di.fc.ul.pt/pt/lxservicesconjugatorpt.html).**

<b>ser</b>	
<b>Indicativo</b>	
<b>Presente</b> eu <u>sou</u> tu <u>és</u> ele/ ela/ você <u>é</u> nós <u>somos</u> vós <u>sois</u> eles/ elas/ vocês <u>são</u>	<b>Pretérito Perfeito Composto</b> eu <u>hei/ tenho sido</u> tu <u>hás/ tens sido</u> ele/ ela/ você <u>há/ tem sido</u> nós <u>havemos/ temos sido</u> vós <u>haveis/ tendes sido</u> eles/ elas/ vocês <u>hã/ têm sido</u>
<b>Pretérito Imperfeito</b> eu <u>era</u> tu <u>eras</u> ele/ ela/ você <u>era</u> nós <u>éramos</u> vós <u>éreis</u> eles/ elas/ vocês <u>eram</u>	<b>Pretérito Mais-que-Perfeito Composto</b> eu <u>havia/ tinha sido</u> tu <u>havia/ tinha sido</u> ele/ ela/ você <u>havia/ tinha sido</u> nós <u>havíamos/ tínhamos sido</u> vós <u>havíeis/ tínheis sido</u> eles/ elas/ vocês <u>havam/ tinham sido</u>

<b>Pretérito Perfeito Simples</b> eu <u>fui</u> tu <u>foste</u> ele/ ela/ você <u>foi</u> nós <u>fomos</u> vós <u>fostes</u> eles/ elas/ vocês <u>foram</u>	
<b>Pretérito Mais-Que-Perfeito Simples</b> eu <u>fora</u> tu <u>foras</u> ele/ ela/ você <u>fora</u> nós <u>fôramos</u> vós <u>fôreis</u> eles/ elas/ vocês <u>foram</u>	<b>Pretérito Mais-que-Perfeito Anterior</b> eu <u>houvera/ tivera sido</u> tu <u>houveras/ tiveras sido</u> ele/ ela/ você <u>houvera/ tivera sido</u> nós <u>houvéramos/ tivéramos sido</u> vós <u>houvéreis/ tivéreis sido</u> eles/ elas/ vocês <u>houveram/ tiveram sido</u>
<b>Futuro do Presente Simples</b> eu <u>serei</u> tu <u>serás</u> ele/ ela/ você <u>será</u> nós <u>seremos</u> vós <u>sereis</u> eles/ elas/ vocês <u>serão</u>	<b>Futuro do Presente Composto</b> eu <u>haverei/ terei sido</u> tu <u>haverás/ terás sido</u> ele/ ela/ você <u>haverá/ terá sido</u> nós <u>haveremos/ teremos sido</u> vós <u>havereis/ tereis sido</u> eles/ elas/ vocês <u>haverão/ terão sido</u>
<b>Futuro do Pretérito Simples</b> eu <u>seria</u> tu <u>serias</u> ele/ ela/ você <u>seria</u> nós <u>seríamos</u> vós <u>seríeis</u> eles/ elas/ vocês <u>seriam</u>	<b>Futuro do Pretérito Composto</b> eu <u>haveria/ teria sido</u> tu <u>haverias/ terias sido</u> ele/ ela/ você <u>haveria/ teria sido</u> nós <u>haveríamos/ teríamos sido</u> vós <u>haveríeis/ teríeis sido</u> eles/ elas/ vocês <u>haveriam/ teriam sido</u>
<b>Conjuntivo</b>	
<b>Presente</b> eu <u>seja</u> tu <u>sejas</u> ele/ ela/ você <u>seja</u> nós <u>sejamos</u> vós <u>sejais</u> eles/ elas/ vocês <u>sejam</u>	<b>Pretérito Perfeito</b> eu <u>haja/ tenha sido</u> tu <u>hajas/ tenham sido</u> ele/ ela/ você <u>haja/ tenha sido</u> nós <u>hajamos/ tenhamos sido</u> vós <u>hajais/ tenhamos sido</u> eles/ elas/ vocês <u>hajam/ tenham sido</u>
<b>Pretérito Imperfeito</b> eu <u>fosse</u> tu <u>fosses</u> ele/ ela/ você <u>fosse</u> nós <u>fôssemos</u> vós <u>fôsseis</u> eles/ elas/ vocês <u>fossem</u>	<b>Pretérito Mais-Que-Perfeito</b> eu <u>houvesse/ tivesse sido</u> tu <u>houvesse/ tivesse sido</u> ele/ ela/ você <u>houvesse/ tivesse sido</u> nós <u>houvéssemos/ tivéssemos sido</u> vós <u>houvésseis/ tivésseis sido</u> eles/ elas/ vocês <u>houvessem/ tivessem sido</u>

<b>Futuro Simples</b> eu <a href="#">for</a> tu <a href="#">fores</a> ele/ ela/ você <a href="#">for</a> nós <a href="#">formos</a> vós <a href="#">fordes</a> eles/ elas/ vocês <a href="#">forem</a>	<b>Futuro Composto</b> eu <a href="#">houver/ tiver sido</a> tu <a href="#">houveres/ tiveres sido</a> ele/ ela/ você <a href="#">houver/ tiver sido</a> nós <a href="#">houvermos/ tivermos sido</a> vós <a href="#">houverdes/ tiverdes sido</a> eles/ elas/ vocês <a href="#">houverem/ tiverem sido</a>
<b>Imperativo</b>	
<b>Afirmativo</b> <a href="#">sê</a> tu <a href="#">seja</a> você <a href="#">sejamos</a> nós <a href="#">sede</a> vós <a href="#">sejam</a> vocês	<b>Negativo</b> não <a href="#">sejas</a> tu não <a href="#">seja</a> você não <a href="#">sejamos</a> nós não <a href="#">sejais</a> vós não <a href="#">sejam</a> vocês
<b>Infinitivo</b>	
<b>Infinitivo Pessoal Presente</b> eu <a href="#">ser</a> tu <a href="#">seres</a> ele/ ela/ você <a href="#">ser</a> nós <a href="#">sermos</a> vós <a href="#">serdes</a> eles/ elas/ vocês <a href="#">serem</a>	<b>Infinitivo Pessoal Pretérito</b> eu <a href="#">haver/ ter sido</a> tu <a href="#">haveres/ teres sido</a> ele/ ela/ você <a href="#">haver/ ter sido</a> nós <a href="#">havermos/ termos sido</a> vós <a href="#">haverdes/ terdes sido</a> eles/ elas/ vocês <a href="#">haverem/ terem sido</a>
<b>Infinitivo Impessoal Presente</b> <a href="#">ser</a>	<b>Infinitivo Impessoal Pretérito</b> <a href="#">haver/ ter sido</a>
<b>Gerúndio</b>	
<b>Presente</b> <a href="#">sendo</a>	<b>Pretérito</b> <a href="#">havendo/ tendo sido</a>
<b>Particípio Passado</b> <a href="#">sido</a> (inv.)	



## **ANEXO V – EXTRACTORES DE N-GRAMAS**



**EXEMPLO 2.92 - LISTAGEM DE BIGRAMAS NUM FICHEIRO INTRODUIDO COMO INPUT**

Suponhamos que temos um ficheiro denominado 'text.txt' com a seguinte informação:

text.txt
first line of text
second line
and a third line of text

Queremos armazenar os resultados do programa 'count.pl' num ficheiro 'text.cnt'.

Após a execução do comando 'count.pl test.cnt test.txt', o ficheiro 'text.cnt' terá o seguinte conteúdo:

text.cnt
11
line<>of<>2 3 2
of<>text<>2 2 2
second<>line<>1 1 3
line<>and<>1 3 1
and<>a<>1 1 1
a<>third<>1 1 1
first<>line<>1 1 3
third<>line<>1 1 3
text<>second<>1 1 1

Este ficheiro mostra-nos que no total existem onze bigramas no ficheiro de origem.

As restantes linhas mostram os bigramas encontrados.

De notar que cada linha termina com três números.

Por exemplo, na linha 1 do ficheiro temos line<>of<>2 3 2, o que significa que o bigrama

line<>of ocorre duas vezes no ficheiro de input,

o segundo número indica o número de bigramas em que o token 'line' aparece do lado esquerdo.

Neste caso, 'line' aparece três vezes à esquerda de bigramas (duas cópias do bigrama line<>of e uma vez no bigrama line<>and).

A leitura das restantes linhas é efectuada do mesmo modo.

**EXEMPLO 2.93 - LISTAGEM DE TRIGRAMAS NUM FICHEIRO INTRODUIDO COMO INPUT**

Considere-se o mesmo ficheiro de *input* do exemplo explicitado no anexo CXIII.

Após a execução do comando ‘count.pl --ngram 3 test.cnt text.txt’ o ficheiro test.cnt terá o seguinte conteúdo:

```
test.cnt
10
line<>of<>text<>2 3 2 2 2 2
and<>a<>third<>1 1 1 1 1 1
third<>line<>of<>1 1 3 2 1 1 2
second<>line<>and<>1 1 3 1 1 1 1
line<>and<>a<>1 3 1 1 1 1 1
a<>third<>line<>1 1 1 2 1 1 1
text<>second<>line<>1 1 1 2 1 1 1
of<>text<>second<>1 1 1 1 1 1 1
first<>line<>of<>1 1 3 2 1 1 2
```

O número que aparece na primeira linha diz-nos que existem dez trigramas no ficheiro de origem.

O primeiro trígama é ‘line<>of<>text<>’ constituído pelos *tokens* ‘line’, ‘of’ e ‘text’.

De notar que neste caso aparecem mais números a seguir a cada trígama do que no exemplo descrito no anexo CXIII.

Contudo, o primeiro número que aparece a seguir a cada trígama tem o mesmo significado do exemplo anterior, ou seja, o número de vezes que o trígama ocorre no texto.

Os segundo, terceiro e quarto valores dizem respeito ao número de trigramas nos quais, no caso do primeiro trígama, os *tokens* ‘line’, ‘of’ e ‘text’ aparecem nas primeira, segunda e terceira posições respectivamente.

Deste modo, ‘line’ ocorre como o *token* na primeira posição em três trigramas, que correspondem a duas cópias de line<>of<>text<> e a uma cópia de line<>and<>a<>.

Do mesmo modo, os *tokens* ‘of’ e ‘text’ aparecem como o segundo e terceiro *tokens* respectivamente de dois bigramas, que correspondem a duas cópias de line<>of<>text<>.

O quinto número corresponde ao número de bigramas nos quais ‘line’ ocorre como o primeiro *token* e ‘of’ como o segundo *token*.

Existem apenas dois trigramas nesta situação que são as duas cópias de line<>of<>text<>.

O sexto valor corresponde ao número de bigramas nos quais 'line' ocorre como o *token* na terceira posição.

Finalmente, o sétimo valor corresponde ao número de bigramas nos quais o *token* 'of' ocorre na segunda posição e o *token* 'text' ocorre na terceira posição.

### EXEMPLO 2.94 - LISTAGEM DAS COMBINAÇÕES DE UM OU MAIS TOKENS NUMA OU MAIS POSIÇÕES FIXAS

Suponhamos agora que se executa o comando 'count.pl --get\_freq\_combo freq\_combo.txt test.cnt test.txt' onde o ficheiro 'text.txt' é aquele indicado no exemplo 2.92.

Neste caso, temos a geração de dois ficheiros:

freq_combo_txt
0 1
0
1

test.cnt
11
line<>of<>2 3 2
of<>text<>2 2 2
second<>line<>1 1 3
line<>and<>1 3 1
and<>a<>1 1 1
a<>third<>1 1 1
first<>line<>1 1 3
third<>line<>1 1 3
text<>second<>1 1 1

Convém lembrar que quando a opção '--ngram' é omitida, o programa utiliza o valor definido por defeito, 2.

**EXEMPLO 2.95 - FUNCIONAMENTO DO PROGRAMA STATISTIC.PL**

Suponhamos que o nosso ficheiro de *input* é o seguinte:

```
test.cnt
11
line<>of<>2 3 2
of<>text<>2 2 2
second<>line<>1 1 3
line<>and<>1 3 1
and<>a<>1 1 1
a<>third<>1 1 1
first<>line<>1 1 3
third<>line<>1 1 3
text<>second<>1 1 1
```

Depois de executar o comando *statistic.pl dice test.dice test.cnt*, o ficheiro *test.dice* terá o seguinte conteúdo:

```
test.dice
11
of<>text<>1 1.0000 2 2 2
and<>a<>1 1.0000 1 1 1
a<>third<>1 1.0000 1 1 1
text<>second<>1 1.0000 1 1 1
line<>of<>2 0.8000 2 3 2
third<>line<>3 0.5000 1 1 3
line<>and<>3 0.5000 1 3 1
second<>line<>3 0.5000 1 1 3
first<>line<>3 0.5000 1 1 3
```

Tal como nos exemplos 2.92, 2.93 e 2.94, o primeiro valor que aparece, dez, corresponde ao número de bigramas.

O primeiro número que aparece imediatamente a seguir ao bigrama corresponde à sua classificação.

O próximo número que aparece, 1.0000, corresponde ao valor calculado pela estatística *dice*.

Os três números finais são os valores associados a cada bigrama no ficheiro 'test.cnf'.

### EXEMPLO 2.96 – COMBINAÇÃO DE CONTAGENS DE FREQUÊNCIAS A PARTIR DE UM FICHEIRO DE BIGRAMAS GERADO PELO PROGRAMA 'COUNT.PL'

Suponhamos que temos o ficheiro de bigramas gerado pelo programa *count.pl* no exemplo 2.92:

```
text.cnt
```

```
11
```

```
line<>of<>2 3 2
```

```
of<>text<>2 2 2
```

```
second<>line<>1 1 3
```

```
line<>and<>1 3 1
```

```
and<>a<>1 1 1
```

```
a<>third<>1 1 1
```

```
first<>line<>1 1 3
```

```
third<>line<>1 1 3
```

```
text<>second<>1 1 1
```

Se executarmos o comando '*combig.pl text.cnt*' obtemos o seguinte resultado:

```

text.cbg
11
first<>line<>1 1 6
a<>third<>1 2 2
second<>line<>1 2 6
text<>second<>1 3 2
of<>text<>2 4 3
third<>line<>1 2 6
and<>a<>1 2 2
line<>of<>2 6 4
line<>and<>1 6 2
    
```

Estes resultados podem ser interpretados do seguinte modo:

- se os bigramas  $word1<>word2<>n11\ n1p\ np1$  e  $word2<>word1<>m11\ m1p\ mp1$  forem encontrados no ficheiro de *input*, então o programa '*combig.pl*' trata-os como um simples bigrama desordenado

$$word1<>word2<> n11+ m11\ n1p+ mp1\ np1+ m1p$$

onde o novo bigrama mostra uma tabela de contigência combinada, na qual a ordem das palavras não importa.

	<i>word2</i>	$\sim word2$	
<i>word1</i>	$n11+ m11$	$n12+m21$	$n1p+mp1$
$\sim word1$	$n21+m12$	$n22+m22-n$	$n2p+mp2-n$
	$np1+m1p$	$np2+m2p-n$	$n$



**EXEMPLO 2.97 - ORDENAÇÃO ALFABÉTICA DE UM FICHEIRO DE BIGRAMAS CRIADO PELO PROGRAMA ‘COUNT.PL’.**

Consideremos o seguinte comando:

```
count2huge.pl -split 3 teste.cnt /test.
```

Sendo o ficheiro teste.cnt o mesmo que foi considerado nos anexos anteriores, o *output* gerado pelo comando anterior consiste na criação de um ficheiro ‘count2huge’ no directório ‘/test’ com o seguinte conteúdo:

count2huge.output
11
a<>third<>1 1 1
and<>a<>1 1 1
first<>line<>1 1 3
line<>and<>1 3 1
line<>of<>2 3 2
of<>text<>2 2 2
second<>line<>1 1 3
text<>second<>1 1 1
third<>line<>1 1 3

**EXEMPLO 2.98 – EXECUÇÃO DO PROGRAMA ‘HUGE-COUNT.PL’ SOBRE O FICHEIRO ‘TEST.TXT’ INTRODUZIDO.**

Suponhamos que o ficheiro de *input* fornecido ao programa ‘huge-count.pl’ é o seguinte:

Em 1896, no sertão da Bahia, teve início um dos acontecimentos mais impressionantes e sangrentos de toda a história do Brasil: a Campanha de Canudos. Quatro expedições foram enviadas durante um ano contra mais de vinte mil habitantes da região: índios, mulatos, caboclos, pretos... sertanejos dirigidos pelo beato Antônio Conselheiro e munidos apenas de paus, pedras e armas rústicas. Os soldados traziam metralhadoras, granadas e canhões. Estavam poderosamente armados e eram numericamente muitas vezes superiores aos
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

revoltosos, mas perdiam todas as batalhas. A resistência do sertanejo assombrava o país, e a derrota de Canudos tornou-se para o Exército e para a República uma questão de honra nacional.

"Antônio Conselheiro chegou aqui fazendo muitos milagres, todo mundo se incutindo com as graças dele, achando que ele era um verdadeiro profeta de Deus. Mas "deixa" que ele era revoltado contra a República. Ele não queria a república, queria a monarquia, o tempo dos reis da monarquia. Quando achou-se com o povo do nordeste ao lado dele, ele que se revoltou contra o governo. Aí o governo da Bahia teve de pedir auxílio ao governo do Rio de Janeiro, que mandou o canhão de guerra e o exército de lá. Veio com um tal de Marechal Bittencourt, que foi quem venceu Antônio Conselheiro."

Vitalício José dos Santos, romeiro de Monte Santo - nascido em 1936.

Época de mudanças

A passagem da monarquia para a república, no final do século XIX, foi um período de muita agitação nacional. A libertação dos escravos em 1888 fora o golpe fatal na monarquia. No ano seguinte, o golpe militar do dia 15 de novembro, liderado pelo marechal Deodoro da Fonseca, proclamou a República. O novo regime trazia a promessa de uma organização de homens livres e iguais perante a lei. As eleições democráticas dariam a todos o direito político de escolher seus dirigentes, e o trabalho livre traria salários. Eram mudanças radicais, que pareciam acabar com antigos privilégios. Já se esperava um levante monarquista. Mas nunca de um grupo de desvalidos... Canudos representou o imprevisto. Para o governo, o nordeste só poderia se beneficiar com a nova ordem. No entanto, em pleno sertão, homens lutavam até a morte em nome dessa rejeição.

O desconhecido sertão

Até o início da guerra, as elites do litoral e do sul ignoravam o que fosse o sertão: uma estranha pátria sem dono, abandonada pelas leis e instituições, vivendo sob o jugo da terra e dos latifundiários. Para compreender a revolta era necessário que o sertão viesse à tona, numa nova tradução. Foi essa a grande proeza do jornalista e engenheiro militar Euclides da Cunha, ao publicar seu livro "Os Sertões", em 1902. Uma obra contundente, que destruiu o sonho brasileiro da república e da civilização branca europeizada. O livro "Os Sertões" nasceu de uma reportagem sobre a Guerra de Canudos para o jornal "O Estado de São Paulo". Euclides da Cunha foi cobrir o evento, em 1897, como enviado de guerra.

"Um dia, à tardinha, Antônio Conselheiro chegou aqui na Fazenda Rosada, na casa de Joaquim Macambira. Perguntou se por ali não tinha uma capelinha onde se rezasse o terço na boca da noite e o ofício de madrugada. Disseram que tinha, aí avisaram aquele pessoal ao redor e também minha mãe - que naquele tempo ainda era moça - e minha avó. Foi quando o convidaram para vir e fazer a igreja de Santo Antônio. Continuaram a crescer a rua, foram aumentando e começaram a ficar por ali. Tinha muita gente, era bom para fazer negócio, bom pra ganhar dinheiro. O Conselheiro mesmo pagava para trabalhar naquelas obras, que eram muito bem feitas. Minha avó se chamava Josefa Maria de Jesus. E minha mãe, Joana Batista de Jesus. Elas eram de um povo que naquele tempo gostava muito de reza. Ave Maria! Se tinha uma trezena de Santo Antônio, elas vinham as 13 noites! E foi nessa data que minha mãe se casou. Quando teve o primeiro filho, Antônio Conselheiro foi quem batizou."

João Reginaldo de Mattos, "João de Régis" - nascido em 1909.

"Não faltava gente para chegar ao Conselheiro... Ele fez muitos milagres aqui, acompanhando esse povo... era com carregamento de pedras, era com madeiras. Sei até de um caso de parto. Ele dava remédio para mulheres que não se despachavam. Ele dava

remédio e salvava... tudo isso... Ele fazia milagre."  
 João Siqueira Santos, "Ioiô da Professora" - nascido em 1911.

"Meu avô não acreditava nesse milagroso Antônio Conselheiro! Ele não acreditava nisso porque "como é que um Deus, um comedor de feijão vai dizer que é Deus? Essa, não!... Se vocês quiserem ir, que vão só! Eu sou aquele que não vou!"  
 Vitalício José dos Santos.

"Papai casou-se lá em Canudos. Mamãe tinha aquela vozona cheia... Ela era morena, mas era bonita, mesmo. Uma cabocla bonita. Cantava na segunda voz. Minha avó e as outras tiravam hinos e ela respondia. O Conselheiro dizia: 'Cantem meus filhos!' Homens e mulheres, todos cantavam."  
 Josefa dos Santos, "D. Zefinha" - nascida em 1919.

Após a execução do commando `'huge-count.pl --tokenlist --split 3 /home/ze/test text.txt'` obtemos o seguinte resultado:

1060	!<>.<>1 7 79	!<>E<>1 7 2
!<>Ele<>1 7 6	!<>Eu<>1 7 1	!<>Homens<>1 7 1
!<>Se<>1 7 2	!<>Vital<>1 7 2	,<>Ant<>2 58 8
,<>D<>1 58 1	,<>loi<>1 58 1	,<>Jo<>1 58 3
,<>Joana<>1 58 1	,<>a<>1 58 24	,<>abandonada<>1 58 1
,<>achando<>1 58 1	,<>acompanhando<>1 58 1	,<>ao<>1 58 5
,<>as<>1 58 8	,<>bom<>1 58 2	,<>caboclos<>1 58 1
,<>como<>1 58 2	,<>e<>2 58 28	,<>elas<>1 58 1
,<>ele<>1 58 3	,<>em<>3 58 10	,<>era<>2 58 9
,<>foi<>1 58 5	,<>foram<>1 58 2	,<>granadas<>1 58 1
,<>homens<>1 58 2	,<>liderado<>1 58 1	,<>mas<>2 58 2
,<>mesmo<>1 58 2	,<>mulatos<>1 58 1	,<>n<>1 58 7
,<>na<>1 58 5	,<>no<>2 58 2	,<>numa<>1 58 1
,<>o<>3 58 56	,<>pedras<>1 58 2	,<>pretos<>1 58 1
,<>proclamou<>1 58 1	,<>que<>6 58 19	,<>queria<>1 58 2
,<>romeiro<>1 58 1	,<>tardinha<>1 58 1	,<>teve<>1 58 3
,<>todo<>1 58 1	,<>todos<>1 58 2	,<>um<>1 58 11
,<>vivendo<>1 58 1	.<>.<>16 78 79	.<>A<>3 78 4

.<>Ant<>1 78 8	.<>As<>1 78 1	.<>Ave<>1 78 1
.<>Cantava<>1 78 1	.<>Canudos<>1 78 5	.<>Continuaram<>1 78 1
.<>Disseram<>1 78 1	.<>E<>1 78 2	.<>Ela<>1 78 1
.<>Elas<>1 78 1	.<>Ele<>5 78 6	.<>Eram<>1 78 1
.<>Estavam<>1 78 1	.<>Euclides<>1 78 2	.<>Foi<>2 78 2
.<>J<>1 78 1	.<>Jo<>2 78 3	.<>Josefa<>1 78 2
.<>Mam<>1 78 1	.<>Mas<>2 78 2	.<>Meu<>1 78 1
.<>Minha<>2 78 2	.<>N<>1 78 1	.<>No<>2 78 2
.<>O<>5 78 6	.<>Os<>1 78 3	.<>Papai<>1 78 1
.<>Para<>2 78 2	.<>Perguntou<>1 78 1	.<>Quando<>2 78 2
.<>Quatro<>1 78 1	.<>Se<>1 78 2	.<>Sei<>1 78 1
.<>Tinha<>1 78 1	.<>Um<>1 78 1	.<>Uma<>2 78 2
.<>Veio<>1 78 1	.<>Vital<>1 78 2	.<>Zefinha<>1 78 1
.<>era<>1 78 9	.<>poca<>1 78 1	.<>sertanejos<>1 78 1
.<>tudo<>1 78 1	13<>noites<>1 1 1	15<>de<>1 1 39
1888<>fora<>1 1 1	1896<>, <>1 1 58	1897<>, <>1 1 58
1902<>.<>1 1 79	1909<>.<>1 1 79	1911<>.<>1 1 79
1919<>.<>1 1 79	1936<>.<>1 1 79	:<>Cantem<>1 4 1
:<>a<>1 4 24	:<>ndios<>1 4 1	:<>uma<>1 4 6
?<>Essa<>1 1 1	A<>liberta<>1 4 1	A<>o<>1 4 56
A<>passagem<>1 4 1	A<>resist<>1 4 1	Ant<>nio<>8 8 8
As<>elei<>1 1 1	At<>o<>1 1 56	Ave<>Maria<>1 1 2
Bahia<>, <>1 2 58	Bahia<>teve<>1 2 3	Batista<>de<>1 1 39
Bittencourt<>, <>1 1 58	Brasil<>:<>1 1 4	Campanha<>de<>1 1 39
Cantava<>na<>1 1 5	Cantem<>meus<>1 1 1	Canudos<>.<>2 5 79
Canudos<>para<>1 5 9	Canudos<>representou<>1 5 1	Canudos<>tornou<>1 5 1
Conselheiro<>!<>1 9 7	Conselheiro<>.<>2 9 79	Conselheiro<>chegou<>2 9 2

Conselheiro<>dizia<>1 9 1	Conselheiro<>e<>1 9 28	Conselheiro<>foi<>1 9 5
Conselheiro<>mesmo<>1 9 2	Continuaram<>a<>1 1 24	Cunha<>,<>1 2 58
Cunha<>foi<>1 2 5	D<>.<>1 1 79	Deodoro<>da<>1 1 14
Deus<>,<>1 3 58	Deus<>.<>1 3 79	Deus<>?<>1 3 1
Disseram<>que<>1 1 19	E<>foi<>1 2 5	E<>minha<>1 2 4
Ela<>era<>1 1 9	Elas<>eram<>1 1 3	Ele<>dava<>2 6 2
Ele<>fazia<>1 6 1	Ele<>fez<>1 6 1	Ele<>n<>2 6 7
Em<>1896<>1 1 1	Eram<>mudan<>1 1 2	Essa<>,<>1 1 58
Estado<>de<>1 1 39	Estavam<>poderosamente<>1 1 1	Eu<>sou<>1 1 1
Euclides<>da<>2 2 14	Ex<>rcito<>1 1 2	Fazenda<>Rosada<>1 1 1
Foi<>essa<>1 2 1	Foi<>quando<>1 2 1	Fonseca<>,<>1 1 58
Guerra<>de<>1 1 39	Homens<>e<>1 1 28	loi<>da<>1 1 14
J<>se<>1 1 12	Janeiro<>,<>1 1 58	Jesus<>.<>2 2 79
Jo<>o<>3 3 56	Joana<>Batista<>1 1 1	Joaquim<>Macambira<>1 1
Jos<>dos<>2 2 7	Josefa<>Maria<>1 2 2	Josefa<>dos<>1 2 7
Macambira<>.<>1 1 79	Mam<>e<>1 1 28	Marechal<>Bittencourt<>1 1 1
Maria<>!<>1 2 7	Maria<>de<>1 2 39	Mas<>deixa<>1 2 1
Mas<>nunca<>1 2 1	Mattos<>,<>1 1 58	Meu<>av<>1 1 4
Minha<>av<>2 2 4	Monte<>Santo<>1 1 3	N<>o<>1 1 56
No<>ano<>1 2 2	No<>entanto<>1 2 1	O<>Conselheiro<>2 6 9
O<>Estado<>1 6 1	O<>desconhecido<>1 6 1	O<>livro<>1 6 2
O<>novo<>1 6 1	Os<>Sert<>2 3 2	Os<>soldados<>1 3 1
Papai<>casou<>1 1 2	Para<>compreender<>1 2 1	Para<>o<>1 2 56
Paulo<>.<>1 1 79	Perguntou<>se<>1 1 12	Professora<>nascido<>1 1 3

Quando<>achou<>1 2 1	Quando<>teve<>1 2 3	Quatro<>expedi<>1 1 1
R<>gis<>1 1 1	Reginaldo<>de<>1 1 39	Rep<>blica<>3 3 6
Rio<>de<>1 1 39	Rosada<>,<>1 1 58	S<>o<>1 1 56
Santo<>Ant<>2 3 8	Santo<>nascido<>1 3 3	Santos<>,<>3 4 58
Santos<>.<>1 4 79	Se<>tinha<>1 2 4	Se<>voc<>1 2 1
Sei<>at<>1 1 2	Sert<>es<>2 2 6	Siqueira<>Santos<>1 1 4
Tinha<>muita<>1 1 2	Um<>dia<>1 1 2	Uma<>cabocla<>1 2 1
Uma<>obra<>1 2 1	Veio<>com<>1 1 7	Vital<>cio<>2 2 6
XIX<>,<>1 1 58	Zefinha<>nascida<>1 1 1	a<>Campanha<>1 24 1
a<>Guerra<>1 24 1	a<>Rep<>3 24 3	a<>avisaram<>1 24 1
a<>crescer<>1 24 1	a<>derrota<>1 24 1	a<>e<>1 24 28
a<>ficar<>1 24 1	a<>grande<>1 24 1	a<>hist<>1 24 1
a<>igreja<>1 24 1	a<>lei<>1 24 1	a<>monarquia<>1 24 4
a<>morte<>1 24 1	a<>nova<>1 24 2	a<>o<>1 24 56
a<>promessa<>1 24 1	a<>rep<>2 24 3	a<>revolta<>1 24 1
a<>rua<>1 24 1	a<>todos<>1 24 2	abandonada<>pelas<>1 1 1
acabar<>com<>1 1 7	achando<>que<>1 1 19	achou<>se<>1 1 12
acompanhando<>esse<>1 1	acontecimentos<>mais<>1 1 2	acreditava<>nesse<>1 2 1
acreditava<>nisso<>1 2 1	agita<>o<>1 1 56	ainda<>era<>1 1 9
ali<>.<>1 2 79	ali<>n<>1 2 7	ano<>contra<>1 2 3
ano<>seguinte<>1 2 1	antigos<>privil<>1 1 1	ao<>Conselheiro<>1 5 9
ao<>governo<>1 5 4	ao<>lado<>1 5 1	ao<>publicar<>1 5 1
ao<>redor<>1 5 1	aos<>revoltosos<>1 1 1	apenas<>de<>1 1 39
aquela<>vozona<>1 1 1	aquele<>pessoal<>1 2 1	aquele<>que<>1 2 19
aqui<>,<>1 3 58	aqui<>fazendo<>1 3 1	aqui<>na<>1 3 5
aram<>a<>1 1 24	armados<>e<>1 1 28	armas<>r<>1 1 1

as<>13<>1 8 1	as<>A<>1 8 4	as<>batalhas<>1 8 1
as<>dele<>1 8 2	as<>elites<>1 8 1	as<>gra<>1 8 1
as<>outras<>1 8 1	as<>radicais<>1 8 1	assombrava<>o<>1 1 56
at<>a<>1 2 24	at<>de<>1 2 39	aumentando<>e<>1 1 28
aux<>lio<>1 1 1	av<>.<>1 4 79	av<>e<>1 4 28
av<>n<>1 4 7	av<>se<>1 4 12	avisaram<>aquela<>1 1 2
batalhas<>.<>1 1 79	batizou<>.<>1 1 79	beato<>Ant<>1 1 8
bem<>feitas<>1 1 1	beneficiar<>com<>1 1 7	blica<>,<>2 6 58
blica<>.<>2 6 79	blica<>e<>1 6 28	blica<>uma<>1 6 6
boca<>da<>1 1 14	bom<>para<>1 2 9	bom<>pra<>1 2 1
bonita<>,<>1 2 58	bonita<>.<>1 2 79	branca<>europeizada<>1 1 1
brasileiro<>da<>1 1 14	cabocla<>bonita<>1 1 2	caboclos<>,<>1 1 58
canh<>es<>1 2 6	canh<>o<>1 2 56	cantavam<>.<>1 1 79
capelinha<>onde<>1 1 1	carregamento<>de<>1 1 39	casa<>de<>1 1 39
caso<>de<>1 1 39	casou<>.<>1 2 79	casou<>se<>1 2 12
chamava<>Josefa<>1 1 2	chegar<>ao<>1 1 5	chegou<>aqui<>2 2 3
cheia<>.<>1 1 79	cio<>,<>1 6 58	cio<>Jos<>2 6 2
cio<>da<>1 6 14	cio<>de<>1 6 39	cio<>um<>1 6 11
civiliza<>o<>1 1 56	cobrir<>o<>1 1 56	com<>a<>1 7 24
com<>antigos<>1 7 1	com<>as<>1 7 8	com<>carregamento<>1 7 1
com<>madeiras<>1 7 1	com<>o<>1 7 56	com<>um<>1 7 11
come<>aram<>1 1 1	comedor<>de<>1 1 39	como<>enviado<>1 2 1
como<>que<>1 2 19	compreender<>a<>1 1 24	contra<>a<>1 3 24
contra<>mais<>1 3 2	contra<>o<>1 3 56	contundente<>,<>1 1 58
convidaram<>para<>1 1 9	crescer<>a<>1 1 24	culo<>XIX<>1 1 1
da<>Bahia<>2 14 2	da<>Cunha<>2 14 2	da<>Fonseca<>1 14 1
da<>Professora<>1 14 1	da<>civiliza<>1 14 1	da<>guerra<>1 14 3

da<>monarquia<>2 14 4	da<>noite<>1 14 1	da<>regi<>1 14 1
da<>rep<>1 14 3	da<>terra<>1 14 1	dariam<>a<>1 1 24
data<>que<>1 1 19	dava<>rem<>2 2 2	de<>Canudos<>3 39 5
de<>Deus<>1 39 3	de<>Janeiro<>1 39 1	de<>Jesus<>2 39 2
de<>Joaquim<>1 39 1	de<>Marechal<>1 39 1	de<>Mattos<>1 39 1
de<>Monte<>1 39 1	de<>R<>1 39 1	de<>S<>1 39 1
de<>Santo<>2 39 3	de<>desvalidos<>1 39 1	de<>escolher<>1 39 1
de<>feij<>1 39 1	de<>guerra<>2 39 3	de<>homens<>1 39 2
de<>honra<>1 39 1	de<>l<>1 39 2	de<>madrugada<>1 39 1
de<>mudan<>1 39 2	de<>muita<>1 39 2	de<>novembro<>1 39 1
de<>parto<>1 39 1	de<>paus<>1 39 1	de<>pedir<>1 39 1
de<>pedras<>1 39 2	de<>reza<>1 39 1	de<>toda<>1 39 1
de<>um<>3 39 11	de<>uma<>2 39 6	de<>vinte<>1 39 1
deixa<>que<>1 1 19	dele<>,<>2 2 58	democr<>ticas<>1 1 1
derrota<>de<>1 1 39	desconhecido<>sert<>1 1 5	despachavam<>.<>1 1 79
dessa<>rejei<>1 1 1	destru<>a<>1 1 24	desvalidos<>.<>1 1 79
dia<>,<>1 2 58	dia<>15<>1 2 1	dinheiro<>.<>1 1 79
dio<>e<>1 2 28	dio<>para<>1 2 9	direito<>pol<>1 1 1
dirigentes<>,<>1 1 58	dirigidos<>pelo<>1 1 2	dizer<>que<>1 1 19
dizia<>:<>1 1 4	do<>Brasil<>1 9 1	do<>Rio<>1 9 1
do<>dia<>1 9 2	do<>jornalista<>1 9 1	do<>litoral<>1 9 1
do<>nordeste<>1 9 2	do<>s<>1 9 5	do<>sertanejo<>1 9 1
do<>sul<>1 9 1	dono<>,<>1 1 58	dos<>Santos<>3 7 4
dos<>acontecimentos<>1 7 1	dos<>escravos<>1 7 1	dos<>latifundi<>1 7 1
dos<>reis<>1 7 1	durante<>um<>1 1 11	e<>,<>1 28 58
e<>a<>1 28 24	e<>armas<>1 28 1	e<>as<>1 28 8



e<>canh<>1 28 2	e<>come<>1 28 1	e<>da<>1 28 14
e<>do<>1 28 9	e<>dos<>1 28 7	e<>ela<>1 28 1
e<>engenheiro<>1 28 1	e<>eram<>1 28 3	e<>fazer<>1 28 2
e<>iguais<>1 28 1	e<>institui<>1 28 1	e<>minha<>1 28 4
e<>mulheres<>1 28 2	e<>munidos<>1 28 1	e<>o<>3 28 56
e<>para<>1 28 9	e<>que<>1 28 19	e<>salvava<>1 28 1
e<>sangrentos<>1 28 1	e<>se<>1 28 12	e<>tamb<>1 28 1
e<>tinha<>1 28 4	ela<>respondia<>1 1 1	elas<>vinham<>1 1 1
ele<>era<>2 3 9	ele<>que<>1 3 19	elei<>es<>1 1 6
elites<>do<>1 1 9	em<>1888<>1 10 1	em<>1897<>1 10 1
em<>1902<>1 10 1	em<>1909<>1 10 1	em<>1911<>1 10 1
em<>1919<>1 10 1	em<>1936<>1 10 1	em<>Canudos<>1 10 5
em<>nome<>1 10 1	em<>pleno<>1 10 1	engenheiro<>militar<>1 1 2
entanto<>,<>1 1 58	enviadas<>durante<>1 1 1	enviado<>de<>1 1 39
era<>bom<>1 9 2	era<>bonita<>1 9 2	era<>com<>2 9 7
era<>mo<>1 9 1	era<>morena<>1 9 1	era<>necess<>1 9 1
era<>revoltado<>1 9 1	era<>um<>1 9 11	eram<>de<>1 3 39
eram<>muito<>1 3 2	eram<>numericamente<>1 3 1	es<>,<>2 6 58
es<>.<>1 6 79	es<>democr<>1 6 1	es<>foram<>1 6 2
es<>nasceu<>1 6 1	escolher<>seus<>1 1 1	escravos<>em<>1 1 10
esperava<>um<>1 1 11	essa<>a<>1 1 24	esse<>povo<>1 1 3
estranha<>p<>1 1 1	uropeizada<>.<>1 1 79	evento<>,<>1 1 58
ex<>rcito<>1 1 2	expedi<>es<>1 1 6	faltava<>gente<>1 1 2
fatal<>na<>1 1 5	fazendo<>muitos<>1 1 2	fazer<>a<>1 2 24
fazer<>neg<>1 2 1	fazia<>milagre<>1 1 1	feij<>o<>1 1 56
feitas<>.<>1 1 79	fez<>muitos<>1 1 2	ficar<>por<>1 1 2

filho<>,<>1 1 58	filhos<>!<>1 1 7	final<>do<>1 1 9
foi<>cobrir<>1 5 1	foi<>nessa<>1 5 1	foi<>quem<>2 5 2
foi<>um<>1 5 11	fora<>o<>1 1 56	foram<>aumentando<>1 2 1
foram<>enviadas<>1 2 1	fosse<>o<>1 1 56	ganhar<>dinheiro<>1 1 1
gente<>,<>1 2 58	gente<>para<>1 2 9	gios<>.<>1 1 79
gis<>nascido<>1 1 3	golpe<>fatal<>1 2 1	golpe<>militar<>1 2 2
gostava<>muito<>1 1 2	governo<>,<>1 4 58	governo<>.<>1 4 79
governo<>da<>1 4 14	governo<>do<>1 4 9	gra<>as<>1 1 8
granadas<>e<>1 1 28	grande<>proeza<>1 1 1	grupo<>de<>1 1 39
guerra<>,<>1 3 58	guerra<>.<>1 3 79	guerra<>e<>1 3 28
habitantes<>da<>1 1 14	hinos<>e<>1 1 28	hist<>ria<>1 1 1
homens<>livres<>1 2 1	homens<>lutavam<>1 2 1	honra<>nacional<>1 1 2
ignoravam<>o<>1 1 56	igreja<>de<>1 1 39	iguais<>perante<>1 1 1
impressionantes<>e<>1 1 28	imprevisto<>.<>1 1 79	in<>cio<>2 2 6
incutindo<>com<>1 1 7	institui<>es<>1 1 6	ir<>,<>1 1 58
isso<>.<>1 1 79	jornal<>O<>1 1 6	jornalista<>e<>1 1 28
jugo<>da<>1 1 14	l<>.<>1 2 79	l<>em<>1 2 10
lado<>dele<>1 1 2	latifundi<>rios<>1 1 2	lei<>.<>1 1 79
leis<>e<>1 1 28	levante<>monarquista<>1 1 1	liberta<>o<>1 1 56
liderado<>pelo<>1 1 2	lio<>ao<>1 1 5	litoral<>e<>1 1 28
livre<>traria<>1 1 1	livres<>e<>1 1 28	livro<>Os<>2 2 3
lutavam<>at<>1 1 2	m<>e<>3 4 28	m<>minha<>1 4 4
madeiras<>.<>1 1 79	madrugada<>.<>1 1 79	mais<>de<>1 2 39
mais<>impressionantes<>1 2 1	mandou<>o<>1 1 56	marechal<>Deodoro<>1 1 1
mas<>era<>1 2 9	mas<>perdiam<>1 2 1	mesmo<>.<>1 2 79
mesmo<>pagava<>1 2 1	metralhadoras<>,<>1 1 58	meus<>filhos<>1 1 1

mil<>habitantes<>1 1 1	milagre<>.<>1 1 79	milagres<>,<>1 2 58
milagres<>aqui<>1 2 3	milagroso<>Ant<>1 1 8	militar<>Euclides<>1 2 2
militar<>do<>1 2 9	minha<>av<>1 4 4	minha<>m<>3 4 4
mo<>a<>1 1 24	monarquia<>,<>1 4 58	monarquia<>.<>2 4 79
monarquia<>para<>1 4 9	monarquista<>.<>1 1 79	morena<>,<>1 1 58
morte<>em<>1 1 10	mudan<>as<>2 2 8	muita<>agita<>1 2 1
muita<>gente<>1 2 2	muitas<>vezes<>1 1 1	muito<>bem<>1 2 1
muito<>de<>1 2 39	muitos<>milagres<>2 2 2	mulatos<>,<>1 1 58
mulheres<>,<>1 2 58	mulheres<>que<>1 2 19	munido<>se<>1 1 12
munidos<>apenas<>1 1 1	n<>o<>7 7 56	na<>Fazenda<>1 5 1
na<>boca<>1 5 1	na<>casa<>1 5 1	na<>monarquia<>1 5 4
na<>segunda<>1 5 1	nacional<>.<>2 2 79	naquelas<>obras<>1 1 1
naquele<>tempo<>2 2 3	nasceu<>de<>1 1 39	nascida<>em<>1 1 10
nascido<>em<>3 3 10	ncia<>do<>1 1 9	ndios<>,<>1 1 58
necess<>rio<>1 1 1	neg<>cio<>1 1 6	nessa<>data<>1 1 1
nesse<>milagroso<>1 1 1	nio<>,<>1 8 58	nio<>.<>1 8 79
nio<>Conselheiro<>6 8 9	nisso<>porque<>1 1 1	no<>final<>1 2 1
no<>sert<>1 2 5	noite<>e<>1 1 28	noites<>!<>1 1 7
nome<>dessa<>1 1 1	nordeste<>ao<>1 2 5	nordeste<>s<>1 2 5
nova<>ordem<>1 2 1	nova<>tradu<>1 2 1	novembro<>,<>1 1 58
novo<>regime<>1 1 1	numa<>nova<>1 1 2	numericamente<>muitas<>1 1 1
nunca<>de<>1 1 39	o<>!<>1 56 7	o<>,<>1 56 58
o<>.<>2 56 79	o<>:<>2 56 4	o<>At<>1 56 1
o<>Ex<>1 56 1	o<>Paulo<>1 56 1	o<>Reginaldo<>1 56 1
o<>Siqueira<>1 56 1	o<>acreditava<>2 56 2	o<>branca<>1 56 1
o<>canh<>1 56 2	o<>convidaram<>1 56 1	o<>da<>1 56 14

o<>de<>4 56 39	o<>direito<>1 56 1	o<>dos<>1 56 7
o<>evento<>1 56 1	o<>ex<>1 56 1	o<>faltava<>1 56 1
o<>golpe<>2 56 2	o<>governo<>3 56 4	o<>imprevisto<>1 56 1
o<>in<>1 56 2	o<>jornal<>1 56 1	o<>jugo<>1 56 1
o<>na<>1 56 5	o<>nacional<>1 56 2	o<>nordeste<>1 56 2
o<>of<>1 56 1	o<>pa<>1 56 1	o<>povo<>1 56 3
o<>primeiro<>1 56 1	o<>que<>1 56 19	o<>queria<>1 56 2
o<>s<>1 56 5	o<>se<>1 56 12	o<>sert<>2 56 5
o<>sonho<>1 56 1	o<>tempo<>1 56 3	o<>ter<>1 56 1
o<>tinha<>1 56 4	o<>trabalho<>1 56 1	o<>vai<>1 56 1
o<>viesse<>1 56 1	o<>vou<>1 56 1	obra<>contundente<>1 1 1
obras<>, <>1 1 58	odo<>de<>1 1 39	of<>cio<>1 1 6
onde<>se<>1 1 12	ordem<>.<>1 1 79	organiza<>o<>1 1 56
outras<>tiravam<>1 1 1	p<>tria<>1 1 1	pa<>s<>1 1 5
pagava<>para<>1 1 9	para<>a<>2 9 24	para<>chegar<>1 9 1
para<>fazer<>1 9 2	para<>mulheres<>1 9 2	para<>o<>2 9 56
para<>trabalhar<>1 9 1	para<>vir<>1 9 1	pareciam<>acabar<>1 1 1
parto<>.<>1 1 79	passagem<>da<>1 1 14	paus<>, <>1 1 58
pedir<>aux<>1 1 1	pedras<>, <>1 2 58	pedras<>e<>1 2 28
pelas<>leis<>1 1 1	pelo<>beato<>1 2 1	pelo<>marechal<>1 2 1
per<>odo<>1 1 1	perante<>a<>1 1 24	perdiam<>todas<>1 1 1
peessoal<>ao<>1 1 5	pleno<>sert<>1 1 5	poca<>de<>1 1 39
poderia<>se<>1 1 12	poderosamente<>armados<>1 1 1	pol<>tico<>1 1 1
por<>ali<>2 2 2	porque<>como<>1 1 2	povo<>.<>1 3 79
povo<>do<>1 3 9	povo<>que<>1 3 19	pra<>ganhar<>1 1 1
pretos<>.<>1 1 79	primeiro<>filho<>1 1 1	privil<>gios<>1 1 1

proclamou<>a<>1 1 24	proeza<>do<>1 1 9	profeta<>de<>1 1 39
promessa<>de<>1 1 39	publicar<>seu<>1 1 1	quando<>o<>1 1 56
que<>Deus<>1 1 9 3	que<>destru<>1 1 9 1	que<>ele<>2 1 9 3
que<>eram<>1 1 9 3	que<>foi<>1 1 9 5	que<>fosse<>1 1 9 1
que<>mandou<>1 1 9 1	que<>minha<>1 1 9 4	que<>n<>2 1 9 7
que<>naquele<>2 1 9 2	que<>o<>1 1 9 56	que<>pareciam<>1 1 9 1
que<>se<>1 1 9 12	que<>tinha<>1 1 9 4	que<>um<>1 1 9 11
que<>v<>1 1 9 1	quem<>batizou<>1 2 1	quem<>venceu<>1 2 1
queria<>a<>2 2 24	quest<>o<>1 1 56	quiserem<>ir<>1 1 1
r<>sticas<>1 1 1	radicais<>,<>1 1 58	rcito<>de<>1 2 39
rcito<>e<>1 2 28	redor<>e<>1 1 28	regi<>o<>1 1 56
regime<>trazia<>1 1 1	reis<>da<>1 1 14	rejei<>o<>1 1 56
rem<>dio<>2 2 2	rep<>blica<>3 3 6	reportagem<>sobre<>1 1 1
representou<>o<>1 1 56	resist<>ncia<>1 1 1	respondia<>.<>1 1 79
revolta<>era<>1 1 9	revoltado<>contra<>1 1 3	revoltosos<>,<>1 1 58
revoltou<>contra<>1 1 3	reza<>.<>1 1 79	rezasse<>o<>1 1 56
ria<>do<>1 1 9	rio<>que<>1 1 19	rios<>.<>2 2 79
romeiro<>de<>1 1 39	rua<>,<>1 1 58	s<>!<>1 5 7
s<>,<>1 5 58	s<>culo<>1 5 1	s<>poderia<>1 5 1
s<>quiserem<>1 5 1	sal<>rios<>1 1 2	salvava<>.<>1 1 79
sangrentos<>de<>1 1 39	se<>beneficiar<>1 12 1	se<>casou<>1 12 2
se<>chamava<>1 12 1	se<>com<>1 12 7	se<>despachavam<>1 12 1
se<>esperava<>1 12 1	se<>incutindo<>1 12 1	se<>l<>1 12 2
se<>para<>1 12 9	se<>por<>1 12 2	se<>revoltou<>1 12 1
se<>rezasse<>1 12 1	seguinte<>,<>1 1 58	segunda<>voz<>1 1 1
sem<>dono<>1 1 1	sert<>o<>5 5 56	sertanejo<>assombrava<>1 1 1

sertanejos<>dirigidos<>1 1 1	seu<>livro<>1 1 2	seus<>dirigentes<>1 1 1
sob<>o<>1 1 56	sobre<>a<>1 1 24	soldados<>traziam<>1 1 1
sonho<>brasileiro<>1 1 1	sou<>aquela<>1 1 2	sticas<>.<>1 1 79
sul<>ignoravam<>1 1 1	superiores<>aos<>1 1 1	tal<>de<>1 1 39
tamb<>m<>1 1 4	tardinha<>,<>1 1 58	tempo<>ainda<>1 3 1
tempo<>dos<>1 3 7	tempo<>gostava<>1 3 1	ter<>o<>1 1 56
terra<>e<>1 1 28	teve<>de<>1 3 39	teve<>in<>1 3 2
teve<>o<>1 3 56	ticas<>dariam<>1 1 1	tico<>de<>1 1 39
tinha<>,<>1 4 58	tinha<>aquela<>1 4 1	tinha<>uma<>2 4 6
tiravam<>hinos<>1 1 1	toda<>a<>1 1 24	todas<>as<>1 1 8
todo<>mundo<>1 1 1	todos<>cantavam<>1 2 1	todos<>o<>1 2 56
tona<>,<>1 1 58	tornou<>se<>1 1 12	trabalhar<>naquelas<>1 1 1
trabalho<>livre<>1 1 1	tradu<>o<>1 1 56	traria<>sal<>1 1 1
trazia<>a<>1 1 24	traziam<>metralhadoras<>1 1 1	trezena<>de<>1 1 39
tria<>sem<>1 1 1	tudo<>isso<>1 1 1	um<>Deus<>1 1 1 3
um<>ano<>1 1 1 2	um<>caso<>1 1 1 1	um<>comedor<>1 1 1 1
um<>dos<>1 1 1 7	um<>grupo<>1 1 1 1	um<>levante<>1 1 1 1
um<>per<>1 1 1 1	um<>povo<>1 1 1 3	um<>tal<>1 1 1 1
um<>verdadeiro<>1 1 1 1	uma<>capelinha<>1 6 1	uma<>estranha<>1 6 1
uma<>organiza<>1 6 1	uma<>quest<>1 6 1	uma<>reportagem<>1 6 1
uma<>trezena<>1 6 1	v<>o<>1 1 56	vai<>dizer<>1 1 1
venceu<>Ant<>1 1 8	verdadeiro<>profeta<>1 1 1	vezes<>superiores<>1 1 1
viesses<>tona<>1 1 1	vinham<>as<>1 1 8	vinte<>mil<>1 1 1
vir<>e<>1 1 28	vivendo<>sob<>1 1 1	voc<>s<>1 1 5
vou<>!<>1 1 7	voz<>.<>1 1 79	vozona<>cheia<>1 1 1

## EXEMPLO 2.99 - ELIMINAÇÃO DE BIGRAMAS CUJAS FREQUÊNCIAS SE SITUAM ENTRE UM DADO LIMITE INFERIOR E UM DADO LIMITE SUPERIOR

Se executarmos o comando `'huge-delete.pl --frequency 3 --ufrequency 50 test.txt /test'`, onde `'test.txt'` contém o *output* gerado no exemplo 2.95, que se encontra no anexo CXVI, e `'/test'` é o directório onde serão armazenados os resultados do *output* do comando especificado, obteremos naquele directório um ficheiro de *output* com o seguinte conteúdo:

1060	,<em>3 58 10	,<que>6 58 19
.<.<>16 78 79	.<A>3 78 4	.<Ele>5 78 6
.<O>5 78 6	Ant<nio>8 8 8	Jo<o>3 3 56
Rep<blica>3 3 6	Santos<,<>3 4 58	a<Rep>3 24 3
de<Canudos>3 39 5	de<um>3 39 11	dos<Santos>3 7 4
e<o>3 28 56	m<e>3 4 28	minha<m>3 4 4
n<o>7 7 56	nascido<em>3 3 10	nio<Conselheiro>6 8 9
o<de>4 56 39		

Como se pode observar, foram eliminados todos os bigramas cujas frequências eram inferiores a três e superiores a cinquenta.

**EXEMPLO 2.100 - JUNÇÃO DE FICHEIROS DE BIGRAMAS EXISTENTES NO DIRECTÓRIO****/TEST**

Suponhamos que temos os seguintes ficheiros no directório /test:

test1.txt

```
,<>caboclos<>1 9 1
,<>e<>1 9 7
,<>granadas<>1 9 1
,<>mas<>1 9 1
,<>mulatos<>1 9 1
,<>no<>1 9 1
,<>pedras<>1 9 1
,<>pretos<>1 9 1
,<>teve<>1 9 1
.<>.<>2 7 8
.<>A<>1 7 1
.<>Estavam<>1 7 1
.<>Os<>1 7 1
.<>Quatro<>1 7 1
.<>sertanejos<>1 7 1
1896<>,<>1 1 9
:<>a<>1 2 4
:<>ndios<>1 2 1
```

test2.txt

```
A<>resist<>1 1 1
Ant<>nio<>1 1 1
Bahia<>,<>1 1 9
Brasil<>:<>1 1 2
Campanha<>de<>1 1 6
Canudos<>.<>1 2 8
Canudos<>tornou<>1 2 1
Conselheiro<>e<>1 1 7
Em<>1896<>1 1 1
Estavam<>poderosamente<>1 1 1
Ex<>rcito<>1 1 1
Os<>soldados<>1 1 1
Quatro<>expedi<>1 1 1
Rep<>blica<>1 1 1
a<>Campanha<>1 4 1
```



a<>Rep<>1 4 1  
 a<>derrota<>1 4 1  
 a<>hist<>1 4 1  
 acontecimentos<>mais<>1 1 2  
 ano<>contra<>1 1 1  
 aos<>revoltosos<>1 1 1  
 apenas<>de<>1 1 6  
 armados<>e<>1 1 7  
 armas<>r<>1 1 1  
 as<>batalhas<>1 1 1  
 assombrosa<>o<>1 1 5  
 batalhas<>.<>1 1 8  
 beato<>Ant<>1 1 1  
 blica<>uma<>1 1 1  
 caboclos<>,<>1 1 9  
 canh<>es<>1 1 2  
 cio<>um<>1 1 2  
 contra<>mais<>1 1 2

test3.txt

da<>Bahia<>1 2 1  
 da<>regi<>1 2 1  
 de<>Canudos<>2 6 2  
 de<>honra<>1 6 1  
 de<>paus<>1 6 1  
 de<>toda<>1 6 1  
 de<>vinte<>1 6 1  
 derrota<>de<>1 1 6  
 dirigidos<>pelo<>1 1 1  
 do<>Brasil<>1 2 1  
 do<>sertanejo<>1 2 1  
 dos<>acontecimentos<>1 1 1  
 durante<>um<>1 1 2  
 e<>a<>1 7 4  
 e<>armas<>1 7 1  
 e<>canh<>1 7 1  
 e<>eram<>1 7 1  
 e<>munidos<>1 7 1  
 e<>para<>1 7 2  
 e<>sangrentos<>1 7 1  
 enviadas<>durante<>1 1 1  
 eram<>numericamente<>1 1 1  
 es<>.<>1 2 8  
 es<>foram<>1 2 1

test4.txt

expedi<>es<>1 1 2  
 foram<>enviadas<>1 1 1  
 granadas<>e<>1 1 7  
 habitantes<>da<>1 1 2  
 hist<>ria<>1 1 1  
 honra<>nacional<>1 1 1  
 impressionantes<>e<>1 1 7  
 in<>cio<>1 1 1  
 mais<>de<>1 2 6  
 mais<>impressionantes<>1 2 1  
 mas<>perdiam<>1 1 1  
 metralhadoras<>,<>1 1 9  
 mil<>habitantes<>1 1 1  
 muitas<>vezes<>1 1 1  
 mulatos<>,<>1 1 9

test5.txt

unidos<>apenas<>1 1 1  
 nacional<>.<>1 1 8  
 ncia<>do<>1 1 2  
 ndios<>,<>1 1 9  
 nio<>Conselheiro<>1 1 1  
 no<>sert<>1 1 1  
 numericamente<>muitas<>1 1 1  
 o<>:<>1 5 2  
 o<>Ex<>1 5 1  
 o<>da<>1 5 2  
 o<>de<>1 5 6  
 o<>pa<>1 5 1  
 pa<>s<>1 1 1  
 para<>a<>1 2 4  
 para<>o<>1 2 5  
 paus<>,<>1 1 9  
 pedras<>e<>1 1 7  
 pelo<>beato<>1 1 1  
 perdiam<>todas<>1 1 1  
 poderosamente<>armados<>1 1 1  
 pretos<>.<>1 1 8  
 quest<>o<>1 1 5  
 r<>sticas<>1 1 1  
 rcito<>e<>1 1 7

```

regi<>o<>1 1 5
resist<>ncia<>1 1 1
revoltosos<>,<>1 1 9
ria<>do<>1 1 2
s<>,<>1 1 9

```

test6.txt

```

sangrentos<>de<>1 1 6
se<>para<>1 1 2
sert<>o<>1 1 5
sertanejo<>assombrava<>1 1 1
sertanejos<>dirigidos<>1 1 1
soldados<>traziam<>1 1 1
sticas<>.<>1 1 8
superiores<>aos<>1 1 1
teve<>in<>1 1 1
toda<>a<>1 1 4
todas<>as<>1 1 1
tornou<>se<>1 1 1
traziam<>metralhadoras<>1 1 1
um<>ano<>1 2 1
um<>dos<>1 2 1
uma<>quest<>1 1 1
vezes<>superiores<>1 1 1
vinte<>mil<>1 1 1

```

Ao executarmos o comando `'huge-merge /test/'` obtemos o seguinte resultado:

output.merged		
,<>caboclos<>1 9 1	,<>e<>1 9 7	,<>granadas<>1 9 1
,<>mas<>1 9 1	,<>mulatos<>1 9 1	,<>no<>1 9 1
,<>pedras<>1 9 1	,<>pretos<>1 9 1	,<>teve<>1 9 1
.<>.<>2 7 8	.<>A<>1 7 1	.<>Estavam<>1 7 1
.<>Os<>1 7 1	.<>Quatro<>1 7 1	.<>sertanejos<>1 7 1
1896<>,<>1 1 9	:<>a<>1 2 4	:<>ndios<>1 2 1
A<>resist<>1 1 1	Ant<>nio<>1 1 1	Bahia<>,<>1 1 9

Brasil<>:<>1 1 2	Campanha<>de<>1 1 6	Canudos<>.<>1 2 8
Canudos<>tornou<>1 2 1	Conselheiro<>e<>1 1 7	Em<>1896<>1 1 1
Estavam<>poderosamente<> 1 1 1	Ex<>rcito<>1 1 1	Os<>soldados<>1 1 1
Quatro<>expedi<>1 1 1	Rep<>blica<>1 1 1	a<>Campanha<>1 4 1
a<>Rep<>1 4 1	a<>derrota<>1 4 1	a<>hist<>1 4 1
acontecimentos<>mais<>1 1 2	ano<>contra<>1 1 1	aos<>revoltosos<>1 1 1
apenas<>de<>1 1 6	armados<>e<>1 1 7	armas<>r<>1 1 1
as<>batalhas<>1 1 1	assombrosa<>o<>1 1 5	batalhas<>.<>1 1 8
beato<>Ant<>1 1 1	blica<>uma<>1 1 1	caboclos<>,<>1 1 9
canh<>es<>1 1 2	cio<>um<>1 1 2	contra<>mais<>1 1 2
da<>Bahia<>1 2 1	da<>regi<>1 2 1	de<>Canudos<>2 6 2
de<>honra<>1 6 1	de<>paus<>1 6 1	de<>toda<>1 6 1
de<>vinte<>1 6 1	derrota<>de<>1 1 6	dirigidos<>pelo<>1 1 1
do<>Brasil<>1 2 1	do<>sertanejo<>1 2 1	dos<>acontecimentos<>1 1 1
durante<>um<>1 1 2	e<>a<>1 7 4	e<>armas<>1 7 1
e<>canh<>1 7 1	e<>eram<>1 7 1	e<>munidos<>1 7 1
e<>para<>1 7 2	e<>sangrentos<>1 7 1	enviadas<>durante<>1 1 1
eram<>numericamente<>1 1 1	es<>.<>1 2 8	es<>foram<>1 2 1
expedi<>es<>1 1 2	foram<>enviadas<>1 1 1	granadas<>e<>1 1 7
habitantes<>da<>1 1 2	hist<>ria<>1 1 1	honra<>nacional<>1 1 1
impressionantes<>e<>1 1 7	in<>cio<>1 1 1	mais<>de<>1 2 6
mais<>impressionantes<>1 2 1	mas<>perdiam<>1 1 1	metralhadoras<>,<>1 1 9
mil<>habitantes<>1 1 1	muitas<>vezes<>1 1 1	mulatos<>,<>1 1 9

munidos<>apenas<>1 1 1	nacional<>.<>1 1 8	ncia<>do<>1 1 2
ndios<>,<>1 1 9	nio<>Conselheiro<>1 1 1	no<>sert<>1 1 1
numericamente<>muitas<>1 1 1	o<>:<>1 5 2	o<>Ex<>1 5 1
o<>da<>1 5 2	o<>de<>1 5 6	o<>pa<>1 5 1
pa<>s<>1 1 1	para<>a<>1 2 4	para<>o<>1 2 5
paus<>,<>1 1 9	pedras<>e<>1 1 7	pelo<>beato<>1 1 1
perdiam<>todas<>1 1 1	poderosamente<>armados<>1 1 1	pretos<>.<>1 1 8
quest<>o<>1 1 5	r<>sticas<>1 1 1	rcito<>e<>1 1 7
regi<>o<>1 1 5	resist<>ncia<>1 1 1	revoltosos<>,<>1 1 9
ria<>do<>1 1 2	s<>,<>1 1 9	sangrentos<>de<>1 1 6
se<>para<>1 1 2	sert<>o<>1 1 5	sertanejo<>assombrava<>1 1 1
sertanejos<>dirigidos<>1 1 1	soldados<>traziam<>1 1 1	sticas<>.<>1 1 8
superiores<>aos<>1 1 1	teve<>in<>1 1 1	toda<>a<>1 1 4
todas<>as<>1 1 1	tornou<>se<>1 1 1	traziam<>metralhadoras<>1 1 1
um<>ano<>1 2 1	um<>dos<>1 2 1	uma<>quest<>1 1 1
vezes<>superiores<>1 1 1	vinte<>mil<>1 1 1	

**EXEMPLO 2.101 - ORDENAÇÃO ALFABÉTICA DE UM FICHEIRO DE BIGRAMAS**

Considere-se o seguinte ficheiro de bigramas gerado pelo programa 'count.pl' com a opção '-- tokenlist':

test.txt		
Em<>1896<>	1896<>,<>	,<>no<>
no<>sert<>	sert<>o<>	o<>da<>
da<>Bahia<>	Bahia<>,<>	,<>teve<>
teve<>in<>	in<>cio<>	cio<>um<>
um<>dos<>	dos<>acontecimentos<>	acontecimentos<>mais<>
mais<>impressionantes<>	impressionantes<>e<>	e<>sangrentos<>
sangrentos<>de<>	de<>toda<>	toda<>a<>
a<>hist<>	hist<>ria<>	ria<>do<>
do<>Brasil<>	Brasil<>:<>	:<>a<>
a<>Campanha<>	Campanha<>de<>	de<>Canudos<>
Canudos<>.<>	.<>Quatro<>	Quatro<>expedi<>
expedi<>es<>	es<>foram<>	foram<>enviadas<>
enviadas<>durante<>	durante<>um<>	um<>ano<>
ano<>contra<>	contra<>mais<>	mais<>de<>
de<>vinte<>	vinte<>mil<>	mil<>habitantes<>
habitantes<>da<>	da<>regi<>	regi<>o<>
o<>:<>	:<>ndios<>	ndios<>,<>
,<>mulatos<>	mulatos<>,<>	,<>caboclos<>
caboclos<>,<>	,<>pretos<>	pretos<>.<>
.<>.<>	.<>sertanejos<>	sertanejos<>dirigidos<>
dirigidos<>pelo<>	pelo<>beato<>	beato<>Ant<>
Ant<>nio<>	nio<>Conselheiro<>	Conselheiro<>e<>
e<>munidos<>	munidos<>apenas<>	apenas<>de<>

de<>paus<>	paus<>,<>	,<>pedras<>
pedras<>e<>	e<>armas<>	armas<>r<>
r<>sticas<>	sticas<>.<>	.<>Os<>
Os<>soldados<>	soldados<>traziam<>	traziam<>metralhadoras<>
metralhadoras<>,<>	,<>granadas<>	granadas<>e<>
e<>canh<>	canh<>es<>	es<>.<>
.<>Estavam<>	Estavam<>poderosamente<>	poderosamente<>armados<>
armados<>e<>	e<>eram<>	eram<>numericamente<>
numericamente<> muitas<>	muitas<>vezes<>	vezes<>superiores<>
superiores<>aos<>	aos<>revoltosos<>	revoltosos<>,<>
,<>mas<>	mas<>perdiam<>	perdiam<>todas<>
todas<>as<>	as<>batalhas<>	batalhas<>.<>
.<>A<>	A<>resist<>	resist<>ncia<>
ncia<>do<>	do<>sertanejo<>	sertanejo<>assombrava<>
assombrava<>o<>	o<>pa<>	pa<>s<>
s<>,<>	,<>e<>	e<>a<>
a<>derrota<>	derrota<>de<>	de<>Canudos<>
Canudos<>tornou<>	tornou<>se<>	se<>para<>
para<>o<>	o<>Ex<>	Ex<>rcito<>
rcito<>e<>	e<>para<>	para<>a<>
a<>Rep<>	Rep<>blica<>	blica<>uma<>
uma<>quest<>	quest<>o<>	o<>de<>
de<>honra<>	honra<>nacional<>	nacional<>.<>

Após a execução do comando '*huge-sort.pl test.txt*' obtivemos o seguinte resultado:

test.txt-sorted		
,<>caboclos<>1 9 1	,<>e<>1 9 7	,<>granadas<>1 9 1
,<>mas<>1 9 1	,<>mulatos<>1 9 1	,<>no<>1 9 1
,<>pedras<>1 9 1	,<>pretos<>1 9 1	,<>teve<>1 9 1
.<>.<>2 7 8	.<>A<>1 7 1	.<>Estavam<>1 7 1
.<>Os<>1 7 1	.<>Quatro<>1 7 1	.<>sertanejos<>1 7 1
1896<>, <>1 1 9	:<>a<>1 2 4	:<>ndios<>1 2 1
A<>resist<>1 1 1	Ant<>nio<>1 1 1	Bahia<>, <>1 1 9
Brasil<>:<>1 1 2	Campanha<>de<>1 1 6	Canudos<>.<>1 2 8
Canudos<>tornou<>1 2 1	Conselheiro<>e<>1 1 7	Em<>1896<>1 1 1
Estavam<>poderosamente<> 1 1 1	Ex<>rcito<>1 1 1	Os<>soldados<>1 1 1
Quatro<>expedi<>1 1 1	Rep<>blica<>1 1 1	a<>Campanha<>1 4 1
a<>Rep<>1 4 1	a<>derrota<>1 4 1	a<>hist<>1 4 1
acontecimentos<>mais<>1 1 2	ano<>contra<>1 1 1	aos<>revoltosos<>1 1 1
apenas<>de<>1 1 6	armados<>e<>1 1 7	armas<>r<>1 1 1
as<>batalhas<>1 1 1	assombrava<>o<>1 1 5	batalhas<>.<>1 1 8
beato<>Ant<>1 1 1	blica<>uma<>1 1 1	caboclos<>, <>1 1 9
canh<>es<>1 1 2	cio<>um<>1 1 2	contra<>mais<>1 1 2
da<>Bahia<>1 2 1	da<>regi<>1 2 1	de<>Canudos<>2 6 2
de<>honra<>1 6 1	de<>paus<>1 6 1	de<>toda<>1 6 1
de<>vinte<>1 6 1	derrota<>de<>1 1 6	dirigidos<>pelo<>1 1 1
do<>Brasil<>1 2 1	do<>sertanejo<>1 2 1	dos<>acontecimentos<>1 1 1
durante<>um<>1 1 2	e<>a<>1 7 4	e<>armas<>1 7 1
e<>canh<>1 7 1	e<>eram<>1 7 1	e<>munidos<>1 7 1



eram<>numericamente<>1 1 1	es<>.<>1 2 8	es<>foram<>1 2 1
expedi<>es<>1 1 2	foram<>enviadas<>1 1 1	granadas<>e<>1 1 7
habitantes<>da<>1 1 2	hist<>ria<>1 1 1	honra<>nacional<>1 1 1
impressionantes<>e<>1 1 7	in<>cio<>1 1 1	mais<>de<>1 2 6
mais<>impressionantes<>1 2 1	mas<>perdiam<>1 1 1	metralhadoras<>,<>1 1 9
mil<>habitantes<>1 1 1	muitas<>vezes<>1 1 1	mulatos<>,<>1 1 9
unidos<>apenas<>1 1 1	nacional<>.<>1 1 8	ncia<>do<>1 1 2
ndios<>,<>1 1 9	nio<>Conselheiro<>1 1 1	no<>sert<>1 1 1
numericamente<>muitas<>1 1 1	o<>:<>1 5 2	o<>Ex<>1 5 1
o<>da<>1 5 2	o<>de<>1 5 6	o<>pa<>1 5 1
pa<>s<>1 1 1	para<>a<>1 2 4	para<>o<>1 2 5
paus<>,<>1 1 9	pedras<>e<>1 1 7	pelo<>beato<>1 1 1
perdiam<>todas<>1 1 1	poderosamente<>armados<> 1 1 1	pretos<>.<>1 1 8
quest<>o<>1 1 5	r<>sticas<>1 1 1	rcito<>e<>1 1 7
regi<>o<>1 1 5	resist<>ncia<>1 1 1	revoltosos<>,<>1 1 9
ria<>do<>1 1 2	s<>,<>1 1 9	sangrentos<>de<>1 1 6
se<>para<>1 1 2	sert<>o<>1 1 5	sertanejo<>assombrava<>1 1 1
sertanejos<>dirigidos<>1 1 1	soldados<>traziam<>1 1 1	sticas<>.<>1 1 8
superiores<>aos<>1 1 1	teve<>in<>1 1 1	toda<>a<>1 1 4
todas<>as<>1 1 1	tornou<>se<>1 1 1	traziam<>metralhadoras<> 1 1 1
um<>ano<>1 2 1	um<>dos<>1 2 1	uma<>quest<>1 1 1
vezes<>superiores<>1 1 1	vinte<>mil<>1 1 1	

## EXEMPLO 2.102 - PARTIÇÃO DE UM FICHEIRO DE BIGRAMAS EM FICHEIROS MAIS PEQUENOS

Suponhamos que temos o seguinte ficheiro de bigramas gerado pelo programa 'count.pl':

test.txt		
Em<>1896<>	1896<>,<>	,<>no<>
no<>sert<>	sert<>o<>	o<>da<>
da<>Bahia<>	Bahia<>,<>	,<>teve<>
teve<>in<>	in<>cio<>	cio<>um<>
um<>dos<>	dos<>acontecimentos<>	acontecimentos<>mais<>
mais<>impressionantes<>	impressionantes<>e<>	e<>sangrentos<>
sangrentos<>de<>	de<>toda<>	toda<>a<>
a<>hist<>	hist<>ria<>	ria<>do<>
do<>Brasil<>	Brasil<>:<>	:<>a<>
a<>Campanha<>	Campanha<>de<>	de<>Canudos<>
Canudos<>.<>	.<>Quatro<>	Quatro<>expedi<>
expedi<>es<>	es<>foram<>	foram<>enviadas<>
enviadas<>durante<>	durante<>um<>	um<>ano<>
ano<>contra<>	contra<>mais<>	mais<>de<>
de<>vinte<>	vinte<>mil<>	mil<>habitantes<>
habitantes<>da<>	da<>regi<>	regi<>o<>
o<>:<>	:<>ndios<>	ndios<>,<>
,<>mulatos<>	mulatos<>,<>	,<>caboclos<>
caboclos<>,<>	,<>pretos<>	pretos<>.<>
.<>.<>	.<>sertanejos<>	sertanejos<>dirigidos<>
dirigidos<>pelo<>	pelo<>beato<>	beato<>Ant<>
Ant<>nio<>	nio<>Conselheiro<>	Conselheiro<>e<>

e<>munidos<>	munidos<>apenas<>	apenas<>de<>
de<>paus<>	paus<>,<>	,<>pedras<>
pedras<>e<>	e<>armas<>	armas<>r<>
r<>sticas<>	sticas<>.<>	.<>Os<>
Os<>soldados<>	soldados<>traziam<>	traziam<>metralhadoras<>
metralhadoras<>,<>	,<>granadas<>	granadas<>e<>
e<>canh<>	canh<>es<>	es<>.<>
.<>Estavam<>		

Se quisermos partir este ficheiro em ficheiros mais pequenos, contendo cada um deles trinta bigramas, utilizando o comando '*huge-split.pl -split 30 test.txt*', obtemos o seguinte resultado:

test.txt.1
Em<>1896<>
1896<>,<>
,<>no<>
no<>sert<>
sert<>o<>
o<>da<>
da<>Bahia<>
Bahia<>,<>
,<>teve<>
teve<>in<>
in<>cio<>
cio<>um<>
um<>dos<>
dos<>acontecimentos<>
acontecimentos<>mais<>
mais<>impressionantes<>
impressionantes<>e<>
e<>sangrentos<>
sangrentos<>de<>

de<>toda<>  
toda<>a<>  
a<>hist<>  
hist<>ria<>  
ria<>do<>  
do<>Brasil<>  
Brasil<>:<>  
<>a<>  
a<>Campanha<>  
Campanha<>de<>  
de<>Canudos<>

test.txt.2

Canudos<>.<>  
.<>Quatro<>  
Quatro<>expedi<>  
expedi<>es<>  
es<>foram<>  
foram<>enviadas<>  
enviadas<>durante<>  
durante<>um<>  
um<>ano<>  
ano<>contra<>  
contra<>mais<>  
mais<>de<>  
de<>vinte<>  
vinte<>mil<>  
mil<>habitantes<>  
habitantes<>da<>  
da<>regi<>  
regi<>o<>  
o<>:<>  
<>ndios<>

ndios<>,<>  
,<>mulatos<>  
mulatos<>,<>  
,<>caboclos<>  
caboclos<>,<>  
,<>pretos<>  
pretos<>.<>  
.<>.<>  
.<>.<>  
.<>sertanejos<>

test.txt.3

sertanejos<>dirigidos<>  
dirigidos<>pelo<>  
pelo<>beato<>  
beato<>Ant<>  
Ant<>nio<>  
nio<>Conselheiro<>  
Conselheiro<>e<>  
e<>munidos<>  
munidos<>apenas<>  
apenas<>de<>  
de<>paus<>  
paus<>,<>  
,<>pedras<>  
pedras<>e<>  
e<>armas<>  
armas<>r<>  
r<>sticas<>  
sticas<>.<>  
.<>Os<>  
Os<>soldados<>  
soldados<>traziam<>

traziam<>metralhadoras<>  
metralhadoras<>,<>  
,<>granadas<>  
granadas<>e<>  
e<>canh<>  
canh<>es<>  
es<>.<>  
.<>Estavam<>  
Estavam<>poderosamente<>

test.txt.4

poderosamente<>armados<>  
armados<>e<>  
e<>eram<>  
eram<>numericamente<>  
numericamente<>muitas<>  
muitas<>vezes<>  
vezes<>superiores<>  
superiores<>aos<>  
aos<>revoltosos<>  
revoltosos<>,<>  
,<>mas<>  
mas<>perdiam<>  
perdiam<>todas<>  
todas<>as<>  
as<>batalhas<>  
batalhas<>.<>  
.<>A<>

**EXEMPLO 2.103 - LISTA DAS CO-OCORRÊNCIAS DE UM FICHEIRO DE BIGRAMAS**

Suponhamos novamente o ficheiro de *input* do exemplo 2.102, test.txt.

Se executarmos o comando `'kocos.pl --order 5 --literal 'acontecimentos' output'`, obtemos o seguinte resultado:

teve<>	foram<>	nio<>
r<>	es<>	hist<>
:<>	Quatro<>	pretos<>
sertanejos<>	sticas<>	Os<>
Estavam<>	habitantes<>	1896<>
no<>	Bahia<>	ndios<>
mulatos<>	caboclos<>	metralhadoras<>

**EXEMPLO 2.104 - CÁLCULO DO COEFICIENTE DE CORRELAÇÃO RELATIVAMENTE A DUAS ESTATÍSTICAS DISTINTAS PARA O MESMO CONJUNTO DE DADOS.**

Suponhamos que temos o seguinte ficheiro gerado pelo programa `'count.pl'`:

teste.txt		
1060	.<>.<>16 78 79	Ant<>nio<>8 8 8
n<>o<>7 7 56	nio<>Conselheiro<>6 8 9	,<>que<>6 58 19
sert<>o<>5 5 56	.<>Ele<>5 78 6	.<>O<>5 78 6
o<>de<>4 56 39	minha<>m<>3 4 4	Jo<>o<>3 3 56
nascido<>em<>3 3 10	de<>um<>3 39 11	,<>em<>3 58 10
Rep<>blica<>3 3 6	m<>e<>3 4 28	dos<>Santos<>3 7 4
.<>A<>3 78 4	o<>governo<>3 56 4	rep<>blica<>3 3 6

de<>Canudos<>3 39 5	a<>Rep<>3 24 3	,<>o<>3 58 56
e<>o<>3 28 56	Santos<>,<>3 4 58	da<>Bahia<>2 14 2
dele<>,<>2 2 58	para<>o<>2 9 56	rios<>.<>2 2 79
de<>uma<>2 39 6	tinha<>uma<>2 4 6	era<>com<>2 9 7
naquele<>tempo<>2 2 3	livro<>Os<>2 2 3	queria<>a<>2 2 24
cio<>Jos<>2 6 2	da<>monarquia<>2 14 4	por<>ali<>2 2 2
o<>sert<>2 56 5	muitos<>milagres<>2 2 2	Euclides<>da<>2 2 14
mudan<>as<>2 2 8	Minha<>av<>2 2 4	de<>guerra<>2 39 3
Conselheiro<>.<>2 9 79	blica<>.<>2 6 79	.<>Mas<>2 78 2
.<>Para<>2 78 2	,<>era<>2 58 9	Sert<>es<>2 2 6
o<>.<>2 56 79	ele<>era<>2 3 9	.<>Quando<>2 78 2
que<>naquele<>2 19 2	que<>ele<>2 19 3	Ele<>n<>2 6 7
foi<>quem<>2 5 2	de<>Santo<>2 39 3	.<>Uma<>2 78 2
Conselheiro<>chegou<>2 9 2	o<>golpe<>2 56 2	para<>a<>2 9 24
.<>Minha<>2 78 2	dava<>rem<>2 2 2	da<>Cunha<>2 14 2
o<>acreditava<>2 56 2	Ele<>dava<>2 6 2	.<>Foi<>2 78 2
blica<>,<>2 6 58	Jos<>dos<>2 2 7	,<>mas<>2 58 2
.<>No<>2 78 2	a<>rep<>2 24 3	chegou<>aqui<>2 2 3
monarquia<>.<>2 4 79	,<>Ant<>2 58 8	que<>n<>2 19 7
,<>no<>2 58 2	.<>Jo<>2 78 3	Os<>Sert<>2 3 2
Jesus<>.<>2 2 79	in<>cio<>2 2 6	O<>Conselheiro<>2 6 9
rem<>dio<>2 2 2	o<>:<>2 56 4	de<>Jesus<>2 39 2
nacional<>.<>2 2 79	,<>e<>2 58 28	Canudos<>.<>2 5 79
Vital<>cio<>2 2 6	es<>,<>2 6 58	Santo<>Ant<>2 3 8
mas<>era<>1 2 9	voc<>s<>1 1 5	resist<>ncia<>1 1 1
tiravam<>hinos<>1 1 1	vou<>!<>1 1 7	de<>parto<>1 39 1



e<>come<>1 28 1	sangrentos<>de<>1 1 39	com<>as<>1 7 8
sem<>dono<>1 1 1	rejei<>o<>1 1 56	democr<>ticas<>1 1 1
final<>do<>1 1 9	obra<>contundente<>1 1 1	l<>.<>1 2 79
.<>Disseram<>1 78 1	que<>pareciam<>1 19 1	acompanhando<>esse<>1 1 1
As<>elei<>1 1 1	nascida<>em<>1 1 10	.<>Se<>1 78 2
representou<>o<>1 1 56	,<>Joana<>1 58 1	o<>branca<>1 56 1
em<>1909<>1 10 1	se<>por<>1 12 2	es<>nasceu<>1 6 1
e<>da<>1 28 14	apenas<>de<>1 1 39	cio<>de<>1 6 39
a<>revolta<>1 24 1	cobrir<>o<>1 1 56	aux<>lio<>1 1 1
Josefa<>dos<>1 2 7	de<>madrugada<>1 39 1	1897<>,<>1 1 58
sonho<>brasileiro<>1 1 1	dio<>e<>1 2 28	Cantem<>meus<>1 1 1
dia<>,<>1 2 58	militar<>Euclides<>1 2 2	pretos<>.<>1 1 79
essa<>a<>1 1 24	,<>bom<>1 58 2	1888<>fora<>1 1 1
e<>para<>1 28 9	que<>fosse<>1 19 1	crescer<>a<>1 1 24
quem<>venceu<>1 2 1	monarquia<>para<>1 4 9	Se<>tinha<>1 2 4
o<>At<>1 56 1	de<>paus<>1 39 1	engenheiro<>militar<>1 1 2
nunca<>de<>1 1 39	nio<>,<>1 8 58	para<>vir<>1 9 1
na<>monarquia<>1 5 4	av<>e<>1 4 28	mais<>impressionantes<>1 2 1
,<>Jo<>1 58 3	em<>1897<>1 10 1	perante<>a<>1 1 24
achou<>se<>1 1 12	es<>foram<>1 6 2	dos<>escravos<>1 7 1
mil<>habitantes<>1 1 1	v<>o<>1 1 56	enviado<>de<>1 1 39
Deus<>?<>1 3 1	Bittencourt<>,<>1 1 58	se<>esperava<>1 12 1
Professora<>nascido<>1 1 3	as<>outras<>1 8 1	casa<>de<>1 1 39
o<>tempo<>1 56 3	e<>munidos<>1 28 1	tria<>sem<>1 1 1
liderado<>pelo<>1 1 2	a<>nova<>1 24 2	do<>Rio<>1 9 1

as<>A<>1 8 4	era<>morena<>1 9 1	Maria<>!<>1 2 7
S<>o<>1 1 56	Quatro<>expedi<>1 1 1	ao<>redor<>1 5 1
o<>que<>1 56 19	de<>desvalidos<>1 39 1	um<>grupo<>1 11 1
mulheres<>,<>1 2 58	E<>foi<>1 2 5	entanto<>,<>1 1 58
achando<>que<>1 1 19	perdiam<>todas<>1 1 1	porque<>como<>1 1 2
imprevisto<>.<>1 1 79	profeta<>de<>1 1 39	Ave<>Maria<>1 1 2
rua<>,<>1 1 58	dariam<>a<>1 1 24	Conselheiro<>mesmo<>1 9 2
todo<>mundo<>1 1 1	meus<>filhos<>1 1 1	levante<>monarquista<>1 1 1
com<>carregamento<>1 7 1	aumentando<>e<>1 1 28	Uma<>obra<>1 2 1
era<>um<>1 9 11	e<>que<>1 28 19	Sei<>at<>1 1 2
esse<>povo<>1 1 3	viesse<>tona<>1 1 1	o<>primeiro<>1 56 1
Mas<>nunca<>1 2 1	nesse<>milagroso<>1 1 1	uma<>reportagem<>1 6 1
brasileiro<>da<>1 1 14	foi<>nessa<>1 5 1	ao<>lado<>1 5 1
bonita<>.<>1 2 79	incutindo<>com<>1 1 7	Brasil<>:<>1 1 4
na<>boca<>1 5 1	Papai<>casou<>1 1 2	.<>Meu<>1 78 1
Quando<>teve<>1 2 3	um<>Deus<>1 11 3	Mattos<>,<>1 1 58
J<>se<>1 1 12	.<>Elas<>1 78 1	E<>minha<>1 2 4
Tinha<>muita<>1 1 2	muito<>de<>1 2 39	tal<>de<>1 1 39
contra<>o<>1 3 56	e<>tinha<>1 28 4	se<>l<>1 12 2
aquele<>que<>1 2 19	,<>romeiro<>1 58 1	de<>Marechal<>1 39 1
mulatos<>,<>1 1 58	habitantes<>da<>1 1 14	cantavam<>.<>1 1 79
Janeiro<>,<>1 1 58	de<>escolher<>1 39 1	e<>sangrentos<>1 28 1
ela<>respondia<>1 1 1	bom<>pra<>1 2 1	expedi<>es<>1 1 6
,<>ao<>1 58 5	noites<>!<>1 1 7	chegar<>ao<>1 1 5
rio<>que<>1 1 19	enviadas<>durante<>1 1 1	fazendo<>muitos<>1 1 2
se<>rezasse<>1 12 1	:<>a<>1 4 24	vai<>dizer<>1 1 1

militar<>do<>1 2 9	dessa<>rejei<>1 1 1	o<>jornal<>1 56 1
chamava<>Josefa<>1 1 2	Campanha<>de<>1 1 39	se<>para<>1 12 9
uma<>estranha<>1 6 1	eram<>numericamente<>1 3 1	s<>,<>1 5 58
,<>foram<>1 58 2	golpe<>fatal<>1 2 1	.<>J<>1 78 1
milagres<>aqui<>1 2 3	as<>gra<>1 8 1	Cunha<>foi<>1 2 5
se<>despachavam<>1 12 1	ticas<>dariam<>1 1 1	ano<>contra<>1 2 3
as<>13<>1 8 1	beato<>Ant<>1 1 8	boca<>da<>1 1 14
paus<>,<>1 1 58	avisaram<>aquela<>1 1 2	mandou<>o<>1 1 56
abandonada<>pelas<>1 1 1	Deodoro<>da<>1 1 14	toda<>a<>1 1 24
o<>povo<>1 56 3	governo<>,<>1 4 58	gra<>as<>1 1 8
Estado<>de<>1 1 39	1896<>,<>1 1 58	de<>S<>1 39 1
em<>Canudos<>1 10 5	despachavam<>.<>1 1 79	o<>imprevisto<>1 56 1
.<>Estavam<>1 78 1	durante<>um<>1 1 11	todas<>as<>1 1 8
dos<>acontecimentos<>1 7 1	em<>nome<>1 10 1	O<>Estado<>1 6 1
A<>o<>1 4 56	Rosada<>,<>1 1 58	de<>homens<>1 39 2
No<>ano<>1 2 2	liberta<>o<>1 1 56	nasceu<>de<>1 1 39
poca<>de<>1 1 39	foram<>enviadas<>1 2 1	de<>Joaquim<>1 39 1
Santos<>.<>1 4 79	de<>pedras<>1 39 2	noite<>e<>1 1 28
sul<>ignoravam<>1 1 1	ali<>n<>1 2 7	.<>Os<>1 78 3
1911<>.<>1 1 79	na<>Fazenda<>1 5 1	.<>As<>1 78 1
como<>enviado<>1 2 1	acreditava<>nesse<>1 2 1	leis<>e<>1 1 28
.<>Papai<>1 78 1	aos<>revoltosos<>1 1 1	.<>Continuaram<>1 78 1
mesmo<>.<>1 2 79	aqui<>fazendo<>1 3 1	Maria<>de<>1 2 39
da<>Professora<>1 14 1	seguinte<>,<>1 1 58	a<>Guerra<>1 24 1
e<>armas<>1 28 1	Rio<>de<>1 1 39	todos<>cantavam<>1 2 1
a<>ficar<>1 24 1	Perguntou<>se<>1 1 12	que<>mandou<>1 19 1

tudo<>isso<>1 1 1	igreja<>de<>1 1 39	novo<>regime<>1 1 1
branca<>europeizada<>1 1 1	aquela<>vozona<>1 1 1	Macambira<>.<>1 1 79
Bahia<>,<>1 2 58	governo<>.<>1 4 79	Veio<>com<>1 1 7
que<>minha<>1 19 4	eram<>muito<>1 3 2	reportagem<>sobre<>1 1 1
o<>queria<>1 56 2	em<>1911<>1 10 1	gios<>.<>1 1 79
com<>madeiras<>1 7 1	o<>pa<>1 56 1	estranha<>p<>1 1 1
e<>do<>1 28 9	come<>aram<>1 1 1	Reginaldo<>de<>1 1 39
a<>morte<>1 24 1	com<>antigos<>1 7 1	Em<>1896<>1 1 1
culo<>XIX<>1 1 1	ncia<>do<>1 1 9	aram<>a<>1 1 24
,<>mesmo<>1 58 2	dinheiro<>.<>1 1 79	lutavam<>at<>1 1 2
com<>a<>1 7 24	dirigidos<>pelo<>1 1 2	,<>foi<>1 58 5
Joaquim<>Macambira<>1 1 1	fora<>o<>1 1 56	um<>povo<>1 11 3
muito<>bem<>1 2 1	Guerra<>de<>1 1 39	a<>promessa<>1 24 1
desvalidos<>.<>1 1 79	armas<>r<>1 1 1	que<>se<>1 19 12
ndios<>,<>1 1 58	aqui<>na<>1 3 5	.<>Eram<>1 78 1
Foi<>essa<>1 2 1	isso<>.<>1 1 79	contra<>a<>1 3 24
grupo<>de<>1 1 39	jornal<>O<>1 1 6	,<>granadas<>1 58 1
homens<>lutavam<>1 2 1	Ela<>era<>1 1 9	a<>todos<>1 24 2
at<>a<>1 2 24	reis<>da<>1 1 14	de<>pedir<>1 39 1
fez<>muitos<>1 1 2	.<>Vital<>1 78 2	radicais<>,<>1 1 58
,<>na<>1 58 5	pleno<>sert<>1 1 5	ordem<>.<>1 1 79
13<>noites<>1 1 1	que<>o<>1 19 56	pelas<>leis<>1 1 1
hist<>ria<>1 1 1	governo<>da<>1 4 14	dio<>para<>1 2 9
O<>livro<>1 6 2	destru<>a<>1 1 24	rcito<>de<>1 2 39
de<>mudan<>1 39 2	era<>bonita<>1 9 2	filho<>,<>1 1 58

acabar<>com<>1 1 7	loi<>da<>1 1 14	numericamente<>muitas<>1 1 1
Marechal<>Bittencourt<>1 1 1	Canudos<>para<>1 5 9	,<>numa<>1 5 8 1
,<>n<>1 5 8 7	tardinha<>,<>1 1 5 8	numa<>nova<>1 1 2
a<>lei<>1 2 4 1	o<>Siqueira<>1 5 6 1	guerra<>,<>1 3 5 8
blica<>uma<>1 6 6	e<>engenheiro<>1 2 8 1	,<>um<>1 5 8 11
Ele<>fez<>1 6 1	A<>passagem<>1 4 1	revoltou<>contra<>1 1 3
Ele<>fazia<>1 6 1	de<>Monte<>1 3 9 1	,<>ele<>1 5 8 3
trabalhar<>naquelas<>1 1 1 1	a<>rua<>1 2 4 1	m<>minha<>1 4 4
foi<>cobrir<>1 5 1	nome<>dessa<>1 1 1	as<>dele<>1 8 2
de<>toda<>1 3 9 1	escravos<>em<>1 1 1 10	R<>gis<>1 1 1
armados<>e<>1 1 2 8	canh<>o<>1 2 5 6	.<>Perguntou<>1 7 8 1
monarquista<>.<>1 1 7 9	impressionantes<>e<>1 1 2 8	pareciam<>acabar<>1 1 1 1
feij<>o<>1 1 5 6	at<>de<>1 2 3 9	se<>revoltou<>1 1 2 1
ali<>.<>1 2 7 9	dos<>reis<>1 7 1	para<>mulheres<>1 9 2
seus<>dirigentes<>1 1 1 1	do<>litoral<>1 9 1	se<>casou<>1 1 2 2
o<>na<>1 5 6 5	O<>novo<>1 6 1	ano<>seguinte<>1 2 1
.<>Ant<>1 7 8 8	Essa<>,<>1 1 5 8	Monte<>Santo<>1 1 3
ao<>publicar<>1 5 1	elites<>do<>1 1 9	de<>Mattos<>1 3 9 1
morte<>em<>1 1 1 10	.<>Canudos<>1 7 8 5	que<>foi<>1 1 9 5
que<>eram<>1 1 9 3	of<>cio<>1 1 6	,<>teve<>1 5 8 3
nio<>.<>1 8 7 9	respondia<>.<>1 1 7 9	.<>Tinha<>1 7 8 1
,<>vivendo<>1 5 8 1	o<>of<>1 5 6 1	cheia<>.<>1 1 7 9
livre<>traria<>1 1 1 1	tornou<>se<>1 1 1 12	do<>jornalista<>1 9 1
se<>com<>1 1 2 7	.<>Quatro<>1 7 8 1	tinha<>,<>1 4 5 8
capelinha<>onde<>1 1 1	neg<>cio<>1 1 6	,<>abandonada<>1 5 8 1

tempo<>ainda<>1 3 1	madrugada<>.<>1 1 79	cio<>da<>1 6 14
era<>mo<>1 9 1	com<>o<>1 7 56	organiza<>o<>1 1 56
e<>se<>1 28 12	europalizada<>.<>1 1 79	.<>era<>1 78 9
civiliza<>o<>1 1 56	nova<>ordem<>1 2 1	ele<>que<>1 3 19
.<>poca<>1 78 1	para<>fazer<>1 9 2	privil<>gios<>1 1 1
muita<>agita<>1 2 1	,<>loi<>1 58 1	tamb<>m<>1 1 4
nisso<>porque<>1 1 1	s<>culo<>1 5 1	filhos<>!<>1 1 7
povo<>.<>1 3 79	Estavam<>poderosamente<>1 1 1	muita<>gente<>1 2 2
proeza<>do<>1 1 9	ir<>,<>1 1 58	gente<>para<>1 2 9
metralhadoras<>,<>1 1 58	para<>trabalhar<>1 9 1	.<>Veio<>1 78 1
o<>ter<>1 56 1	sertanejos<>dirigidos<>1 1 1	verdadeiro<>profeta<>1 1 1
!<>Eu<>1 7 1	Foi<>quando<>1 2 1	escolher<>seus<>1 1 1
e<>ela<>1 28 1	foram<>aumentando<>1 2 1	?<>Essa<>1 1 1
com<>um<>1 7 11	cio<>,<>1 6 58	dia<>15<>1 2 1
soldados<>traziam<>1 1 1	av<>.<>1 4 79	,<>queria<>1 58 2
bom<>para<>1 2 9	sob<>o<>1 1 56	Cantava<>na<>1 1 5
Uma<>cabocla<>1 2 1	sou<>aquela<>1 1 2	gente<>,<>1 2 58
,<>achando<>1 58 1	e<>institui<>1 28 1	que<>um<>1 19 11
lei<>.<>1 1 79	uma<>organiza<>1 6 1	o<>sonho<>1 56 1
um<>ano<>1 11 2	de<>muita<>1 39 2	teve<>de<>1 3 39
o<>s<>1 56 5	as<>radicais<>1 8 1	antigos<>privil<>1 1 1
minha<>av<>1 4 4	um<>tal<>1 11 1	aqui<>,<>1 3 58
reza<>.<>1 1 79	peçoal<>ao<>1 1 5	institui<>es<>1 1 6
contudente<>,<>1 1 58	ficar<>por<>1 1 2	Mas<>deixa<>1 2 1
novembro<>,<>1 1 58	madeiras<>.<>1 1 79	obras<>,<>1 1 58
que<>destru<>1 19 1	A<>liberta<>1 4 1	de<>Deus<>1 39 3

e<>fazer<>1 28 2	a<>Campanha<>1 24 1	traria<>sal<>1 1 1
tempo<>gostava<>1 3 1	Bahia<>teve<>1 2 3	No<>entanto<>1 2 1
a<>igreja<>1 24 1	Quando<>achou<>1 2 1	desconhecido<>sert<>1 1 5
e<>minha<>1 28 4	necess<>rio<>1 1 1	XIX<>, <>1 1 58
!<>Homens<>1 7 1	gostava<>muito<>1 1 2	do<>nordeste<>1 9 2
o<>da<>1 56 14	sticas<>.<>1 1 79	pelo<>beato<>1 2 1
ainda<>era<>1 1 9	caso<>de<>1 1 39	as<>batalhas<>1 8 1
governo<>do<>1 4 9	no<>final<>1 2 1	o<>direito<>1 56 1
e<>dos<>1 28 7	se<>chamava<>1 12 1	ignoravam<>o<>1 1 56
!<>E<>1 7 2	cabocla<>bonita<>1 1 2	a<>monarquia<>1 24 4
, <>D<>1 58 1	Os<>soldados<>1 3 1	o<>nacional<>1 56 2
da<>civiliza<>1 14 1	s<>!<>1 5 7	teve<>o<>1 3 56
Eu<>sou<>1 1 1	Zefinha<>nascida<>1 1 1	milagres<>, <>1 2 58
, <>proclamou<>1 58 1	gis<>nascido<>1 1 3	vezes<>superiores<>1 1 1
tempo<>dos<>1 3 7	nova<>tradu<>1 2 1	s<>poderia<>1 5 1
acreditava<>nisso<>1 2 1	o<>canh<>1 56 2	vir<>e<>1 1 28
de<>R<>1 39 1	ter<>o<>1 1 56	r<>sticas<>1 1 1
elei<>es<>1 1 6	da<>guerra<>1 14 3	em<>1902<>1 10 1
O<>desconhecido<>1 6 1	eram<>de<>1 3 39	e<>, <>1 28 58
ria<>do<>1 1 9	blica<>e<>1 6 28	trabalho<>livre<>1 1 1
.<>Josefa<>1 78 2	data<>que<>1 1 19	evento<>, <>1 1 58
Joana<>Batista<>1 1 1	promessa<>de<>1 1 39	, <>todos<>1 58 2
mesmo<>pagava<>1 2 1	era<>bom<>1 9 2	vinte<>mil<>1 1 1
1936<>.<>1 1 79	sal<>rios<>1 1 2	.<>Euclides<>1 78 2
era<>revoltado<>1 9 1	milagre<>.<>1 1 79	pra<>ganhar<>1 1 1
.<>sertanejos<>1 78 1	do<>s<>1 9 5	mas<>perdiam<>1 2 1
naquelas<>obras<>1 1 1	de<>Janeiro<>1 39 1	Se<>voc<>1 2 1

jornalista<>e<>1 1 28	poderia<>se<>1 1 12	hinos<>e<>1 1 28
revoltado<>contra<>1 1 3	se<>beneficiar<>1 12 1	do<>sul<>1 9 1
era<>necess<>1 9 1	agita<>o<>1 1 56	a<>o<>1 24 56
casou<>.<>1 2 79	pedir<>aux<>1 1 1	pedras<>,<>1 2 58
revoltosos<>,<>1 1 58	beneficiar<>com<>1 1 7	pedras<>e<>1 2 28
feitas<>.<>1 1 79	muitas<>vezes<>1 1 1	da<>rep<>1 14 3
revolta<>era<>1 1 9	povo<>que<>1 3 19	dirigentes<>,<>1 1 58
,<>caboclos<>1 58 1	salvava<>.<>1 1 79	mo<>a<>1 1 24
,<>pretos<>1 58 1	!<>Se<>1 7 2	sertanejo<>assombrava<>1 1 1
quest<>o<>1 1 56	o<>in<>1 56 2	para<>chegar<>1 9 1
Mam<>e<>1 1 28	lado<>dele<>1 1 2	1902<>.<>1 1 79
convidaram<>para<>1 1 9	monarquia<>,<>1 4 58	a<>e<>1 24 28
bonita<>,<>1 2 58	Deus<>.<>1 3 79	primeiro<>filho<>1 1 1
dono<>,<>1 1 58	.<>N<>1 78 1	,<>elas<>1 58 1
.<>Mam<>1 78 1	fazer<>a<>1 2 24	compreender<>a<>1 1 24
do<>Brasil<>1 9 1	o<>convidaram<>1 56 1	proclamou<>a<>1 1 24
elas<>vinham<>1 1 1	em<>pleno<>1 10 1	Homens<>e<>1 1 28
o<>vai<>1 56 1	1909<>.<>1 1 79	,<>a<>1 58 24
de<> <>1 39 2	.<>tudo<>1 78 1	es<>.<>1 6 79
que<>tinha<>1 19 4	.<>Cantava<>1 78 1	!<>Vital<>1 7 2
o<>jugo<>1 56 1	uma<>trezena<>1 6 1	pol<>tico<>1 1 1
fazer<>neg<>1 2 1	aquele<>pessoal<>1 2 1	o<>Reginaldo<>1 56 1
jugo<>da<>1 1 14	ganhar<>dinheiro<>1 1 1	da<>Fonseca<>1 14 1
Conselheiro<>!<>1 9 7	mundo<>se<>1 1 12	,<>todo<>1 58 1
passagem<>da<>1 1 14	da<>noite<>1 14 1	dos<>latifundi<>1 7 1
vozona<>cheia<>1 1 1	bem<>feitas<>1 1 1	honra<>nacional<>1 1 2



:<>uma<>1 4 6	,<>mulatos<>1 58 1	per<>odo<>1 1 1
quem<>batizou<>1 2 1	o<>Ex<>1 56 1	o<>nordeste<>1 56 2
grande<>proeza<>1 1 1	iguais<>perante<>1 1 1	Para<>o<>1 2 56
o<>,<>1 56 58	batizou<>.<>1 1 79	es<>democr<>1 6 1
o<>viesse<>1 56 1	fazia<>milagre<>1 1 1	.<>Sei<>1 78 1
do<>sertanejo<>1 9 1	trazia<>a<>1 1 24	s<>quiserem<>1 5 1
o<>trabalho<>1 56 1	na<>segunda<>1 5 1	l<>em<>1 2 10
av<>se<>1 4 12	povo<>do<>1 3 9	regi<>o<>1 1 56
sobre<>a<>1 1 24	pelo<>marechal<>1 2 1	e<>mulheres<>1 28 2
Conselheiro<>dizia<>1 9 1	um<>comedor<>1 11 1	munidos<>apenas<>1 1 1
regime<>trazia<>1 1 1	Para<>compreender<>1 2 1	granadas<>e<>1 1 28
parto<>.<>1 1 79	terra<>e<>1 1 28	At<>o<>1 1 56
latifundi<>rios<>1 1 2	,<>homens<>1 58 2	trezena<>de<>1 1 39
publicar<>seu<>1 1 1	fatal<>na<>1 1 5	de<>honra<>1 39 1
na<>casa<>1 5 1	ex<>rcito<>1 1 2	um<>dos<>1 11 7
nordeste<>s<>1 2 5	em<>1919<>1 10 1	Fonseca<>,<>1 1 58
morena<>,<>1 1 58	.<>Ave<>1 78 1	,<>como<>1 58 2
em<>1888<>1 10 1	golpe<>militar<>1 2 2	Siqueira<>Santos<>1 1 4
de<>feij<>1 39 1	.<>E<>1 78 2	Deus<>,<>1 3 58
um<>verdadeiro<>1 11 1	e<>eram<>1 28 3	um<>levante<>1 11 1
o<>!<>1 56 7	Batista<>de<>1 1 39	cio<>um<>1 6 11
dizer<>que<>1 1 19	tradu<>o<>1 1 56	derrota<>de<>1 1 39
em<>1936<>1 10 1	o<>faltava<>1 56 1	uma<>quest<>1 6 1
quiserem<>ir<>1 1 1	,<>as<>1 58 8	de<>novembro<>1 39 1
a<>crescer<>1 24 1	o<>se<>1 56 12	foi<>um<>1 5 11
Disseram<>que<>1 1 19	venceu<>Ant<>1 1 8	lio<>ao<>1 1 5
como<>que<>1 2 19	casou<>se<>1 2 12	pagava<>para<>1 1 9

p<>tria<>1 1 1	,<>acompanhando<>1 58 1	guerra<>.<>1 3 79
marechal<>Deodoro<>1 1 1	a<>grande<>1 24 1	Elas<>eram<>1 1 3
deixa<>que<>1 1 19	que<>Deus<>1 19 3	,<>liderado<>1 58 1
todos<>o<>1 2 56	quando<>o<>1 1 56	Continuaram<>a<>1 1 24
litoral<>e<>1 1 28	assombrava<>o<>1 1 56	Meu<>av<>1 1 4
direito<>pol<>1 1 1	e<>iguais<>1 28 1	de<>reza<>1 39 1
redor<>e<>1 1 28	Canudos<>tornou<>1 5 1	Cunha<>,<>1 2 58
N<>o<>1 1 56	ao<>governo<>1 5 4	.<>Zefinha<>1 78 1
.<>Ela<>1 78 1	15<>de<>1 1 39	traziam<>metralhadoras<>1 1 1
seu<>livro<>1 1 2	e<>as<>1 28 8	tinha<>aquela<>1 4 1
mais<>de<>1 2 39	de<>vinte<>1 39 1	a<>avisaram<>1 24 1
se<>incutindo<>1 12 1	a<>hist<>1 24 1	e<>canh<>1 28 2
,<>pedras<>1 58 2	o<>Paulo<>1 56 1	batalhas<>.<>1 1 79
um<>caso<>1 11 1	1919<>.<>1 1 79	da<>regi<>1 14 1
nessa<>data<>1 1 1	homens<>livres<>1 2 1	tona<>,<>1 1 58
do<>dia<>1 9 2	Canudos<>representou<>1 5 1	segunda<>voz<>1 1 1
Eram<>mudan<>1 1 2	vivendo<>sob<>1 1 1	voz<>.<>1 1 79
uma<>capelinha<>1 6 1	rcito<>e<>1 2 28	milagroso<>Ant<>1 1 8
vinham<>as<>1 1 8	Paulo<>.<>1 1 79	e<>tamb<>1 28 1
onde<>se<>1 1 12	mulheres<>que<>1 2 19	ao<>Conselheiro<>1 5 9
o<>evento<>1 56 1	acontecimentos<>mais<>1 1 2	Josefa<>Maria<>1 2 2
canh<>es<>1 2 6	Fazenda<>Rosada<>1 1 1	o<>dos<>1 56 7
rezasse<>o<>1 1 56	Santo<>nascido<>1 3 3	faltava<>gente<>1 1 2
Um<>dia<>1 1 2	esperava<>um<>1 1 11	fosse<>o<>1 1 56
no<>sert<>1 2 5	pa<>s<>1 1 5	.<>Um<>1 78 1
,<>tardinha<>1 58 1	caboclos<>,<>1 1 58	e<>salvava<>1 28 1

comedor<>de<>1 1 39	Conselheiro<>foi<>1 9 5	e<>a<>1 28 24
Ex<>rcito<>1 1 2	poderosamente<>armados<>1 1 1	o<>tinha<>1 56 4
teve<>in<>1 3 2	!<>.<>1 7 79	Conselheiro<>e<>1 9 28
odo<>de<>1 1 39	outras<>tiravam<>1 1 1	que<>v<>1 19 1
:<>ndios<>1 4 1	av<>n<>1 4 7	da<>terra<>1 14 1
dizia<>:<>1 1 4	a<>derrota<>1 24 1	superiores<>aos<>1 1 1
livres<>e<>1 1 28	nordeste<>ao<>1 2 5	contra<>mais<>1 3 2
D<>.<>1 1 79	:<>Cantem<>1 4 1	tico<>de<>1 1 39
o<>ex<>1 56 1	guerra<>e<>1 3 28	o<>vou<>1 56 1
romeiro<>de<>1 1 39	carregamento<>de<>1 1 39	as<>elites<>1 8 1
A<>resist<>1 4 1	um<>per<>1 11 1	!<>Ele<>1 7 6

Depois de submetido este ficheiro às estatísticas *dice* e *qui-quadrado* obtiveram-se os seguintes resultados:

estatística <i>dice</i>	
1060	Ant<>nio<>1 1.0000 8 8 8
por<>ali<>1 1.0000 2 2 2	muitos<>milagres<>1 1.0000 2 2 2
dava<>rem<>1 1.0000 2 2 2	rem<>dio<>1 1.0000 2 2 2
resist<>ncia<>1 1.0000 1 1 1	tiravam<>hinos<>1 1.0000 1 1 1
sem<>dono<>1 1.0000 1 1 1	democr<>ticas<>1 1.0000 1 1 1
obra<>contundente<>1 1.0000 1 1 1	acompanhando<>esse<>1 1.0000 1 1 1
As<>elei<>1 1.0000 1 1 1	aux<>lio<>1 1.0000 1 1 1
sonho<>brasileiro<>1 1.0000 1 1 1	Cantem<>meus<>1 1.0000 1 1 1
1888<>fora<>1 1.0000 1 1 1	mil<>habitantes<>1 1.0000 1 1 1
tria<>sem<>1 1.0000 1 1 1	Quatro<>expedi<>1 1.0000 1 1 1
perdiam<>todas<>1 1.0000 1 1 1	todo<>mundo<>1 1.0000 1 1 1
meus<>filhos<>1 1.0000 1 1 1	levante<>monarquista<>1 1.0000 1 1 1

viesses<>tona<>1 1.0000 1 1 1	nesses<>milagrosos<>1 1.0000 1 1 1
ela<>respondia<>1 1.0000 1 1 1	enviadas<>durante<>1 1.0000 1 1 1
vai<>dizer<>1 1.0000 1 1 1	dessa<>rejei<>1 1.0000 1 1 1
ticas<>dariam<>1 1.0000 1 1 1	abandonada<>pelas<>1 1.0000 1 1 1
sul<>ignoravam<>1 1.0000 1 1 1	aos<>revoltosos<>1 1.0000 1 1 1
tudo<>isso<>1 1.0000 1 1 1	novo<>regime<>1 1.0000 1 1 1
branca<>europeizada<>1 1.0000 1 1 1	aquela<>vozona<>1 1.0000 1 1 1
reportagem<>sobre<>1 1.0000 1 1 1	estranha<>p<>1 1.0000 1 1 1
come<>aram<>1 1.0000 1 1 1	Em<>1896<>1 1.0000 1 1 1
culo<>XIX<>1 1.0000 1 1 1	Joaquim<>Macambira<>1 1.0000 1 1 1
armas<>r<>1 1.0000 1 1 1	13<>noites<>1 1.0000 1 1 1
pelas<>leis<>1 1.0000 1 1 1	hist<>ria<>1 1.0000 1 1 1
numericamente<>muitas<>1 1.0000 1 1 1	Marechal<>Bittencourt<>1 1.0000 1 1 1
trabalhar<>naquelas<>1 1.0000 1 1 1	nome<>dessa<>1 1.0000 1 1 1
R<>gis<>1 1.0000 1 1 1	pareciam<>acabar<>1 1.0000 1 1 1
seus<>dirigentes<>1 1.0000 1 1 1	livre<>traria<>1 1.0000 1 1 1
capelinha<>onde<>1 1.0000 1 1 1	privil<>gios<>1 1.0000 1 1 1
nisso<>porque<>1 1.0000 1 1 1	Estavam<>poderosamente<>1 1.0000 1 1 1
sertanejos<>dirigidos<>1 1.0000 1 1 1	verdadeiro<>profeta<>1 1.0000 1 1 1
escolher<>seus<>1 1.0000 1 1 1	?<>Essa<>1 1.0000 1 1 1
soldados<>traziam<>1 1.0000 1 1 1	antigos<>privil<>1 1.0000 1 1 1
traria<>sal<>1 1.0000 1 1 1	necess<>rio<>1 1.0000 1 1 1
Eu<>sou<>1 1.0000 1 1 1	Zefinha<>nascida<>1 1.0000 1 1 1
vezes<>superiores<>1 1.0000 1 1 1	r<>sticas<>1 1.0000 1 1 1
trabalho<>livre<>1 1.0000 1 1 1	Joana<>Batista<>1 1.0000 1 1 1
vinte<>mil<>1 1.0000 1 1 1	pra<>ganhar<>1 1.0000 1 1 1
naquelas<>obras<>1 1.0000 1 1 1	pedir<>aux<>1 1.0000 1 1 1

muitas<>vezes<>1 1.0000 1 1 1	sertanejo<>assombrava<>1 1.0000 1 1 1
primeiro<>filho<>1 1.0000 1 1 1	elas<>vinham<>1 1.0000 1 1 1
pol<>tico<>1 1.0000 1 1 1	ganhar<>dinheiro<>1 1.0000 1 1 1
vozona<>cheia<>1 1.0000 1 1 1	bem<>feitas<>1 1.0000 1 1 1
per<>odo<>1 1.0000 1 1 1	grande<>proeza<>1 1.0000 1 1 1
iguais<>perante<>1 1.0000 1 1 1	fazia<>milagre<>1 1.0000 1 1 1
munidos<>apenas<>1 1.0000 1 1 1	regime<>trazia<>1 1.0000 1 1 1
publicar<>seu<>1 1.0000 1 1 1	quiserem<>ir<>1 1.0000 1 1 1
p<>tria<>1 1.0000 1 1 1	marechal<>Deodoro<>1 1.0000 1 1 1
direito<>pol<>1 1.0000 1 1 1	traziam<>metralhadoras<>1 1.0000 1 1 1
nessa<>data<>1 1.0000 1 1 1	segunda<>voz<>1 1.0000 1 1 1
vivendo<>sob<>1 1.0000 1 1 1	Fazenda<>Rosada<>1 1.0000 1 1 1
poderosamente<>armados<>1 1.0000 1 1 1	outras<>tiravam<>1 1.0000 1 1 1
superiores<>aos<>1 1.0000 1 1 1	naquele<>tempo<>2 0.8000 2 2 3
livro<>Os<>2 0.8000 2 2 3	chegou<>aqui<>2 0.8000 2 2 3
Os<>Sert<>2 0.8000 2 3 2	minha<>m<>3 0.7500 3 4 4
nio<>Conselheiro<>4 0.7059 6 8 9	Rep<>blica<>5 0.6667 3 3 6
rep<>blica<>5 0.6667 3 3 6	Minha<>av<>5 0.6667 2 2 4
quem<>venceu<>5 0.6667 1 2 1	engenheiro<>militar<>5 0.6667 1 1 2
mais<>impressionantes<>5 0.6667 1 2 1	liderado<>pelo<>5 0.6667 1 1 2
porque<>como<>5 0.6667 1 1 2	Ave<>Maria<>5 0.6667 1 1 2
Uma<>obra<>5 0.6667 1 2 1	Sei<>at<>5 0.6667 1 1 2
Mas<>nunca<>5 0.6667 1 2 1	Papai<>casou<>5 0.6667 1 1 2
Tinha<>muita<>5 0.6667 1 1 2	bom<>pra<>5 0.6667 1 2 1
fazendo<>muitos<>5 0.6667 1 1 2	chamava<>Josefa<>5 0.6667 1 1 2
golpe<>fatal<>5 0.6667 1 2 1	avisaram<>aquela<>5 0.6667 1 1 2
foram<>enviadas<>5 0.6667 1 2 1	como<>enviado<>5 0.6667 1 2 1

acreditava<>nesse<>5 0.6667 1 2 1	todos<>cantavam<>5 0.6667 1 2 1
lutavam<>at<>5 0.6667 1 1 2	dirigidos<>pelo<>5 0.6667 1 1 2
muito<>bem<>5 0.6667 1 2 1	Foi<>essa<>5 0.6667 1 2 1
homens<>lutavam<>5 0.6667 1 2 1	fez<>muitos<>5 0.6667 1 1 2
numa<>nova<>5 0.6667 1 1 2	ano<>seguinte<>5 0.6667 1 2 1
nova<>ordem<>5 0.6667 1 2 1	muita<>agita<>5 0.6667 1 2 1
Foi<>quando<>5 0.6667 1 2 1	foram<>aumentando<>5 0.6667 1 2 1
dia<>15<>5 0.6667 1 2 1	Uma<>cabocla<>5 0.6667 1 2 1
sou<>aquele<>5 0.6667 1 1 2	ficar<>por<>5 0.6667 1 1 2
Mas<>deixa<>5 0.6667 1 2 1	No<>entanto<>5 0.6667 1 2 1
Quando<>achou<>5 0.6667 1 2 1	gostava<>muito<>5 0.6667 1 1 2
pelo<>beato<>5 0.6667 1 2 1	no<>final<>5 0.6667 1 2 1
cabocla<>bonita<>5 0.6667 1 1 2	nova<>tradu<>5 0.6667 1 2 1
acreditava<>nisso<>5 0.6667 1 2 1	mesmo<>pagava<>5 0.6667 1 2 1
nova<>ordem<>5 0.6667 1 2 1	muita<>agita<>5 0.6667 1 2 1
sal<>rios<>5 0.6667 1 1 2	mas<>perdiam<>5 0.6667 1 2 1
Se<>voc<>5 0.6667 1 2 1	lado<>dele<>5 0.6667 1 1 2
fazer<>neg<>5 0.6667 1 2 1	aquele<>pessoal<>5 0.6667 1 2 1
honra<>nacional<>5 0.6667 1 1 2	quem<>batizou<>5 0.6667 1 2 1
pelo<>marechal<>5 0.6667 1 2 1	Para<>compreender<>5 0.6667 1 2 1
latifundi<>rios<>5 0.6667 1 1 2	ex<>rcito<>5 0.6667 1 1 2
seu<>livro<>5 0.6667 1 1 2	homens<>livres<>5 0.6667 1 2 1
Eram<>mudan<>5 0.6667 1 1 2	acontecimentos<>mais<>5 0.6667 1 1 2
faltava<>gente<>5 0.6667 1 1 2	Um<>dia<>5 0.6667 1 1 2
Ex<>rcito<>5 0.6667 1 1 2	foi<>quem<>6 0.5714 2 5 2
dos<>Santos<>7 0.5455 3 7 4	cio<>Jos<>8 0.5000 2 6 2
Sert<>es<>8 0.5000 2 2 6	Ele<>dava<>8 0.5000 2 6 2

in<>cio<>8 0.5000 2 2 6	Vital<>cio<>8 0.5000 2 2 6
militar<>Euclides<>8 0.5000 1 2 2	Deus<>?<>8 0.5000 1 3 1
Professora<>nascido<>8 0.5000 1 1 3	esse<>povo<>8 0.5000 1 1 3
eram<>numericamente<>8 0.5000 1 3 1	No<>ano<>8 0.5000 1 2 2
aqui<>fazendo<>8 0.5000 1 3 1	revoltou<>contra<>8 0.5000 1 1 3
Monte<>Santo<>8 0.5000 1 1 3	tempo<>ainda<>8 0.5000 1 3 1
muita<>gente<>8 0.5000 1 2 2	tempo<>gostava<>8 0.5000 1 3 1
Os<>soldados<>8 0.5000 1 3 1	gis<>nascido<>8 0.5000 1 1 3
revoltado<>contra<>8 0.5000 1 1 3	golpe<>militar<>8 0.5000 1 2 2
Elas<>eram<>8 0.5000 1 1 3	Josefa<>Maria<>8 0.5000 1 2 2
nascido<>em<>9 0.4615 3 3 10	Jos<>dos<>10 0.4444 2 2 7
tinha<>uma<>11 0.4000 2 4 6	mudan<>as<>11 0.4000 2 2 8
Brasil<>:<>11 0.4000 1 1 4	Quando<>teve<>11 0.4000 1 2 3
milagres<>aqui<>11 0.4000 1 2 3	ano<>contra<>11 0.4000 1 2 3
eram<>muito<>11 0.4000 1 3 2	A<>passagem<>11 0.4000 1 4 1
tamb<>m<>11 0.4000 1 1 4	A<>liberta<>11 0.4000 1 4 1
Bahia<>teve<>11 0.4000 1 2 3	Siqueira<>Santos<>11 0.4000 1 1 4
Meu<>av<>11 0.4000 1 1 4	tinha<>aquela<>11 0.4000 1 4 1
teve<>in<>11 0.4000 1 3 2	:<>ndios<>11 0.4000 1 4 1
dizia<>:<>11 0.4000 1 1 4	contra<>mais<>11 0.4000 1 3 2
:<>Cantem<>11 0.4000 1 4 1	A<>resist<>11 0.4000 1 4 1
Conselheiro<>chegou<>12 0.3636 2 9 2	Santo<>Ant<>12 0.3636 2 3 8
ele<>era<>13 0.3333 2 3 9	voc<>s<>13 0.3333 1 1 5
Se<>tinha<>13 0.3333 1 2 4	ao<>redor<>13 0.3333 1 5 1
foi<>nessa<>13 0.3333 1 5 1	ao<>lado<>13 0.3333 1 5 1
na<>boca<>13 0.3333 1 5 1	E<>minha<>13 0.3333 1 2 4
chegar<>ao<>13 0.3333 1 1 5	na<>Fazenda<>13 0.3333 1 5 1

pleno<>sert<>13 0.3333 1 1 5	foi<>cobrir<>13 0.3333 1 5 1
ao<>publicar<>13 0.3333 1 5 1	s<>culo<>13 0.3333 1 5 1
Cantava<>na<>13 0.3333 1 1 5	peçoal<>ao<>13 0.3333 1 1 5
desconhecido<>sert<>13 0.3333 1 1 5	s<>poderia<>13 0.3333 1 5 1
s<>quiserem<>13 0.3333 1 5 1	na<>segunda<>13 0.3333 1 5 1
fatal<>na<>13 0.3333 1 1 5	na<>casa<>13 0.3333 1 5 1
lio<>ao<>13 0.3333 1 1 5	Canudos<>tornou<>13 0.3333 1 5 1
Canudos<>representou<>13 0.3333 1 5 1	Santo<>nascido<>13 0.3333 1 3 3
pa<>s<>13 0.3333 1 1 5	Ele<>n<>14 0.3077 2 6 7
es<>nasceu<>15 0.2857 1 6 1	E<>foi<>15 0.2857 1 2 5
uma<>reportagem<>15 0.2857 1 6 1	expedi<>es<>15 0.2857 1 1 6
uma<>estranha<>15 0.2857 1 6 1	Cunha<>foi<>15 0.2857 1 2 5
O<>Estado<>15 0.2857 1 6 1	jornal<>O<>15 0.2857 1 1 6
Ele<>fez<>15 0.2857 1 6 1	Ele<>fazia<>15 0.2857 1 6 1
O<>novo<>15 0.2857 1 6 1	of<>cio<>15 0.2857 1 1 6
neg<>cio<>15 0.2857 1 1 6	uma<>organiza<>15 0.2857 1 6 1
institui<>es<>15 0.2857 1 1 6	elei<>es<>15 0.2857 1 1 6
O<>desconhecido<>15 0.2857 1 6 1	uma<>trezena<>15 0.2857 1 6 1
nordeste<>s<>15 0.2857 1 2 5	uma<>quest<>15 0.2857 1 6 1
uma<>capelinha<>15 0.2857 1 6 1	no<>sert<>15 0.2857 1 2 5
nDepPatternordeste<>ao<>15 0.2857 1 2 5	O<>Conselheiro<>16 0.2667 2 6 9
da<>Bahia<>17 0.2500 2 14 2	era<>com<>17 0.2500 2 9 7
Euclides<>da<>17 0.2500 2 2 14	da<>Cunha<>17 0.2500 2 14 2
vou<>!<>17 0.2500 1 1 7	es<>foram<>17 0.2500 1 6 2
dos<>escravos<>17 0.2500 1 7 1	com<>carregamento<>17 0.2500 1 7 1
incutindo<>com<>17 0.2500 1 1 7	noites<>!<>17 0.2500 1 1 7
dos<>acontecimentos<>17 0.2500 1 7 1	Veio<>com<>17 0.2500 1 1 7



com<>madeiras<>17 0.2500 1 7 1	com<>antigos<>17 0.2500 1 7 1
aqui<>na<>17 0.2500 1 3 5	O<>livro<>17 0.2500 1 6 2
acabar<>com<>17 0.2500 1 1 7	m<>minha<>17 0.2500 1 4 4
dos<>reis<>17 0.2500 1 7 1	filhos<>!<>17 0.2500 1 1 7
!<>Eu<>17 0.2500 1 7 1	minha<>av<>17 0.2500 1 4 4
!<>Homens<>17 0.2500 1 7 1	beneficiar<>com<>17 0.2500 1 1 7
dos<>latifundi<>17 0.2500 1 7 1	canh<>es<>17 0.2500 1 2 6
n<>o<>18 0.2222 7 7 56	a<>Rep<>18 0.2222 3 24 3
da<>monarquia<>18 0.2222 2 14 4	Josefa<>dos<>18 0.2222 1 2 7
na<>monarquia<>18 0.2222 1 5 4	as<>outras<>18 0.2222 1 8 1
Maria<>!<>18 0.2222 1 2 7	as<>gra<>18 0.2222 1 8 1
as<>13<>18 0.2222 1 8 1	beato<>Ant<>18 0.2222 1 1 8
gra<>as<>18 0.2222 1 1 8	todas<>as<>18 0.2222 1 1 8
ali<>n<>18 0.2222 1 2 7	as<>radicais<>18 0.2222 1 8 1
as<>batalhas<>18 0.2222 1 8 1	!<>E<>18 0.2222 1 7 2
!<>Se<>18 0.2222 1 7 2	!<>Vital<>18 0.2222 1 7 2
venceu<>Ant<>18 0.2222 1 1 8	ao<>governo<>18 0.2222 1 5 4
milagroso<>Ant<>18 0.2222 1 1 8	vinham<>as<>18 0.2222 1 1 8
as<>elites<>18 0.2222 1 8 1	.<>.<>19 0.2038 16 78 79
final<>do<>20 0.2000 1 1 9	para<>vir<>20 0.2000 1 9 1
do<>Rio<>20 0.2000 1 9 1	era<>morena<>20 0.2000 1 9 1
ncia<>do<>20 0.2000 1 1 9	Ela<>era<>20 0.2000 1 1 9
as<>dele<>20 0.2000 1 8 2	do<>litoral<>20 0.2000 1 9 1
elites<>do<>20 0.2000 1 1 9	do<>jornalista<>20 0.2000 1 9 1
era<>mo<>20 0.2000 1 9 1	proeza<>do<>20 0.2000 1 1 9
para<>trabalhar<>20 0.2000 1 9 1	ainda<>era<>20 0.2000 1 1 9
tempo<>dos<>20 0.2000 1 3 7	ria<>do<>20 0.2000 1 1 9

era<>revoltado<>20 0.2000 1 9 1	do<>sul<>20 0.2000 1 9 1
era<>necess<>20 0.2000 1 9 1	revolta<>era<>20 0.2000 1 1 9
para<>chegar<>20 0.2000 1 9 1	convidaram<>para<>20 0.2000 1 1 9
do<>Brasil<>20 0.2000 1 9 1	:<>uma<>20 0.2000 1 4 6
do<>sertanejo<>20 0.2000 1 9 1	Conselheiro<>dizia<>20 0.2000 1 9 1
pagava<>para<>20 0.2000 1 1 9	que<>naquele<>21 0.1905 2 19 2
m<>e<>22 0.1875 3 4 28	que<>ele<>23 0.1818 2 19 3
mas<>era<>23 0.1818 1 2 9	nascida<>em<>23 0.1818 1 1 10
em<>1909<>23 0.1818 1 10 1	em<>1897<>23 0.1818 1 10 1
Conselheiro<>mesmo<>23 0.1818 1 9 2	militar<>do<>23 0.1818 1 2 9
em<>nome<>23 0.1818 1 10 1	em<>1911<>23 0.1818 1 10 1
dio<>para<>23 0.1818 1 2 9	era<>bonita<>23 0.1818 1 9 2
escravos<>em<>23 0.1818 1 1 10	para<>mulheres<>23 0.1818 1 9 2
morte<>em<>23 0.1818 1 1 10	para<>fazer<>23 0.1818 1 9 2
gente<>para<>23 0.1818 1 2 9	bom<>para<>23 0.1818 1 2 9
do<>nordeste<>23 0.1818 1 9 2	em<>1902<>23 0.1818 1 10 1
era<>bom<>23 0.1818 1 9 2	em<>pleno<>23 0.1818 1 10 1
em<>1919<>23 0.1818 1 10 1	em<>1888<>23 0.1818 1 10 1
em<>1936<>23 0.1818 1 10 1	do<>dia<>23 0.1818 1 9 2
av<>n<>23 0.1818 1 4 7	as<>A<>24 0.1667 1 8 4
um<>grupo<>24 0.1667 1 11 1	durante<>um<>24 0.1667 1 1 11
blica<>uma<>24 0.1667 1 6 6	um<>tal<>24 0.1667 1 11 1
s<>!<>24 0.1667 1 5 7	l<>em<>24 0.1667 1 2 10
povo<>do<>24 0.1667 1 3 9	um<>comedor<>24 0.1667 1 11 1
um<>verdadeiro<>24 0.1667 1 11 1	um<>levante<>24 0.1667 1 11 1
um<>caso<>24 0.1667 1 11 1	esperava<>um<>24 0.1667 1 1 11
um<>per<>24 0.1667 1 11 1	sert<>o<>25 0.1639 5 5 56

,<>que<>26 0.1558 6 58 19	queria<>a<>27 0.1538 2 2 24
que<>n<>27 0.1538 2 19 7	monarquia<>para<>27 0.1538 1 4 9
achou<>se<>27 0.1538 1 1 12	se<>esperava<>27 0.1538 1 12 1
J<>se<>27 0.1538 1 1 12	se<>rezasse<>27 0.1538 1 12 1
se<>despachavam<>27 0.1538 1 12 1	Perguntou<>se<>27 0.1538 1 1 12
se<>revoltou<>27 0.1538 1 12 1	tornou<>se<>27 0.1538 1 1 12
um<>ano<>27 0.1538 1 11 2	governo<>do<>27 0.1538 1 4 9
se<>chamava<>27 0.1538 1 12 1	poderia<>se<>27 0.1538 1 1 12
se<>beneficiar<>27 0.1538 1 12 1	mundo<>se<>27 0.1538 1 1 12
se<>incutindo<>27 0.1538 1 12 1	onde<>se<>27 0.1538 1 1 12
!<>Ele<>27 0.1538 1 7 6	a<>rep<>28 0.1481 2 24 3
se<>por<>29 0.1429 1 12 2	um<>Deus<>29 0.1429 1 11 3
se<>l<>29 0.1429 1 12 2	um<>povo<>29 0.1429 1 11 3
Canudos<>para<>29 0.1429 1 5 9	se<>casou<>29 0.1429 1 12 2
do<>s<>29 0.1429 1 9 5	casou<>se<>29 0.1429 1 2 12
se<>por<>29 0.1429 1 12 2	um<>Deus<>29 0.1429 1 11 3
ao<>Conselheiro<>29 0.1429 1 5 9	Conselheiro<>foi<>29 0.1429 1 9 5
de<>Canudos<>30 0.1364 3 39 5	com<>as<>31 0.1333 1 7 8
brasileiro<>da<>31 0.1333 1 1 14	habitantes<>da<>31 0.1333 1 1 14
boca<>da<>31 0.1333 1 1 14	Deodoro<>da<>31 0.1333 1 1 14
em<>Canudos<>31 0.1333 1 10 5	da<>Professora<>31 0.1333 1 14 1
reis<>da<>31 0.1333 1 1 14	loi<>da<>31 0.1333 1 1 14
da<>civiliza<>31 0.1333 1 14 1	jugo<>da<>31 0.1333 1 1 14
da<>Fonseca<>31 0.1333 1 14 1	passagem<>da<>31 0.1333 1 1 14
da<>noite<>31 0.1333 1 14 1	da<>regi<>31 0.1333 1 14 1
da<>terra<>31 0.1333 1 14 1	Conselheiro<>!<>32 0.1250 1 9 7
av<>se<>32 0.1250 1 4 12	foi<>um<>32 0.1250 1 5 11

para<a>33 0.1212 2 9 24	de<um>34 0.1200 3 39 11
.<Ele>35 0.1190 5 78 6	.<O>35 0.1190 5 78 6
da<guerra>36 0.1176 1 14 3	da<rep>36 0.1176 1 14 3
cio<um>36 0.1176 1 6 11	governo<da>37 0.1111 1 4 14
com<um>37 0.1111 1 7 11	um<dos>37 0.1111 1 11 7
se<com>38 0.1053 1 12 7	Jo<o>39 0.1017 3 3 56
o<governo>40 0.1000 3 56 4	que<pareciam>40 0.1000 1 19 1
que<fosse>40 0.1000 1 19 1	achando<que>40 0.1000 1 1 19
era<um>40 0.1000 1 9 11	rio<que>40 0.1000 1 1 19
que<mandou>40 0.1000 1 19 1	cio<da>40 0.1000 1 6 14
que<destru>40 0.1000 1 19 1	data<que>40 0.1000 1 1 19
dizer<que>40 0.1000 1 1 19	Disseram<que>40 0.1000 1 1 19
deixa<que>40 0.1000 1 1 19	que<v>40 0.1000 1 19 1
de<Jesus>41 0.0976 2 39 2	Santos<,>42 0.0968 3 4 58
de<guerra>43 0.0952 2 39 3	de<Santo>43 0.0952 2 39 3
aquele<que>43 0.0952 1 2 19	se<para>43 0.0952 1 12 9
como<que>43 0.0952 1 2 19	mulheres<que>43 0.0952 1 2 19
que<eram>44 0.0909 1 19 3	ele<que>44 0.0909 1 3 19
povo<que>44 0.0909 1 3 19	que<Deus>44 0.0909 1 19 3
de<uma>45 0.0889 2 39 6	,<em>46 0.0882 3 58 10
que<minha>47 0.0870 1 19 4	que<tinha>47 0.0870 1 19 4
o<de>48 0.0842 4 56 39	que<foi>49 0.0833 1 19 5
a<revolta>50 0.0800 1 24 1	essa<a>50 0.0800 1 1 24
crescer<a>50 0.0800 1 1 24	perante<a>50 0.0800 1 1 24
dariam<a>50 0.0800 1 1 24	toda<a>50 0.0800 1 1 24
a<Guerra>50 0.0800 1 24 1	a<ficar>50 0.0800 1 24 1
a<morte>50 0.0800 1 24 1	aram<a>50 0.0800 1 1 24

a<>promessa<>50 0.0800 1 24 1	destru<>a<>50 0.0800 1 1 24
a<>lei<>50 0.0800 1 24 1	a<>rua<>50 0.0800 1 24 1
a<>Campanha<>50 0.0800 1 24 1	a<>igreja<>50 0.0800 1 24 1
mo<>a<>50 0.0800 1 1 24	compreender<>a<>50 0.0800 1 1 24
proclamou<>a<>50 0.0800 1 1 24	trazia<>a<>50 0.0800 1 1 24
sobre<>a<>50 0.0800 1 1 24	a<>crescer<>50 0.0800 1 24 1
a<>grande<>50 0.0800 1 24 1	Continuaram<>a<>50 0.0800 1 1 24
a<>avisaram<>50 0.0800 1 24 1	a<>hist<>50 0.0800 1 24 1
a<>derrota<>50 0.0800 1 24 1	a<>nova<>51 0.0769 1 24 2
a<>todos<>51 0.0769 1 24 2	at<>a<>51 0.0769 1 2 24
fazer<>a<>51 0.0769 1 2 24	contra<>a<>52 0.0741 1 3 24
.<>A<>53 0.0732 3 78 4	e<>o<>54 0.0714 3 28 56
:<>a<>54 0.0714 1 4 24	a<>monarquia<>54 0.0714 1 24 4
o<>golpe<>55 0.0690 2 56 2	o<>acreditava<>55 0.0690 2 56 2
e<>come<>55 0.0690 1 28 1	e<>munidos<>55 0.0690 1 28 1
aumentando<>e<>55 0.0690 1 1 28	e<>sangrentos<>55 0.0690 1 28 1
noite<>e<>55 0.0690 1 1 28	leis<>e<>55 0.0690 1 1 28
e<>armas<>55 0.0690 1 28 1	e<>engenheiro<>55 0.0690 1 28 1
armados<>e<>55 0.0690 1 1 28	impressionantes<>e<>55 0.0690 1 1 28
e<>ela<>55 0.0690 1 28 1	e<>institui<>55 0.0690 1 28 1
vir<>e<>55 0.0690 1 1 28	jornalista<>e<>55 0.0690 1 1 28
hinos<>e<>55 0.0690 1 1 28	Mam<>e<>55 0.0690 1 1 28
Homens<>e<>55 0.0690 1 1 28	granadas<>e<>55 0.0690 1 1 28
terra<>e<>55 0.0690 1 1 28	litoral<>e<>55 0.0690 1 1 28
e<>iguais<>55 0.0690 1 28 1	redor<>e<>55 0.0690 1 1 28
e<>tamb<>55 0.0690 1 28 1	e<>salvava<>55 0.0690 1 28 1
livres<>e<>55 0.0690 1 1 28	dele<>, <>56 0.0667 2 2 58

,<>mas<>56 0.0667 2 58 2	,<>no<>56 0.0667 2 58 2
o<>:<>56 0.0667 2 56 4	dio<>e<>56 0.0667 1 2 28
que<>um<>56 0.0667 1 19 11	e<>fazer<>56 0.0667 1 28 2
pedras<>e<>56 0.0667 1 2 28	e<>mulheres<>56 0.0667 1 28 2
e<>canh<>56 0.0667 1 28 2	rcito<>e<>56 0.0667 1 2 28
o<>sert<>57 0.0656 2 56 5	com<>a<>58 0.0645 1 7 24
que<>se<>58 0.0645 1 19 12	e<>eram<>58 0.0645 1 28 3
guerra<>e<>58 0.0645 1 3 28	blica<>,<>59 0.0625 2 6 58
es<>,<>59 0.0625 2 6 58	av<>e<>59 0.0625 1 4 28
e<>tinha<>59 0.0625 1 28 4	e<>minha<>59 0.0625 1 28 4
para<>o<>60 0.0615 2 9 56	,<>Ant<>61 0.0606 2 58 8
,<>era<>62 0.0597 2 58 9	blica<>e<>63 0.0588 1 6 28
e<>dos<>64 0.0571 1 28 7	e<>as<>65 0.0556 1 28 8
e<>para<>66 0.0541 1 28 9	e<>do<>66 0.0541 1 28 9
Conselheiro<>e<>66 0.0541 1 9 28	,<>o<>67 0.0526 3 58 56
.<>Mas<>68 0.0500 2 78 2	.<>Para<>68 0.0500 2 78 2
.<>Quando<>68 0.0500 2 78 2	.<>Uma<>68 0.0500 2 78 2
.<>Minha<>68 0.0500 2 78 2	.<>Foi<>68 0.0500 2 78 2
.<>No<>68 0.0500 2 78 2	de<>parto<>68 0.0500 1 39 1
sangrentos<>de<>68 0.0500 1 1 39	apenas<>de<>68 0.0500 1 1 39
de<>madrugada<>68 0.0500 1 39 1	de<>paus<>68 0.0500 1 39 1
nunca<>de<>68 0.0500 1 1 39	enviado<>de<>68 0.0500 1 1 39
casa<>de<>68 0.0500 1 1 39	de<>desvalidos<>68 0.0500 1 39 1
profeta<>de<>68 0.0500 1 1 39	tal<>de<>68 0.0500 1 1 39
de<>Marechal<>68 0.0500 1 39 1	de<>escolher<>68 0.0500 1 39 1
Campanha<>de<>68 0.0500 1 1 39	Estado<>de<>68 0.0500 1 1 39
de<>S<>68 0.0500 1 39 1	nasceu<>de<>68 0.0500 1 1 39

poca<>de<>68 0.0500 1 1 39	de<>Joaquim<>68 0.0500 1 39 1
Rio<>de<>68 0.0500 1 1 39	igreja<>de<>68 0.0500 1 1 39
Reginaldo<>de<>68 0.0500 1 1 39	Guerra<>de<>68 0.0500 1 1 39
grupo<>de<>68 0.0500 1 1 39	de<>pedir<>68 0.0500 1 39 1
de<>Monte<>68 0.0500 1 39 1	de<>toda<>68 0.0500 1 39 1
de<>Mattos<>68 0.0500 1 39 1	e<>se<>68 0.0500 1 28 12
caso<>de<>68 0.0500 1 1 39	de<>R<>68 0.0500 1 39 1
promessa<>de<>68 0.0500 1 1 39	de<>Janeiro<>68 0.0500 1 39 1
trezena<>de<>68 0.0500 1 1 39	de<>honra<>68 0.0500 1 39 1
de<>feij<>68 0.0500 1 39 1	Batista<>de<>68 0.0500 1 1 39
derrota<>de<>68 0.0500 1 1 39	de<>novembro<>68 0.0500 1 39 1
de<>reza<>68 0.0500 1 39 1	15<>de<>68 0.0500 1 1 39
de<>vinte<>68 0.0500 1 39 1	comedor<>de<>68 0.0500 1 1 39
odo<>de<>68 0.0500 1 1 39	tico<>de<>68 0.0500 1 1 39
romeiro<>de<>68 0.0500 1 1 39	carregamento<>de<>68 0.0500 1 1 39
rios<>. <>69 0.0494 2 2 79	.<>Jo<>69 0.0494 2 78 3
Jesus<>. <>69 0.0494 2 2 79	nacional<>. <>69 0.0494 2 2 79
muito<>de<>70 0.0488 1 2 39	de<>homens<>70 0.0488 1 39 2
de<>pedras<>70 0.0488 1 39 2	Maria<>de<>70 0.0488 1 2 39
rcito<>de<>70 0.0488 1 2 39	de<>mudan<>70 0.0488 1 39 2
at<>de<>70 0.0488 1 2 39	de<>muita<>70 0.0488 1 39 2
de<>l<>70 0.0488 1 39 2	mais<>de<>70 0.0488 1 2 39
monarquia<>. <>71 0.0482 2 4 79	Canudos<>. <>72 0.0476 2 5 79
e<>da<>72 0.0476 1 28 14	teve<>de<>72 0.0476 1 3 39
de<>Deus<>72 0.0476 1 39 3	eram<>de<>72 0.0476 1 3 39
blica<>. <>73 0.0471 2 6 79	, <>e<>74 0.0465 2 58 28
Conselheiro<>. <>75 0.0455 2 9 79	cio<>de<>76 0.0444 1 6 39

e<>que<>77 0.0426 1 28 19	a<>e<>78 0.0385 1 24 28
e<>a<>78 0.0385 1 28 24	rejei<>o<>79 0.0351 1 1 56
representou<>o<>79 0.0351 1 1 56	o<>branca<>79 0.0351 1 56 1
cobrir<>o<>79 0.0351 1 1 56	o<>At<>79 0.0351 1 56 1
v<>o<>79 0.0351 1 1 56	S<>o<>79 0.0351 1 1 56
o<>primeiro<>79 0.0351 1 56 1	o<>jornal<>79 0.0351 1 56 1
mandou<>o<>79 0.0351 1 1 56	o<>imprevisto<>79 0.0351 1 56 1
liberta<>o<>79 0.0351 1 1 56	o<>pa<>79 0.0351 1 56 1
fora<>o<>79 0.0351 1 1 56	o<>Siqueira<>79 0.0351 1 56 1
feij<>o<>79 0.0351 1 1 56	o<>of<>79 0.0351 1 56 1
organiza<>o<>79 0.0351 1 1 56	civiliza<>o<>79 0.0351 1 1 56
o<>ter<>79 0.0351 1 56 1	sob<>o<>79 0.0351 1 1 56
o<>sonho<>79 0.0351 1 56 1	o<>direito<>79 0.0351 1 56 1
ignoravam<>o<>79 0.0351 1 1 56	ter<>o<>79 0.0351 1 1 56
agita<>o<>79 0.0351 1 1 56	quest<>o<>79 0.0351 1 1 56
o<>convidaram<>79 0.0351 1 56 1	o<>vai<>79 0.0351 1 56 1
o<>jugo<>79 0.0351 1 56 1	o<>Reginaldo<>79 0.0351 1 56 1
o<>Ex<>79 0.0351 1 56 1	o<>viesses<>79 0.0351 1 56 1
o<>trabalho<>79 0.0351 1 56 1	regi<>o<>79 0.0351 1 1 56
At<>o<>79 0.0351 1 1 56	tradu<>o<>79 0.0351 1 1 56
o<>faltava<>79 0.0351 1 56 1	quando<>o<>79 0.0351 1 1 56
assombrava<>o<>79 0.0351 1 1 56	N<>o<>79 0.0351 1 1 56
o<>Paulo<>79 0.0351 1 56 1	o<>evento<>79 0.0351 1 56 1
rezasse<>o<>79 0.0351 1 1 56	fosse<>o<>79 0.0351 1 1 56
o<>ex<>79 0.0351 1 56 1	o<>vou<>79 0.0351 1 56 1
o<>queria<>80 0.0345 1 56 2	canh<>o<>80 0.0345 1 2 56
o<>nacional<>80 0.0345 1 56 2	o<>canh<>80 0.0345 1 56 2



o<>in<>80 0.0345 1 56 2	o<>nordeste<>80 0.0345 1 56 2
Para<>o<>80 0.0345 1 2 56	todos<>o<>80 0.0345 1 2 56
,<>Joana<>81 0.0339 1 58 1	1897<>,<>81 0.0339 1 1 58
Bittencourt<>,<>81 0.0339 1 1 58	o<>tempo<>81 0.0339 1 56 3
entanto<>,<>81 0.0339 1 1 58	rua<>,<>81 0.0339 1 1 58
Mattos<>,<>81 0.0339 1 1 58	contra<>o<>81 0.0339 1 3 56
,<>romeiro<>81 0.0339 1 58 1	mulatos<>,<>81 0.0339 1 1 58
Janeiro<>,<>81 0.0339 1 1 58	paus<>,<>81 0.0339 1 1 58
o<>povo<>81 0.0339 1 56 3	1896<>,<>81 0.0339 1 1 58
Rosada<>,<>81 0.0339 1 1 58	seguinte<>,<>81 0.0339 1 1 58
ndios<>,<>81 0.0339 1 1 58	,<>granadas<>81 0.0339 1 58 1
radicais<>,<>81 0.0339 1 1 58	filho<>,<>81 0.0339 1 1 58
,<>numa<>81 0.0339 1 58 1	tardinha<>,<>81 0.0339 1 1 58
Essa<>,<>81 0.0339 1 1 58	,<>vivendo<>81 0.0339 1 58 1
,<>abandonada<>81 0.0339 1 58 1	,<>loi<>81 0.0339 1 58 1
ir<>,<>81 0.0339 1 1 58	metralhadoras<>,<>81 0.0339 1 1 58
,<>achando<>81 0.0339 1 58 1	contundente<>,<>81 0.0339 1 1 58
novembro<>,<>81 0.0339 1 1 58	obras<>,<>81 0.0339 1 1 58
XIX<>,<>81 0.0339 1 1 58	,<>D<>81 0.0339 1 58 1
teve<>o<>81 0.0339 1 3 56	,<>proclamou<>81 0.0339 1 58 1
evento<>,<>81 0.0339 1 1 58	revoltosos<>,<>81 0.0339 1 1 58
dirigentes<>,<>81 0.0339 1 1 58	,<>caboclos<>81 0.0339 1 58 1
,<>pretos<>81 0.0339 1 58 1	dono<>,<>81 0.0339 1 1 58
,<>elas<>81 0.0339 1 58 1	,<>todo<>81 0.0339 1 58 1
,<>mulatos<>81 0.0339 1 58 1	Fonseca<>,<>81 0.0339 1 1 58
morena<>,<>81 0.0339 1 1 58	,<>acompanhando<>81 0.0339 1 58 1
,<>liderado<>81 0.0339 1 58 1	tona<>,<>81 0.0339 1 1 58

,<>tardinha<>81 0.0339 1 58 1	caboclos<>,<>81 0.0339 1 1 58
dia<>,<>82 0.0333 1 2 58	,<>bom<>82 0.0333 1 58 2
mulheres<>,<>82 0.0333 1 2 58	,<>foram<>82 0.0333 1 58 2
A<>o<>82 0.0333 1 4 56	Bahia<>,<>82 0.0333 1 2 58
,<>mesmo<>82 0.0333 1 58 2	,<>queria<>82 0.0333 1 58 2
gente<>,<>82 0.0333 1 2 58	milagres<>,<>82 0.0333 1 2 58
,<>todos<>82 0.0333 1 58 2	pedras<>,<>82 0.0333 1 2 58
bonita<>,<>82 0.0333 1 2 58	,<>homens<>82 0.0333 1 58 2
,<>como<>82 0.0333 1 58 2	Cunha<>,<>82 0.0333 1 2 58
,<>pedras<>82 0.0333 1 58 2	o<>tinha<>82 0.0333 1 56 4
,<>Jo<>83 0.0328 1 58 3	guerra<>,<>83 0.0328 1 3 58
,<>ele<>83 0.0328 1 58 3	o<>na<>83 0.0328 1 56 5
,<>teve<>83 0.0328 1 58 3	o<>s<>83 0.0328 1 56 5
aqui<>,<>83 0.0328 1 3 58	Deus<>,<>83 0.0328 1 3 58
governo<>,<>84 0.0323 1 4 58	tinha<>,<>84 0.0323 1 4 58
monarquia<>,<>84 0.0323 1 4 58	,<>ao<>85 0.0317 1 58 5
s<>,<>85 0.0317 1 5 58	,<>foi<>85 0.0317 1 58 5
,<>na<>85 0.0317 1 58 5	com<>o<>85 0.0317 1 7 56
o<>!<>85 0.0317 1 56 7	o<>dos<>85 0.0317 1 56 7
cio<>,<>86 0.0312 1 6 58	,<>n<>87 0.0308 1 58 7
nio<>,<>88 0.0303 1 8 58	,<>as<>88 0.0303 1 58 8
o<>.<>89 0.0296 2 56 79	o<>se<>90 0.0294 1 56 12
,<>um<>91 0.0290 1 58 11	o<>da<>92 0.0286 1 56 14
o<>que<>93 0.0267 1 56 19	que<>o<>93 0.0267 1 19 56
.<>Disseram<>94 0.0253 1 78 1	.<>Meu<>94 0.0253 1 78 1
.<>Elas<>94 0.0253 1 78 1	.<>J<>94 0.0253 1 78 1
.<>Estavam<>94 0.0253 1 78 1	.<>As<>94 0.0253 1 78 1

.<>Papai<>94 0.0253 1 78 1	.<>Continuaram<>94 0.0253 1 78 1
.<>Eram<>94 0.0253 1 78 1	.<>Perguntou<>94 0.0253 1 78 1
.<>Tinha<>94 0.0253 1 78 1	.<>Quatro<>94 0.0253 1 78 1
.<>poca<>94 0.0253 1 78 1	.<>Veio<>94 0.0253 1 78 1
.<>sertanejos<>94 0.0253 1 78 1	.<>N<>94 0.0253 1 78 1
.<>Mam<>94 0.0253 1 78 1	.<>tudo<>94 0.0253 1 78 1
.<>Cantava<>94 0.0253 1 78 1	.<>Sei<>94 0.0253 1 78 1
.<>Ave<>94 0.0253 1 78 1	.<>Zefinha<>94 0.0253 1 78 1
.<>Ela<>94 0.0253 1 78 1	.<>Um<>94 0.0253 1 78 1
.<>Se<>95 0.0250 1 78 2	pretos<>. <>95 0.0250 1 1 79
imprevisto<>. <>95 0.0250 1 1 79	cantavam<>. <>95 0.0250 1 1 79
despachavam<>. <>95 0.0250 1 1 79	1911<>. <>95 0.0250 1 1 79
Macambira<>. <>95 0.0250 1 1 79	gios<>. <>95 0.0250 1 1 79
dinheiro<>. <>95 0.0250 1 1 79	desvalidos<>. <>95 0.0250 1 1 79
isso<>. <>95 0.0250 1 1 79	.<>Vital<>95 0.0250 1 78 2
ordem<>. <>95 0.0250 1 1 79	monarquista<>. <>95 0.0250 1 1 79
respondia<>. <>95 0.0250 1 1 79	cheia<>. <>95 0.0250 1 1 79
madrugada<>. <>95 0.0250 1 1 79	europalizada<>. <>95 0.0250 1 1 79
lei<>. <>95 0.0250 1 1 79	reza<>. <>95 0.0250 1 1 79
madeiras<>. <>95 0.0250 1 1 79	sticas<>. <>95 0.0250 1 1 79
.<>Josefa<>95 0.0250 1 78 2	1936<>. <>95 0.0250 1 1 79
.<>Euclides<>95 0.0250 1 78 2	milagre<>. <>95 0.0250 1 1 79
a<>o<>95 0.0250 1 24 56	feitas<>. <>95 0.0250 1 1 79
salvava<>. <>95 0.0250 1 1 79	1902<>. <>95 0.0250 1 1 79
1909<>. <>95 0.0250 1 1 79	batizou<>. <>95 0.0250 1 1 79
parto<>. <>95 0.0250 1 1 79	.<>E<>95 0.0250 1 78 2
batalhas<>. <>95 0.0250 1 1 79	1919<>. <>95 0.0250 1 1 79

voz<>.<>95 0.0250 1 1 79	Paulo<>.<>95 0.0250 1 1 79
D<>.<>95 0.0250 1 1 79	l<>.<>96 0.0247 1 2 79
bonita<>.<>96 0.0247 1 2 79	.<>Os<>96 0.0247 1 78 3
mesmo<>.<>96 0.0247 1 2 79	ali<>.<>96 0.0247 1 2 79
casou<>.<>96 0.0247 1 2 79	povo<>.<>97 0.0244 1 3 79
Deus<>.<>97 0.0244 1 3 79	,<>a<>97 0.0244 1 58 24
guerra<>.<>97 0.0244 1 3 79	Santos<>.<>98 0.0241 1 4 79
governo<>.<>98 0.0241 1 4 79	.<>Canudos<>98 0.0241 1 78 5
av<>.<>98 0.0241 1 4 79	es<>.<>99 0.0235 1 6 79
.<>Ant<>100 0.0233 1 78 8	e<>,<>100 0.0233 1 28 58
!<>.<>100 0.0233 1 7 79	nio<>.<>101 0.0230 1 8 79
.<>era<>101 0.0230 1 78 9	o<>,<>102 0.0175 1 56 58

<i>estatística qui-quadrado</i>	
1060	Ant<>nio<>1 1060.0000 8 8 8
por<>ali<>1 1060.0000 2 2 2	muitos<>milagres<>1 1060.0000 2 2 2
dava<>rem<>1 1060.0000 2 2 2	rem<>dio<>1 1060.0000 2 2 2
resist<>ncia<>1 1060.0000 1 1 1	tiravam<>hinos<>1 1060.0000 1 1 1
sem<>dono<>1 1060.0000 1 1 1	democr<>ticas<>1 1060.0000 1 1 1
obra<>contundente<>1 1060.0000 1 1 1	acompanhando<>esse<>1 1060.0000 1 1 1
As<>elei<>1 1060.0000 1 1 1	aux<>lio<>1 1060.0000 1 1 1
sonho<>brasileiro<>1 1060.0000 1 1 1	Cantem<>meus<>1 1060.0000 1 1 1
1888<>fora<>1 1060.0000 1 1 1	mil<>habitantes<>1 1060.0000 1 1 1
tria<>sem<>1 1060.0000 1 1 1	Quatro<>expedi<>1 1060.0000 1 1 1
perdiam<>todas<>1 1060.0000 1 1 1	todo<>mundo<>1 1060.0000 1 1 1
meus<>filhos<>1 1060.0000 1 1 1	levante<>monarquista<>1 1060.0000 1 1 1
viesses<>tona<>1 1060.0000 1 1 1	nesse<>milagroso<>1 1060.0000 1 1 1

ela<>respondia<>1 1060.0000 1 1 1	enviadas<>durante<>1 1060.0000 1 1 1
vai<>dizer<>1 1060.0000 1 1 1	dessa<>rejei<>1 1060.0000 1 1 1
ticas<>dariam<>1 1060.0000 1 1 1	abandonada<>pelas<>1 1060.0000 1 1 1
sul<>ignoravam<>1 1060.0000 1 1 1	aos<>revoltosos<>1 1060.0000 1 1 1
tudo<>isso<>1 1060.0000 1 1 1	novo<>regime<>1 1060.0000 1 1 1
branca<>europeizada<>1 1060.0000 1 1 1	aquela<>vozona<>1 1060.0000 1 1 1
reportagem<>sobre<>1 1060.0000 1 1 1	estranha<>p<>1 1060.0000 1 1 1
come<>aram<>1 1060.0000 1 1 1	Em<>1896<>1 1060.0000 1 1 1
culo<>XIX<>1 1060.0000 1 1 1	Joaquim<>Macambira<>1 1060.0000 1 1 1
armas<>r<>1 1060.0000 1 1 1	13<>noites<>1 1060.0000 1 1 1
pelas<>leis<>1 1060.0000 1 1 1	hist<>ria<>1 1060.0000 1 1 1
numericamente<>muitas<>1 1060.0000 1 1 1	Marechal<>Bittencourt<>1 1060.0000 1 1 1
trabalhar<>naquelas<>1 1060.0000 1 1 1	nome<>dessa<>1 1060.0000 1 1 1
R<>gis<>1 1060.0000 1 1 1	pareciam<>acabar<>1 1060.0000 1 1 1
seus<>dirigentes<>1 1060.0000 1 1 1	livre<>traria<>1 1060.0000 1 1 1
capelinha<>onde<>1 1060.0000 1 1 1	privil<>gios<>1 1060.0000 1 1 1
nisso<>porque<>1 1060.0000 1 1 1	Estavam<>poderosamente<>1 1060.0000 1 1 1
sertanejos<>dirigidos<>1 1060.0000 1 1 1	verdadeiro<>profeta<>1 1060.0000 1 1 1
escolher<>seus<>1 1060.0000 1 1 1	?<>Essa<>1 1060.0000 1 1 1
soldados<>traziam<>1 1060.0000 1 1 1	antigos<>privil<>1 1060.0000 1 1 1
traria<>sal<>1 1060.0000 1 1 1	necess<>rio<>1 1060.0000 1 1 1
Eu<>sou<>1 1060.0000 1 1 1	Zefinha<>nascida<>1 1060.0000 1 1 1
vezes<>superiores<>1 1060.0000 1 1 1	r<>sticas<>1 1060.0000 1 1 1
trabalho<>livre<>1 1060.0000 1 1 1	Joana<>Batista<>1 1060.0000 1 1 1
vinte<>mil<>1 1060.0000 1 1 1	pra<>ganhar<>1 1060.0000 1 1 1
naquelas<>obras<>1 1060.0000 1 1 1	pedir<>aux<>1 1060.0000 1 1 1

muitas<>vezes<>1 1060.0000 1 1 1	sertanejo<>assombrava<>1 1060.0000 1 1 1
primeiro<>filho<>1 1060.0000 1 1 1	elas<>vinham<>1 1060.0000 1 1 1
pol<>tico<>1 1060.0000 1 1 1	ganhar<>dinheiro<>1 1060.0000 1 1 1
vozona<>cheia<>1 1060.0000 1 1 1	bem<>feitas<>1 1060.0000 1 1 1
per<>odo<>1 1060.0000 1 1 1	grande<>proeza<>1 1060.0000 1 1 1
iguais<>perante<>1 1060.0000 1 1 1	fazia<>milagre<>1 1060.0000 1 1 1
munidos<>apenas<>1 1060.0000 1 1 1	regime<>trazia<>1 1060.0000 1 1 1
publicar<>seu<>1 1060.0000 1 1 1	quiserem<>ir<>1 1060.0000 1 1 1
p<>tria<>1 1060.0000 1 1 1	marechal<>Deodoro<>1 1060.0000 1 1 1
direito<>pol<>1 1060.0000 1 1 1	traziam<>metralhadoras<>1 1060.0000 1 1 1
nessa<>data<>1 1060.0000 1 1 1	segunda<>voz<>1 1060.0000 1 1 1
vivendo<>sob<>1 1060.0000 1 1 1	Fazenda<>Rosada<>1 1060.0000 1 1 1
poderosamente<>armados<>1 1060.0000 1 1 1	outras<>tiravam<>1 1060.0000 1 1 1
superiores<>aos<>1 1060.0000 1 1 1	naquele<>tempo<>2 705.9987 2 2 3
livro<>Os<>2 705.9987 2 2 3	chegou<>aqui<>2 705.9987 2 2 3
Os<>Sert<>2 705.9987 2 3 2	minha<>m<>3 594.7453 3 4 4
quem<>venceu<>4 529.4995 1 2 1	engenheiro<>militar<>4 529.4995 1 1 2
mais<>impressionantes<>4 529.4995 1 2 1	liderado<>pelo<>4 529.4995 1 1 2
porque<>como<>4 529.4995 1 1 2	Ave<>Maria<>4 529.4995 1 1 2
Uma<>obra<>4 529.4995 1 2 1	Sei<>at<>4 529.4995 1 1 2
Mas<>nunca<>4 529.4995 1 2 1	Papai<>casou<>4 529.4995 1 1 2
Tinha<>muita<>4 529.4995 1 1 2	bom<>pra<>4 529.4995 1 2 1
fazendo<>muitos<>4 529.4995 1 1 2	chamava<>Josefa<>4 529.4995 1 1 2
golpe<>fatal<>4 529.4995 1 2 1	avisaram<>aquela<>4 529.4995 1 1 2
foram<>enviadas<>4 529.4995 1 2 1	como<>enviado<>4 529.4995 1 2 1
acreditava<>nesse<>4 529.4995 1 2 1	todos<>cantavam<>4 529.4995 1 2 1

lutavam<>at<>4 529.4995 1 1 2	dirigidos<>pele<>4 529.4995 1 1 2
muito<>bem<>4 529.4995 1 2 1	Foi<>essa<>4 529.4995 1 2 1
homens<>lutavam<>4 529.4995 1 2 1	fez<>muitos<>4 529.4995 1 1 2
numa<>nova<>4 529.4995 1 1 2	ano<>seguinte<>4 529.4995 1 2 1
nova<>ordem<>4 529.4995 1 2 1	muita<>agita<>4 529.4995 1 2 1
Foi<>quando<>4 529.4995 1 2 1	foram<>aumentando<>4 529.4995 1 2 1
dia<>15<>4 529.4995 1 2 1	Uma<>cabocla<>4 529.4995 1 2 1
sou<>aquela<>4 529.4995 1 1 2	ficar<>por<>4 529.4995 1 1 2
Mas<>deixa<>4 529.4995 1 2 1	No<>entanto<>4 529.4995 1 2 1
Quando<>achou<>4 529.4995 1 2 1	gostava<>muito<>4 529.4995 1 1 2
pele<>beato<>4 529.4995 1 2 1	no<>final<>4 529.4995 1 2 1
cabocla<>bonita<>4 529.4995 1 1 2	nova<>tradu<>4 529.4995 1 2 1
acreditava<>nisso<>4 529.4995 1 2 1	mesmo<>pagava<>4 529.4995 1 2 1
sal<>rios<>4 529.4995 1 1 2	mas<>perdiam<>4 529.4995 1 2 1
Se<>voc<>4 529.4995 1 2 1	lado<>dele<>4 529.4995 1 1 2
fazer<>neg<>4 529.4995 1 2 1	aquela<>pessoal<>4 529.4995 1 2 1
honra<>nacional<>4 529.4995 1 1 2	quem<>batizou<>4 529.4995 1 2 1
pele<>marechal<>4 529.4995 1 2 1	Para<>compreender<>4 529.4995 1 2 1
latifundi<>rios<>4 529.4995 1 1 2	ex<>rcito<>4 529.4995 1 1 2
seu<>livro<>4 529.4995 1 1 2	homens<>livres<>4 529.4995 1 2 1
Eram<>mudan<>4 529.4995 1 1 2	acontecimentos<>mais<>4 529.4995 1 1 2
faltava<>gente<>4 529.4995 1 1 2	Um<>dia<>4 529.4995 1 1 2
Ex<>rcito<>4 529.4995 1 1 2	Minha<>av<>5 528.9981 2 2 4
Rep<>blica<>6 528.4957 3 3 6	rep<>blica<>6 528.4957 3 3 6
nio<>Conselheiro<>7 526.4777 6 8 9	foi<>quem<>8 422.7977 2 5 2
Deus<>?<>9 352.6660 1 3 1	Professora<>nascido<>9 352.6660 1 1 3
esse<>povo<>9 352.6660 1 1 3	eram<>numericamente<>9 352.6660 1 3 1

aqui<>fazendo<>9 352.6660 1 3 1	revoltou<>contra<>9 352.6660 1 1 3
Monte<>Santo<>9 352.6660 1 1 3	tempo<>ainda<>9 352.6660 1 3 1
tempo<>gostava<>9 352.6660 1 3 1	Os<>soldados<>9 352.6660 1 3 1
gis<>nascido<>9 352.6660 1 1 3	revoltado<>contra<>9 352.6660 1 1 3
Elas<>eram<>9 352.6660 1 1 3	cio<>Jos<>10 351.9975 2 6 2
Sert<>es<>10 351.9975 2 2 6	Ele<>dava<>10 351.9975 2 6 2
in<>cio<>10 351.9975 2 2 6	Vital<>cio<>10 351.9975 2 2 6
dos<>Santos<>11 338.2423 3 7 4	nascido<>em<>12 315.8940 3 3 10
Jos<>dos<>13 301.4259 2 2 7	Brasil<>:<>14 264.2493 1 1 4
A<>passagem<>14 264.2493 1 4 1	tamb<>m<>14 264.2493 1 1 4
A<>liberta<>14 264.2493 1 4 1	Siqueira<>Santos<>14 264.2493 1 1 4
Meu<>av<>14 264.2493 1 1 4	tinha<>aquela<>14 264.2493 1 4 1
:<>ndios<>14 264.2493 1 4 1	dizia<>:<>14 264.2493 1 1 4
:<>Cantem<>14 264.2493 1 4 1	A<>resist<>14 264.2493 1 4 1
militar<>Euclides<>15 263.9991 1 2 2	No<>ano<>15 263.9991 1 2 2
muita<>gente<>15 263.9991 1 2 2	golpe<>militar<>15 263.9991 1 2 2
Josefa<>Maria<>15 263.9991 1 2 2	mudan<>as<>16 263.4972 2 2 8
Conselheiro<>chegou<>17 233.9971 2 9 2	voc<>s<>18 211.1992 1 1 5
ao<>redor<>18 211.1992 1 5 1	foi<>nessa<>18 211.1992 1 5 1
ao<>lado<>18 211.1992 1 5 1	na<>boca<>18 211.1992 1 5 1
chegar<>ao<>18 211.1992 1 1 5	na<>Fazenda<>18 211.1992 1 5 1
pleno<>sert<>18 211.1992 1 1 5	foi<>cobrir<>18 211.1992 1 5 1
ao<>publicar<>18 211.1992 1 5 1	s<>culo<>18 211.1992 1 5 1
Cantava<>na<>18 211.1992 1 1 5	peessoal<>ao<>18 211.1992 1 1 5
desconhecido<>sert<>18 211.1992 1 1 5	s<>poderia<>18 211.1992 1 5 1
s<>quiserem<>18 211.1992 1 5 1	na<>segunda<>18 211.1992 1 5 1
fatal<>na<>18 211.1992 1 1 5	na<>casa<>18 211.1992 1 5 1



lio<>ao<>18 211.1992 1 1 5	Canudos<>tornou<>18 211.1992 1 5 1
Canudos<>representou<>18 211.1992 1 5 1	pa<>s<>18 211.1992 1 1 5
es<>nasceu<>19 175.8325 1 6 1	uma<>reportagem<>19 175.8325 1 6 1
expedi<>es<>19 175.8325 1 1 6	uma<>estranha<>19 175.8325 1 6 1
O<>Estado<>19 175.8325 1 6 1	jornal<>O<>19 175.8325 1 1 6
Ele<>fez<>19 175.8325 1 6 1	Ele<>fazia<>19 175.8325 1 6 1
O<>novo<>19 175.8325 1 6 1	of<>cio<>19 175.8325 1 1 6
neg<>cio<>19 175.8325 1 1 6	uma<>organiza<>19 175.8325 1 6 1
institui<>es<>19 175.8325 1 1 6	elei<>es<>19 175.8325 1 1 6
O<>desconhecido<>19 175.8325 1 6 1	uma<>trezena<>19 175.8325 1 6 1
es<>democr<>19 175.8325 1 6 1	uma<>quest<>19 175.8325 1 6 1
uma<>capelinha<>19 175.8325 1 6 1	Quando<>teve<>20 175.4992 1 2 3
milagres<>aqui<>20 175.4992 1 2 3	ano<>contra<>20 175.4992 1 2 3
eram<>muito<>20 175.4992 1 3 2	Bahia<>teve<>20 175.4992 1 2 3
teve<>in<>20 175.4992 1 3 2	contra<>mais<>20 175.4992 1 3 2
Santo<>Ant<>21 174.4964 2 3 8	tinha<>uma<>22 174.3302 2 4 6
ele<>era<>23 154.8114 2 3 9	vou<>!<>24 150.5706 1 1 7
dos<>escravos<>24 150.5706 1 7 1	com<>carregamento<>24 150.5706 1 7 1
incutindo<>com<>24 150.5706 1 1 7	noites<>!<>24 150.5706 1 1 7
dos<>acontecimentos<>24 150.5706 1 7 1	Veio<>com<>24 150.5706 1 1 7
com<>madeiras<>24 150.5706 1 7 1	com<>antigos<>24 150.5706 1 7 1
acabar<>com<>24 150.5706 1 1 7	dos<>reis<>24 150.5706 1 7 1
filhos<>!<>24 150.5706 1 1 7	!<>Eu<>24 150.5706 1 7 1
!<>Homens<>24 150.5706 1 7 1	beneficiar<>com<>24 150.5706 1 1 7
dos<>latifundi<>24 150.5706 1 7 1	da<>Bahia<>25 149.7110 2 14 2
Euclides<>da<>25 149.7110 2 2 14	da<>Cunha<>25 149.7110 2 14 2
as<>outras<>26 131.6242 1 8 1	as<>gra<>26 131.6242 1 8 1

as<>13<>26 131.6242 1 8 1	beato<>Ant<>26 131.6242 1 1 8
gra<>as<>26 131.6242 1 1 8	todas<>as<>26 131.6242 1 1 8
as<>radicais<>26 131.6242 1 8 1	as<>batalhas<>26 131.6242 1 8 1
venceu<>Ant<>26 131.6242 1 1 8	milagroso<>Ant<>26 131.6242 1 1 8
vinham<>as<>26 131.6242 1 1 8	as<>elites<>26 131.6242 1 8 1
Se<>tinha<>27 131.2495 1 2 4	E<>minha<>27 131.2495 1 2 4
a<>Rep<>28 129.8675 3 24 3	n<>o<>29 126.3343 7 7 56
final<>do<>30 116.8880 1 1 9	para<>vir<>30 116.8880 1 9 1
do<>Rio<>30 116.8880 1 9 1	era<>morena<>30 116.8880 1 9 1
ncia<>do<>30 116.8880 1 1 9	Ela<>era<>30 116.8880 1 1 9
do<>litoral<>30 116.8880 1 9 1	elites<>do<>30 116.8880 1 1 9
do<>jornalista<>30 116.8880 1 9 1	era<>mo<>30 116.8880 1 9 1
proeza<>do<>30 116.8880 1 1 9	para<>trabalhar<>30 116.8880 1 9 1
ainda<>era<>30 116.8880 1 1 9	ria<>do<>30 116.8880 1 1 9
era<>revoltado<>30 116.8880 1 9 1	do<>sul<>30 116.8880 1 9 1
era<>necess<>30 116.8880 1 9 1	revolta<>era<>30 116.8880 1 1 9
para<>chegar<>30 116.8880 1 9 1	convidaram<>para<>30 116.8880 1 1 9
do<>Brasil<>30 116.8880 1 9 1	do<>sertanejo<>30 116.8880 1 9 1
Conselheiro<>dizia<>30 116.8880 1 9 1	pagava<>para<>30 116.8880 1 1 9
Santo<>nascido<>31 116.4445 1 3 3	que<>naquele<>32 109.7861 2 19 2
nascida<>em<>33 105.0992 1 1 10	em<>1909<>33 105.0992 1 10 1
em<>1897<>33 105.0992 1 10 1	em<>nome<>33 105.0992 1 10 1
em<>1911<>33 105.0992 1 10 1	escravos<>em<>33 105.0992 1 1 10
morte<>em<>33 105.0992 1 1 10	em<>1902<>33 105.0992 1 10 1
em<>pleno<>33 105.0992 1 10 1	em<>1919<>33 105.0992 1 10 1
em<>1888<>33 105.0992 1 10 1	em<>1936<>33 105.0992 1 10 1
E<>foi<>34 104.6999 1 2 5	Cunha<>foi<>34 104.6999 1 2 5

nordeste<>s<>34 104.6999 1 2 5	no<>sert<>34 104.6999 1 2 5
nordeste<>ao<>34 104.6999 1 2 5	Ele<>n<>35 98.1926 2 6 7
um<>grupo<>36 95.4537 1 11 1	durante<>um<>36 95.4537 1 11 1
um<>tal<>36 95.4537 1 11 1	um<>comedor<>36 95.4537 1 11 1
um<>verdadeiro<>36 95.4537 1 11 1	um<>levante<>36 95.4537 1 11 1
um<>caso<>36 95.4537 1 11 1	esperava<>um<>36 95.4537 1 11 1
um<>per<>36 95.4537 1 11 1	sert<>o<>37 90.0677 5 5 56
achou<>se<>38 87.4158 1 1 12	se<>esperava<>38 87.4158 1 12 1
J<>se<>38 87.4158 1 1 12	se<>rezasse<>38 87.4158 1 12 1
se<>despachavam<>38 87.4158 1 12 1	Perguntou<>se<>38 87.4158 1 1 12
se<>revoltou<>38 87.4158 1 12 1	tornou<>se<>38 87.4158 1 1 12
se<>chamava<>38 87.4158 1 12 1	poderia<>se<>38 87.4158 1 1 12
se<>beneficiar<>38 87.4158 1 12 1	mundo<>se<>38 87.4158 1 1 12
se<>incutindo<>38 87.4158 1 12 1	onde<>se<>38 87.4158 1 1 12
es<>foram<>39 87.0003 1 6 2	O<>livro<>39 87.0003 1 6 2
canh<>es<>39 87.0003 1 2 6	queria<>a<>40 86.4965 2 2 24
m<>e<>41 81.7438 3 4 28	O<>Conselheiro<>42 75.6361 2 6 9
brasileiro<>da<>43 74.7848 1 1 14	habitantes<>da<>43 74.7848 1 1 14
boca<>da<>43 74.7848 1 1 14	Deodoro<>da<>43 74.7848 1 1 14
da<>Professora<>43 74.7848 1 14 1	reis<>da<>43 74.7848 1 1 14
loi<>da<>43 74.7848 1 1 14	da<>civiliza<>43 74.7848 1 14 1
jugo<>da<>43 74.7848 1 1 14	da<>Fonseca<>43 74.7848 1 14 1
passagem<>da<>43 74.7848 1 1 14	da<>noite<>43 74.7848 1 14 1
da<>regi<>43 74.7848 1 14 1	da<>terra<>43 74.7848 1 14 1
Josefa<>dos<>44 74.3579 1 2 7	Maria<>!<>44 74.3579 1 2 7
ali<>n<>44 74.3579 1 2 7	!<>E<>44 74.3579 1 7 2
!<>Se<>44 74.3579 1 7 2	!<>Vital<>44 74.3579 1 7 2

da<>monarquia<>45 73.0032 2 14 4	que<>ele<>46 71.9290 2 19 3
aqui<>na<>47 69.2022 1 3 5	as<>dele<>48 64.8762 1 8 2
m<>minha<>49 64.7529 1 4 4	minha<>av<>49 64.7529 1 4 4
era<>com<>50 64.3284 2 9 7	mas<>era<>51 57.5017 1 2 9
Conselheiro<>mesmo<>51 57.5017 1 9 2	militar<>do<>51 57.5017 1 2 9
dio<>para<>51 57.5017 1 2 9	era<>bonita<>51 57.5017 1 9 2
para<>mulheres<>51 57.5017 1 9 2	para<>fazer<>51 57.5017 1 9 2
gente<>para<>51 57.5017 1 2 9	bom<>para<>51 57.5017 1 2 9
do<>nordeste<>51 57.5017 1 9 2	era<>bom<>51 57.5017 1 9 2
do<>dia<>51 57.5017 1 9 2	a<>rep<>52 56.3895 2 24 3
que<>pareciam<>53 54.8412 1 19 1	que<>fosse<>53 54.8412 1 19 1
achando<>que<>53 54.8412 1 1 19	rio<>que<>53 54.8412 1 1 19
que<>mandou<>53 54.8412 1 19 1	que<>destru<>53 54.8412 1 19 1
data<>que<>53 54.8412 1 1 19	dizer<>que<>53 54.8412 1 1 19
Disseram<>que<>53 54.8412 1 1 19	deixa<>que<>53 54.8412 1 1 19
que<>v<>53 54.8412 1 19 1	Jo<>o<>54 53.9384 3 3 56
de<>Jesus<>55 52.4580 2 39 2	l<>em<>56 51.6021 1 2 10
na<>monarquia<>57 51.4548 1 5 4	ao<>governo<>57 51.4548 1 5 4
.<>Ele<>58 51.0930 5 78 6	.<>O<>58 51.0930 5 78 6
tempo<>dos<>59 48.9569 1 3 7	um<>ano<>60 46.7753 1 11 2
de<>Canudos<>61 44.9657 3 39 5	a<>revolta<>62 43.2074 1 24 1
perante<>a<>62 43.2074 1 1 24	dariam<>a<>62 43.2074 1 1 24
toda<>a<>62 43.2074 1 1 24	a<>Guerra<>62 43.2074 1 24 1
a<>ficar<>62 43.2074 1 24 1	a<>morte<>62 43.2074 1 24 1
aram<>a<>62 43.2074 1 1 24	a<>promessa<>62 43.2074 1 24 1
destru<>a<>62 43.2074 1 1 24	a<>lei<>62 43.2074 1 24 1
a<>rua<>62 43.2074 1 24 1	a<>Campanha<>62 43.2074 1 24 1

a<>igreja<>62 43.2074 1 24 1	mo<>a<>62 43.2074 1 1 24
compreender<>a<>62 43.2074 1 1 24	proclamou<>a<>62 43.2074 1 1 24
trazia<>a<>62 43.2074 1 1 24	sobre<>a<>62 43.2074 1 1 24
a<>crescer<>62 43.2074 1 24 1	a<>grande<>62 43.2074 1 24 1
Continuaram<>a<>62 43.2074 1 1 24	a<>avisaram<>62 43.2074 1 24 1
a<>hist<>62 43.2074 1 24 1	a<>derrota<>62 43.2074 1 24 1
se<>por<>63 42.7531 1 12 2	se<>l<>63 42.7531 1 12 2
se<>casou<>63 42.7531 1 12 2	casou<>se<>63 42.7531 1 2 12
:<>uma<>64 42.5902 1 4 6	o<>governo<>65 39.0004 3 56 4
povo<>do<>66 37.7107 1 3 9	Santos<>,<>67 37.5268 3 4 58
e<>come<>68 36.8919 1 28 1	e<>munidos<>68 36.8919 1 28 1
aumentando<>e<>68 36.8919 1 1 28	e<>sangrentos<>68 36.8919 1 28 1
noite<>e<>68 36.8919 1 1 28	leis<>e<>68 36.8919 1 1 28
e<>armas<>68 36.8919 1 28 1	e<>engenheiro<>68 36.8919 1 28 1
armados<>e<>68 36.8919 1 1 28	impressionantes<>e<>68 36.8919 1 1 28
e<>ela<>68 36.8919 1 28 1	e<>institui<>68 36.8919 1 28 1
vir<>e<>68 36.8919 1 1 28	jornalista<>e<>68 36.8919 1 1 28
hinos<>e<>68 36.8919 1 1 28	Mam<>e<>68 36.8919 1 1 28
Homens<>e<>68 36.8919 1 1 28	granadas<>e<>68 36.8919 1 1 28
terra<>e<>68 36.8919 1 1 28	litoral<>e<>68 36.8919 1 1 28
e<>iguais<>68 36.8919 1 28 1	redor<>e<>68 36.8919 1 1 28
e<>tamb<>68 36.8919 1 28 1	e<>salvava<>68 36.8919 1 28 1
livres<>e<>68 36.8919 1 1 28	av<>n<>69 36.2589 1 4 7
o<>golpe<>70 35.9249 2 56 2	o<>acreditava<>70 35.9249 2 56 2
dele<>,<>71 34.6170 2 2 58	,<>mas<>71 34.6170 2 58 2
,<>no<>71 34.6170 2 58 2	de<>guerra<>72 33.6807 2 39 3
de<>Santo<>72 33.6807 2 39 3	as<>A<>73 31.5110 1 8 4

um<>Deus<>74 30.5550 1 11 3	um<>povo<>74 30.5550 1 11 3
que<>n<>75 28.7059 2 19 7	s<>!<>76 28.6421 1 5 7
monarquia<>para<>77 27.8187 1 4 9	governo<>do<>77 27.8187 1 4 9
blica<>uma<>78 27.7921 1 6 6	.<>A<>79 26.9485 3 78 4
aquele<>que<>80 26.4538 1 2 19	como<>que<>80 26.4538 1 2 19
mulheres<>que<>80 26.4538 1 2 19	de<>parto<>81 26.2042 1 39 1
sangrentos<>de<>81 26.2042 1 1 39	apenas<>de<>81 26.2042 1 1 39
de<>madrugada<>81 26.2042 1 39 1	de<>paus<>81 26.2042 1 39 1
nunca<>de<>81 26.2042 1 1 39	enviado<>de<>81 26.2042 1 1 39
casa<>de<>81 26.2042 1 1 39	de<>desvalidos<>81 26.2042 1 39 1
profeta<>de<>81 26.2042 1 1 39	tal<>de<>81 26.2042 1 1 39
de<>Marechal<>81 26.2042 1 39 1	de<>escolher<>81 26.2042 1 39 1
Campanha<>de<>81 26.2042 1 1 39	Estado<>de<>81 26.2042 1 1 39
de<>S<>81 26.2042 1 39 1	nasceu<>de<>81 26.2042 1 1 39
poca<>de<>81 26.2042 1 1 39	de<>Joaquim<>81 26.2042 1 39 1
Rio<>de<>81 26.2042 1 1 39	igreja<>de<>81 26.2042 1 1 39
Reginaldo<>de<>81 26.2042 1 1 39	Guerra<>de<>81 26.2042 1 1 39
grupo<>de<>81 26.2042 1 1 39	de<>pedir<>81 26.2042 1 39 1
de<>Monte<>81 26.2042 1 39 1	de<>toda<>81 26.2042 1 39 1
de<>Mattos<>81 26.2042 1 39 1	caso<>de<>81 26.2042 1 1 39
de<>R<>81 26.2042 1 39 1	promessa<>de<>81 26.2042 1 1 39
de<>Janeiro<>81 26.2042 1 39 1	trezena<>de<>81 26.2042 1 1 39
de<>honra<>81 26.2042 1 39 1	de<>feij<>81 26.2042 1 39 1
Batista<>de<>81 26.2042 1 1 39	derrota<>de<>81 26.2042 1 1 39
de<>novembro<>81 26.2042 1 39 1	de<>reza<>81 26.2042 1 39 1
15<>de<>81 26.2042 1 1 39	de<>vinte<>81 26.2042 1 39 1
comedor<>de<>81 26.2042 1 1 39	odo<>de<>81 26.2042 1 1 39

tico<>de<>81 26.2042 1 1 39	romeiro<>de<>81 26.2042 1 1 39
carregamento<>de<>81 26.2042 1 1 39	,<>que<>82 25.4945 6 58 19
.<>Mas<>83 25.2271 2 78 2	.<>Para<>83 25.2271 2 78 2
.<>Quando<>83 25.2271 2 78 2	.<>Uma<>83 25.2271 2 78 2
.<>Minha<>83 25.2271 2 78 2	.<>Foi<>83 25.2271 2 78 2
.<>No<>83 25.2271 2 78 2	rios<>.<>84 24.8824 2 2 79
Jesus<>.<>84 24.8824 2 2 79	nacional<>.<>84 24.8824 2 2 79
da<>guerra<>85 23.6562 1 14 3	da<>rep<>85 23.6562 1 14 3
!<>Ele<>86 23.5659 1 7 6	Canudos<>para<>87 21.8862 1 5 9
do<>s<>87 21.8862 1 9 5	ao<>Conselheiro<>87 21.8862 1 5 9
Conselheiro<>foi<>87 21.8862 1 9 5	.<>.<>88 20.8205 16 78 79
a<>nova<>89 20.6338 1 24 2	a<>todos<>89 20.6338 1 24 2
at<>a<>89 20.6338 1 2 24	fazer<>a<>89 20.6338 1 2 24
av<>se<>90 20.4362 1 4 12	em<>Canudos<>91 19.5226 1 10 5
rejei<>o<>92 17.9455 1 1 56	representou<>o<>92 17.9455 1 1 56
o<>branca<>92 17.9455 1 56 1	cobrir<>o<>92 17.9455 1 1 56
o<>At<>92 17.9455 1 56 1	v<>o<>92 17.9455 1 1 56
S<>o<>92 17.9455 1 1 56	o<>primeiro<>92 17.9455 1 56 1
o<>jornal<>92 17.9455 1 56 1	mandou<>o<>92 17.9455 1 1 56
o<>imprevisto<>92 17.9455 1 56 1	liberta<>o<>92 17.9455 1 1 56
o<>pa<>92 17.9455 1 56 1	fora<>o<>92 17.9455 1 1 56
o<>Siqueira<>92 17.9455 1 56 1	feij<>o<>92 17.9455 1 1 56
o<>of<>92 17.9455 1 56 1	organiza<>o<>92 17.9455 1 1 56
civiliza<>o<>92 17.9455 1 1 56	o<>ter<>92 17.9455 1 56 1
sob<>o<>92 17.9455 1 1 56	o<>sonho<>92 17.9455 1 56 1
o<>direito<>92 17.9455 1 56 1	ignoravam<>o<>92 17.9455 1 1 56
ter<>o<>92 17.9455 1 1 56	agita<>o<>92 17.9455 1 1 56

quest<>o<>92 17.9455 1 1 56	o<>convidaram<>92 17.9455 1 56 1
o<>vai<>92 17.9455 1 56 1	o<>jugo<>92 17.9455 1 56 1
o<>Reginaldo<>92 17.9455 1 56 1	o<>Ex<>92 17.9455 1 56 1
o<>viesses<>92 17.9455 1 56 1	o<>trabalho<>92 17.9455 1 56 1
regi<>o<>92 17.9455 1 1 56	At<>o<>92 17.9455 1 1 56
tradu<>o<>92 17.9455 1 1 56	o<>faltava<>92 17.9455 1 56 1
quando<>o<>92 17.9455 1 1 56	assombrava<>o<>92 17.9455 1 1 56
N<>o<>92 17.9455 1 1 56	o<>Paulo<>92 17.9455 1 56 1
o<>evento<>92 17.9455 1 56 1	rezasse<>o<>92 17.9455 1 1 56
fosse<>o<>92 17.9455 1 1 56	o<>ex<>92 17.9455 1 56 1
o<>vou<>92 17.9455 1 56 1	foi<>um<>93 17.5893 1 5 11
dio<>e<>94 17.4751 1 2 28	e<>fazer<>94 17.4751 1 28 2
pedras<>e<>94 17.4751 1 2 28	e<>mulheres<>94 17.4751 1 28 2
e<>canh<>94 17.4751 1 28 2	rcito<>e<>94 17.4751 1 2 28
de<>um<>95 17.4594 3 39 11	,<>Joana<>96 17.2922 1 58 1
1897<>,<>96 17.2922 1 1 58	Bittencourt<>,<>96 17.2922 1 1 58
entanto<>,<>96 17.2922 1 1 58	rua<>,<>96 17.2922 1 1 58
Mattos<>,<>96 17.2922 1 1 58	,<>romeiro<>96 17.2922 1 58 1
mulatos<>,<>96 17.2922 1 1 58	Janeiro<>,<>96 17.2922 1 1 58
paus<>,<>96 17.2922 1 1 58	1896<>,<>96 17.2922 1 1 58
Rosada<>,<>96 17.2922 1 1 58	seguinte<>,<>96 17.2922 1 1 58
ndios<>,<>96 17.2922 1 1 58	,<>granadas<>96 17.2922 1 58 1
radicais<>,<>96 17.2922 1 1 58	filho<>,<>96 17.2922 1 1 58
,<>numa<>96 17.2922 1 58 1	tardinha<>,<>96 17.2922 1 1 58
Essa<>,<>96 17.2922 1 1 58	,<>vivendo<>96 17.2922 1 58 1
,<>abandonada<>96 17.2922 1 58 1	,<>loi<>96 17.2922 1 58 1
ir<>,<>96 17.2922 1 1 58	metralhadoras<>,<>96 17.2922 1 1 58



,<>achando<>96 17.2922 1 1 58 1	contudente<>,<>96 17.2922 1 1 58
novembro<>,<>96 17.2922 1 1 58	obras<>,<>96 17.2922 1 1 58
XIX<>,<>96 17.2922 1 1 58	,<>D<>96 17.2922 1 58 1
,<>proclamou<>96 17.2922 1 58 1	evento<>,<>96 17.2922 1 1 58
revoltosos<>,<>96 17.2922 1 1 58	dirigentes<>,<>96 17.2922 1 1 58
,<>caboclos<>96 17.2922 1 58 1	,<>pretos<>96 17.2922 1 58 1
dono<>,<>96 17.2922 1 1 58	,<>elas<>96 17.2922 1 58 1
,<>todo<>96 17.2922 1 58 1	,<>mulatos<>96 17.2922 1 58 1
Fonseca<>,<>96 17.2922 1 1 58	morena<>,<>96 17.2922 1 1 58
,<>acompanhando<>96 17.2922 1 58 1	,<>liderado<>96 17.2922 1 58 1
tona<>,<>96 17.2922 1 1 58	,<>tardinha<>96 17.2922 1 58 1
caboclos<>,<>96 17.2922 1 1 58	governo<>da<>97 17.2739 1 4 14
com<>as<>98 17.2243 1 7 8	que<>eram<>99 17.0023 1 19 3
ele<>que<>99 17.0023 1 3 19	povo<>que<>99 17.0023 1 3 19
que<>Deus<>99 17.0023 1 19 3	para<>a<>100 16.3389 2 9 24
o<>:<>101 16.0449 2 56 4	.<>Jo<>102 15.5234 2 78 3
Conselheiro<>!<>103 15.1121 1 9 7	de<>uma<>104 14.9729 2 39 6
cio<>um<>105 14.3522 1 6 11	contra<>a<>106 13.1236 1 3 24
.<>Disseram<>107 12.6016 1 78 1	.<>Meu<>107 12.6016 1 78 1
.<>Elas<>107 12.6016 1 78 1	.<>J<>107 12.6016 1 78 1
.<>Estavam<>107 12.6016 1 78 1	.<>As<>107 12.6016 1 78 1
.<>Papai<>107 12.6016 1 78 1	.<>Continuaram<>107 12.6016 1 78 1
.<>Eram<>107 12.6016 1 78 1	.<>Perguntou<>107 12.6016 1 78 1
.<>Tinha<>107 12.6016 1 78 1	.<>Quatro<>107 12.6016 1 78 1
.<>poca<>107 12.6016 1 78 1	.<>Veio<>107 12.6016 1 78 1
.<>sertanejos<>107 12.6016 1 78 1	.<>N<>107 12.6016 1 78 1
.<>Mam<>107 12.6016 1 78 1	.<>tudo<>107 12.6016 1 78 1

.<>Cantava<>107 12.6016 1 78 1	.<>Sei<>107 12.6016 1 78 1
.<>Ave<>107 12.6016 1 78 1	.<>Zefinha<>107 12.6016 1 78 1
.<>Ela<>107 12.6016 1 78 1	.<>Um<>107 12.6016 1 78 1
pretos<>.<>108 12.4294 1 1 79	imprevisto<>.<>108 12.4294 1 1 79
cantavam<>.<>108 12.4294 1 1 79	despachavam<>.<>108 12.4294 1 1 79
1911<>.<>108 12.4294 1 1 79	Macambira<>.<>108 12.4294 1 1 79
gios<>.<>108 12.4294 1 1 79	dinheiro<>.<>108 12.4294 1 1 79
desvalidos<>.<>108 12.4294 1 1 79	isso<>.<>108 12.4294 1 1 79
ordem<>.<>108 12.4294 1 1 79	monarquista<>.<>108 12.4294 1 1 79
respondia<>.<>108 12.4294 1 1 79	cheia<>.<>108 12.4294 1 1 79
madrugada<>.<>108 12.4294 1 1 79	europalizada<>.<>108 12.4294 1 1 79
lei<>.<>108 12.4294 1 1 79	reza<>.<>108 12.4294 1 1 79
madeiras<>.<>108 12.4294 1 1 79	sticas<>.<>108 12.4294 1 1 79
1936<>.<>108 12.4294 1 1 79	milagre<>.<>108 12.4294 1 1 79
feitas<>.<>108 12.4294 1 1 79	salvava<>.<>108 12.4294 1 1 79
1902<>.<>108 12.4294 1 1 79	1909<>.<>108 12.4294 1 1 79
batizou<>.<>108 12.4294 1 1 79	parto<>.<>108 12.4294 1 1 79
batalhas<>.<>108 12.4294 1 1 79	1919<>.<>108 12.4294 1 1 79
voz<>.<>108 12.4294 1 1 79	Paulo<>.<>108 12.4294 1 1 79
D<>.<>108 12.4294 1 1 79	que<>minha<>109 12.2848 1 19 4
que<>tinha<>109 12.2848 1 19 4	muito<>de<>110 12.1317 1 2 39
de<>homens<>110 12.1317 1 39 2	de<>pedras<>110 12.1317 1 39 2
Maria<>de<>110 12.1317 1 2 39	rcito<>de<>110 12.1317 1 2 39
de<>mudan<>110 12.1317 1 39 2	at<>de<>110 12.1317 1 2 39
de<>muita<>110 12.1317 1 39 2	de<>l<>110 12.1317 1 39 2
mais<>de<>110 12.1317 1 2 39	o<>sert<>111 12.1003 2 56 5
com<>um<>112 12.0425 1 7 11	um<>dos<>112 12.0425 1 11 7

,<>em<>113 11.7427 3 58 10	e<>eram<>114 11.0197 1 28 3
guerra<>e<>114 11.0197 1 3 28	cio<>da<>115 10.9032 1 6 14
se<>com<>116 10.8927 1 12 7	monarquia<>.<>117 10.5380 2 4 79
que<>foi<>118 9.4609 1 19 5	:<>a<>119 9.3792 1 4 24
a<>monarquia<>119 9.3792 1 24 4	blica<>,<>120 9.0562 2 6 58
es<>,<>120 9.0562 2 6 58	era<>um<>121 8.9689 1 9 11
se<>para<>122 8.0759 1 12 9	o<>queria<>123 8.0073 1 56 2
canh<>o<>123 8.0073 1 2 56	o<>nacional<>123 8.0073 1 56 2
o<>canh<>123 8.0073 1 56 2	o<>in<>123 8.0073 1 56 2
o<>nordeste<>123 8.0073 1 56 2	Para<>o<>123 8.0073 1 2 56
todos<>o<>123 8.0073 1 2 56	av<>e<>124 7.8048 1 4 28
e<>tinha<>124 7.8048 1 28 4	e<>minha<>124 7.8048 1 28 4
Canudos<>.<>125 7.7155 2 5 79	dia<>,<>126 7.6814 1 2 58
,<>bom<>126 7.6814 1 58 2	mulheres<>,<>126 7.6814 1 2 58
,<>foram<>126 7.6814 1 58 2	Bahia<>,<>126 7.6814 1 2 58
,<>mesmo<>126 7.6814 1 58 2	,<>queria<>126 7.6814 1 58 2
gente<>,<>126 7.6814 1 2 58	milagres<>,<>126 7.6814 1 2 58
,<>todos<>126 7.6814 1 58 2	pedras<>,<>126 7.6814 1 2 58
bonita<>,<>126 7.6814 1 2 58	,<>homens<>126 7.6814 1 58 2
,<>como<>126 7.6814 1 58 2	Cunha<>,<>126 7.6814 1 2 58
,<>pedras<>126 7.6814 1 58 2	teve<>de<>127 7.4652 1 3 39
de<>Deus<>127 7.4652 1 39 3	eram<>de<>127 7.4652 1 3 39
,<>Ant<>128 5.9433 2 58 8	blica<>.<>129 5.8597 2 6 79
.<>Se<>130 5.3447 1 78 2	.<>Vital<>130 5.3447 1 78 2
.<>Josefa<>130 5.3447 1 78 2	.<>Euclides<>130 5.3447 1 78 2
.<>E<>130 5.3447 1 78 2	l<>.<>131 5.2590 1 2 79
bonita<>.<>131 5.2590 1 2 79	mesmo<>.<>131 5.2590 1 2 79

ali<>.<>131 5.2590 1 2 79	casou<>.<>131 5.2590 1 2 79
para<>o<>132 5.2050 2 9 56	,<>era<>133 4.9240 2 58 9
o<>tempo<>134 4.7306 1 56 3	contra<>o<>134 4.7306 1 3 56
o<>povo<>134 4.7306 1 56 3	teve<>o<>134 4.7306 1 3 56
blica<>e<>135 4.6154 1 6 28	com<>a<>136 4.6019 1 7 24
,<>Jo<>137 4.5152 1 58 3	guerra<>,<>137 4.5152 1 3 58
,<>ele<>137 4.5152 1 58 3	,<>teve<>137 4.5152 1 58 3
aqui<>,<>137 4.5152 1 3 58	Deus<>,<>137 4.5152 1 3 58
e<>dos<>138 3.7151 1 28 7	que<>um<>139 3.3635 1 19 11
A<>o<>140 3.1194 1 4 56	o<>tinha<>140 3.1194 1 56 4
e<>as<>141 3.0463 1 28 8	.<>Os<>142 2.9776 1 78 3
governo<>,<>143 2.9604 1 4 58	tinha<>,<>143 2.9604 1 4 58
monarquia<>,<>143 2.9604 1 4 58	que<>se<>144 2.9499 1 19 12
povo<>.<>145 2.9215 1 3 79	Deus<>.<>145 2.9215 1 3 79
guerra<>.<>145 2.9215 1 3 79	cio<>de<>146 2.8720 1 6 39
Conselheiro<>.<>147 2.8707 2 9 79	e<>para<>148 2.5319 1 28 9
e<>do<>148 2.5319 1 28 9	Conselheiro<>e<>148 2.5319 1 9 28
o<>na<>149 2.1745 1 56 5	o<>s<>149 2.1745 1 56 5
,<>ao<>150 2.0501 1 58 5	s<>,<>150 2.0501 1 5 58
,<>foi<>150 2.0501 1 58 5	,<>na<>150 2.0501 1 58 5
o<>de<>151 2.0014 4 56 39	Santos<>.<>152 1.7924 1 4 79
governo<>.<>152 1.7924 1 4 79	av<>.<>152 1.7924 1 4 79
e<>o<>153 1.6954 3 28 56	o<>,<>154 1.5530 1 56 58
e<>se<>155 1.5290 1 28 12	cio<>,<>156 1.4621 1 6 58
o<>.<>157 1.2914 2 56 79	.<>Canudos<>158 1.1777 1 78 5
com<>o<>159 1.1413 1 7 56	o<>!<>159 1.1413 1 56 7
o<>dos<>159 1.1413 1 56 7	e<>da<>160 1.1178 1 28 14

,<n>161 1.0584 1 58 7	nio<>,<>162 0.7698 1 8 58
,<as>162 0.7698 1 58 8	es<>.<>163 0.7427 1 6 79
e<>que<>164 0.5170 1 28 19	!<>.<>165 0.4770 1 7 79
.<Ant>166 0.3126 1 78 8	nio<>.<>167 0.2977 1 8 79
,<um>168 0.2815 1 58 11	o<>se<>169 0.2257 1 56 12
a<>e<>170 0.2221 1 24 28	e<>a<>170 0.2221 1 28 24
e<>,<>171 0.2008 1 28 58	.<era>172 0.1875 1 78 9
,<e>173 0.1553 2 58 28	o<>da<>174 0.0981 1 56 14
,<a>175 0.0809 1 58 24	a<>o<>176 0.0612 1 24 56
,<o>177 0.0015 3 58 56	o<>que<>178 0.0000 1 56 19
que<>o<>178 0.0000 1 19 56	

Após a execução do comando '*rank.pl dice qui-quadrado*' obtivemos o seguinte resultado:

Rank correlation coefficient = 0.9430



## **ANEXO VI – FERRAMENTAS ESPECIALIZADAS**





**EXEMPLO 2.105 - GRAMÁTICA COM QUATRO REGRAS BASEADAS EM DEPENDÊNCIAS**

AdjnR: NOUN ADJ

Agr: number, genre

SpecL: DT NOUN

Agr: number, genre

SubjL: NOUN [ADV]\* VERB

Agr: number

DobjR: VERB [ADV]\* NOUN

**EXEMPLO 2.106 - REPRESENTAÇÃO SIMPLES BASEADA EM DEPENDÊNCIAS DE UMA****FRASE**

Considere-se um ficheiro no formato de texto plano como o seguinte:

test.txt
Eu tenho um sonho.

Ao executarmos o comando '*dp.sh -a treetagger pt test.txt > output.txt*' obtemos o seguinte resultado:

```
SENT::Eu_PRO_0_<case:0|gender:0|lemma:eu|number:0|person:0|politeness:0|pos:0|p
ossessor:0|token:Eu|type:0|>
```

```
tenho_VERB_1_<gender:0|lemma:ter|mode:0|number:0|person:0|pos:1|tense:0|token
:tenho|type:0|>
```

```
um_DT_2_<gender:0|lemma:um|number:0|person:0|pos:2|possessor:0|token:um|type
:0|>
```

```
sonho_NOUN_3_<gender:0|lemma:sonho|number:0|person:3|pos:3|token:sonho|type:
C|> ._SENT_4_<lemma:.|pos:4|token:.|>
```

```
(SubjL;ter_VERB_1;eu_PRO_0)
```

```
(SpecL;sonho_NOUN_3;um_DT_2)
```

```
(DobjR;ter_VERB_1;sonho_NOUN_3)
```

Como se pode observar, o *output* obtido consiste em quatro linhas:

- a primeira linha corresponde ao input da estratégia de *parsing* (Inicia com 'SENT::' e contém 5 colunas separadas por espaços. Em cada coluna encontra-se a informação morfo-sintáctica associada a cada *token* do texto dado como *input*.);
- as restantes linhas mostram a análise baseada em dependências (por exemplo, a segunda linha – SubjL;ter\_VERB\_1;eu\_PRO\_0 – representa a dependência SubjL existente entre o pronome 'eu' e o verbo 'ter').

## EXEMPLO 2.107 - ANÁLISE DE DISCURSO

Considere-se o seguinte texto:

'Os resultados da análise de inteligibilidade de córpis podem ajudar a guiar a tarefa de simplificação textual,  
fornecendo quais características realmente tornam um texto mais simples de ser entendido por pessoas dos mais diversos níveis de letramento.[s]  
Apesar de ter sido criada para este fim,  
a ferramenta pode ser utilizada  
para quaisquer fins que necessitem de tais informações.[s][p]'

Após o processamento manual do *input* anteriormente referido, obtiveram-se os seguintes resultados:

- segmentos
  - 1: Os resultados de a análise de inteligibilidade de córpis podem ajudar a guiar a tarefa de simplificação(simplificação)\_n textual(textual)\_adj, ()\_
  - 2: fornecendo quais características realmente tornam um texto mais simples de ser entendido por pessoas de os mais diversos níveis de letramento(letramento)\_n. ()\_
  - 3: Apesar de ter sido criada para este fim(fim)\_n, ()\_
  - 4: a ferramenta pode ser utilizada(utilizar)\_v-pcp. ()\_
  - 5: para quaisquer fins que necessitem de tais informações(informação)\_n. ()\_

- padrões encontrados

Relation: concession | Marker 1:

$^{\wedge}\text{Apesar}\backslash\left([\wedge]\right)^{\wedge}\_[\wedge]^{\wedge}$  | Marker 2:  $.\text{*}\backslash\left([\wedge]\right)^{\wedge}\_[\wedge]^{\wedge}$

Relation: contrast | Marker 1:

$.\text{*}\backslash\left([\wedge]\right)^{\wedge}\_[\wedge]^{\wedge}$  | Marker 2:  $\text{apesar}\backslash\left([\wedge]\right)^{\wedge}\_[\wedge]^{\wedge}$

- relações identificadas

concession,contrast,elaboration,elaboration

hypotactic,paratactic,hypotactic,hypotactic

4,3,2,5

4,5,2,5

3,1,1,4

3,2,1,4

5

- gramática gerada

$s(1,1,\text{leaf},1,1)$ .

$s(2,2,\text{leaf},2,1)$ .

$s(3,3,\text{leaf},3,1)$ .

$s(4,4,\text{leaf},4,1)$ .

$s(5,5,\text{leaf},5,1)$ .

$s(1,2,\text{elaboration},\text{elaboration}(n(N),s(S)),P) \quad :- \quad s(1,1,\text{leaf},N,P1), \quad s(2,2,\text{leaf},S,P2),$   
 $p(\text{elaboration},\text{leaf},n,\text{leaf},s,P3), \quad P \text{ is } P1*P2*P3.$

$s(3,4,\text{concession},\text{concession}(n(N),s(S)),P) \quad :- \quad s(3,3,\text{leaf},N,P1), \quad s(4,4,\text{leaf},S,P2),$   
 $p(\text{concession},\text{leaf},n,\text{leaf},s,P3), \quad P \text{ is } P1*P2*P3.$

$s(4,5,\text{elaboration},\text{elaboration}(n(N),s(S)),P) \quad :- \quad s(4,4,\text{leaf},N,P1), \quad s(5,5,\text{leaf},S,P2),$   
 $p(\text{elaboration},\text{leaf},n,\text{leaf},s,P3), \quad P \text{ is } P1*P2*P3.$

$s(3,5,\text{concession},\text{concession}(n(N),s(S)),P) \quad :- \quad s(3,3,\text{leaf},N,P1), \quad s(4,5,\text{elaboration},S,P2),$   
 $p(\text{concession},\text{leaf},n,\text{elaboration},s,P3), \quad P \text{ is } P1*P2*P3.$

$s(3,5,\text{elaboration},\text{elaboration}(n(N),s(S)),P) \quad :- \quad s(3,4,\text{concession},N,P1), \quad s(5,5,\text{leaf},S,P2),$   
 $p(\text{elaboration},\text{concession},n,\text{leaf},s,P3), \quad P \text{ is } P1*P2*P3.$

- tempos de execução

tagging2.pl 15

Identificador de relações - relate.pl 23

Geração da Gramática - grammar.pl 0

Execução da gramática - execute.pl 14

## EXEMPLO 2.108 - ETIQUETAÇÃO DE UM TEXTO

Suponhamos que introduzimos no *site* [http://label.ist.utl.pt/pt/eelo\\_online\\_pt.php](http://label.ist.utl.pt/pt/eelo_online_pt.php) o texto especificado no exemplo 2.107.

O resultado obtido é o seguinte:

Os.eu.PronPess:L4mp[o]:L3mp[o]

Os.o.DetArtDef:mp

Os.o.PronDem:mp

resultados.resultado.Nome:mp

da.do.Preposition & DetArtDef:fs

da.do.Preposition & PronDem:fs

análise.análise.Nome:fs

de.de.Preposition

inteligibilidade.inteligibilidade.Nome:fs

de.de.Preposition

cópus

podem.podar.Verbo:S4p[i]:S3p[i]:Y4p[icqny]:4p[ix]

podem.poder.Verbo+Vaux:P4p[icqn]:P3p[icqn]

ajudar.ajudar.Verbo:Z4s[icq]:Z3s[icq]:Z1s[icq]:W[icq]:U4s[i]:U3s[i]:U1s[i]

a.a.Nome+Letra:ms

a.a.Preposition

a.eu.PronPess:L4fs[o]:L3fs[o]

a.o.DetArtDef:fs

a.o.PronDem:fs

guiar.guiar.Verbo:Z4s[icq]:Z3s[icq]:Z1s[icq]:W[icq]:U4s[i]:U3s[i]:U1s[i]

a.a.Nome+Letra:ms

a.a.Prep  
a.eu.PronPess:L4fs[o]:L3fs[o]  
a.o.DetArtDef:fs  
a.o.PronDem:fs  
tarefa.tarefa.Nome:fs  
tarefa.tarefar.Verbo:P4s[icqo]:P3s[icqo]:Y2s[icqoy]  
de.de.Prep  
simplificação.simplificação.Nome:fs  
textual.textual.Adj+Rel:ms:fs  
fornecendo.fornecer.Verbo:G[icqo]  
quais.qual.PronExcl:mp:fp  
quais.qual.PronInter:mp:fp  
quais.qual.PronRel:mp:fp  
características.característica.Nome:fp  
características.característico.Adj+Pd:fp  
realmente.realmente.Adv  
tornam.tornar.Verbo:P4p[icq]:P3p[icq]  
um.um.DetNum+Val=1:Cms  
um.um.DetArtIndef:ms  
texto.texto.Nome:ms  
mais.mais.Adv  
mais.mais.Nome:ms:mp  
simples.simples.Adj+Det:ms:fs:mp:fp  
simples.simples.Adj+Pd:ms:fs:mp:fp  
simples.simples.Nome:ms:fs:mp:fp  
de.de.Prep  
ser.ser.Nome:ms  
ser.ser.Verbo+Vaux:Z4s[icq]:Z3s[icq]:Z1s[icq]:W[icq]  
entendido.entendido.Adj+Pd:ms  
entendido.entendido.Nome:ms  
entendido.entender.Verbo:K[i]  
por.por.Prep  
pessoas.pessoa.Nome:fp  
dos.do.Prep & DetArtDef:mp

dos.do.Prep & PronDem:mp  
mais.mais.Adv  
mais.mais.Nome:ms:mp  
diversos.diverso.Adj+Det:mp  
diversos.diverso.Adj+Pd:mp  
níveis.nível.Nome:mp  
de.de.Prep  
letramento  
s.s.Nome+Letra:ms  
Apesar.apesar  
de.de.Prep  
ter.ter.Verbo+Vaux:Z4s[icq]:Z3s[icq]:Z1s[icq]:W[icq]  
sido.ser.Verbo+Vaux:K[i]  
criada.criado.Adj+Pd:fs  
criada.criado.Nome:fs  
para.para.Conj  
para.para.Prep  
para.parir.Verbo:Y4s[icqoy]:4s[ix]:S4s[i]:S3s[i]:S1s[i]  
este.este.DetDem:ms  
este.este.PronDem:ms  
fim.fim.Nome:ms  
a.a.Nome+Letra:ms  
a.a.Prep  
a.eu.PronPess:L4fs[o]:L3fs[o]  
a.o.DetArtDef:fs  
a.o.PronDem:fs  
ferramenta.ferramenta.Nome:fs  
pode.podar.Verbo:S4s[i]:S3s[i]:S1s[i]:Y4s[icqoy]:4s[ix]  
pode.poder.Verbo+Vaux:Y2s[icqoy]:P4s[icqo]:P3s[icqo]  
ser.ser.Nome:ms  
ser.ser.Verbo+Vaux:Z4s[icq]:Z3s[icq]:Z1s[icq]:W[icq]  
utilizada.utilizado.Adj+Pd:fs  
para.para.Conj  
para.para.Prep

para.parrir.Verbo:Y4s[icqoy]:4s[ix]:S4s[i]:S3s[i]:S1s[i]  
 quaisquer.qualquer.Det.+Ind:mp:fp  
 quaisquer.qualquer.PronIndef:mp:fp  
 fins.fim.Nome:mp  
 que.que.Conj  
 que.que.PronExcl  
 que.que.PronInter  
 que.que.PronRel  
 necessitem.necessitar.Verbo:S4p[i]:S3p[i]:Y4p[icqny]:4p[ix]  
 de.de.Prep  
 tais.tais.Nome:ms:mp  
 tais.tal.DetDem:mp:fp  
 tais.tal.Nome:mp  
 tais.tal.PronDem:mp:fp  
 informações.informação.Nome:fp  
 s.s.Nome+Letra:ms  
 p.p.Nome+Letra:ms

## EXEMPLO 2.109 - ANÁLISE ESTATÍSTICA DE CORPORA

```

#!/usr/bin/perl
use strict;
use Text::CStatiBR; &Text::CStatiBR::CSTATIBR("5");
#5 files are analysed.
#Main output
#file created is
#1 (10).txt
  
```

## EXEMPLO 2.110 - FUNCIONAMENTO DA FERRAMENTA LINGUA.SH EXISTENTE NO PACKAGE LINGUAToolKIT

Considere-se o seguinte texto contido num ficheiro com o nome 'teste.txt' sem qualquer formatação:

teste.txt
<p>Numa tarde de Agosto muito quente estava eu na estação de Nine a espera do comboio que me iria levar a viagem fantástica até Paris (França). O comboio chegou a hora prevista que estava marcada no bilhete. Entrei e sentei-me a meio da carruagem foi quase das primeiras pessoas a entrar. Passados cinco minutos, quando o comboio estava prestes a iniciar o seu percurso, uma mulher cega sentou-se a minha beira. Já nos tínhamos andando um bom bocado quando a senhora falou.</p> <p>- Boa tarde, como se chama?</p> <p>- Boa tarde, sou o Hélder e a Senhora?</p> <p>- Sua a Beatriz.</p> <p>- De onde é e para onde vai viajar?</p> <p>- Sou de Barcelos e vou para Paris.</p> <p>- Eu também sou de Barcelos e vou fazer uma viagem incrível a Paris.</p> <p>- Estamos a passar a fronteira, é uma região muito bonita. A sua cegueira é de nascença?</p> <p>- Quando nasci via, mas não era muito bem tinha dificuldades. Os anos foram passando e cada vez mais não conseguia ver. Como é que esta a carruagem? Tem muita gente?</p> <p>- A carruagem esta com muita gente. Mas não tinha cura?</p> <p>- Infelizmente não.</p> <p>Ao fim da Beatriz disser estas palavras tinha-mos chegado a estação de Paris e cada um seguiu a sua viagem.</p>

Ao executarmos o comando `lingua.sh treetagger pt teste.txt 5` obtemos os seguintes ficheiros:

- ✓ `parsed.txt` – que contém o texto analisado sintacticamente;
- ✓ `cooccur.txt` – que contém as co-ocorrências entre as palavras do texto dado como input;
- ✓ onze ficheiros que contêm as medidas de similaridade entre as palavras do texto de input(`__baseline.gz`, `__cosineBin.gz`, `__diceMin.gz`, `__jaccardMax.gz`, `__city.gz`, `__cosine.gz`, `__eucl.gz`, `__js.gz`, `__diceBin.gz`, `__jaccard.gz`, `__lin.gz`);



## EXEMPLO 2.111 - ANÁLISE DE UM TEXTO DADO COMO INPUT COM O MULTILINGUAL DEPENDENCY PARSER

Considere-se o texto:

teste.txt
Eu sou um homem.

Ao executarmos o comando `parser_ml.sh treetagger pt teste.txt` obtemos o seguinte resultado:

```
SENT::<eu_PN_0 ser_VERBFCOP_1 um_DT_2 homem_NOM_3 ._SENT>
(Lobj;ser_VERBF_1;eu_PN_0)
(Spec;homem_NOM_3;um_DT_2)
(Robj;ser_VERBF_1;homem_NOM_3)
```

## EXEMPLO 2.112 - ILUSTRAÇÃO DA UTILIZAÇÃO DO EXTRACTOR DE TERMOS MULTILÍNGUA GALEXTRA

Considere-se como *input* o texto do exemplo 2.110.

Ao executarmos o comando `extractorTermos.sh treetagger pt log teste.txt` obtemos o seguinte resultado:

NA: hora-prevista	NA: mulher-cega
NA: viagem-fantástica	NA: viagem-incrível
AN: Boa-tarde	AN: Sua-Beatriz
AN: bom-bocado	AN: cada-vez
AN: minha-beira	AN: muita-gente
AN: seu-percurso	AN: sua-cegueira
NPN: espera-do-comboio	NPN: estação-de-Paris
NPN: fim-da-Beatriz	NPN: meio-da-carruagem
NPN: tarde-de-Agosto	fim da Beatriz 2.74682509896297 N-P-N
tarde de Agosto2.74682509896297 N-P-N	meio da carruagem 2.74682509896297 N-P-N
estação de Paris N-P-N 2.74682509896297	Sua Beatriz A-N

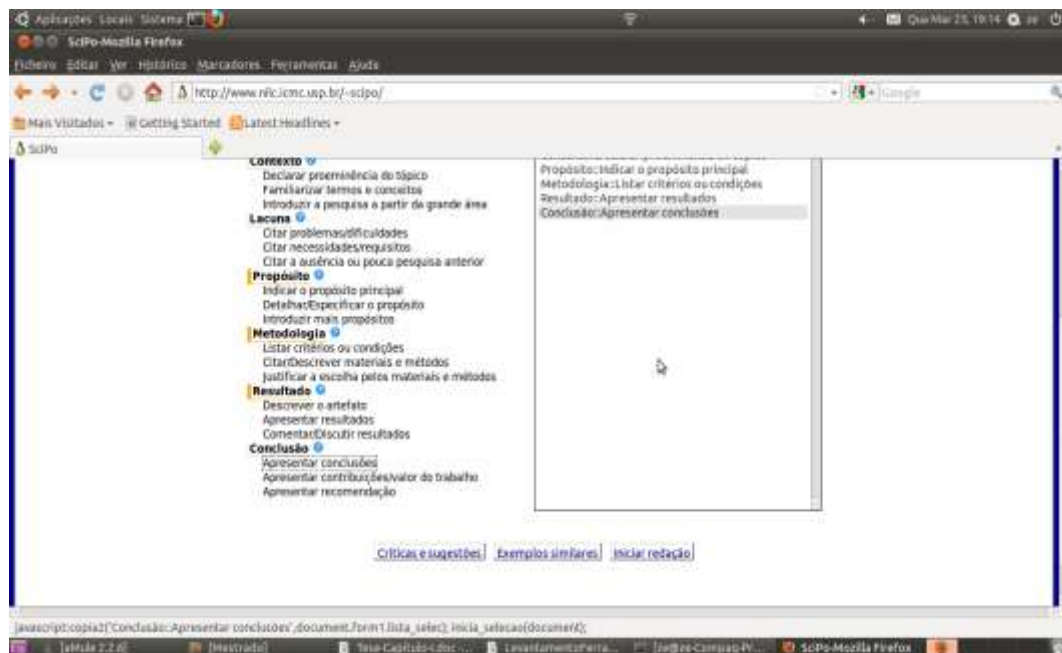
bom bocado	A-N	Boa tarde	A-N
sua cegueira	A-N	hora prevista	N-A
minha beira	A-N	espera do comboio	N-P-N
mulher cega	N-A	seu percurso	A-N
viagem fantástica	N-A	muita gente	A-N
cada vez	A-N	viagem incrível	N-A

### EXEMPLO 2.113 – UTILIZAÇÃO DA FERRAMENTA O CONSTRUCTOR

Ao introduzirmos como *input* a palavra *computador* obtemos a seguinte expressão:

*(computador[pos="N.\*"&pos!="NUM.\*"])*.

### EXEMPLO 2.114 – UTILIZAÇÃO DO SCIPO PARA A ESCRITA DE UM RESUMO



Após a declaração dos parâmetros obteve-se o seguinte resultado:

Atualmente as \_\_\_ devem \_\_\_ e a tendência é que estas \_\_\_ sejam \_\_\_ Para atingir esses objectivos, este trabalho propõe que \_\_\_ sejam \_\_\_ Os \_\_\_ desenvolvidos foram executados em diferentes \_\_\_ tanto em \_\_\_ como em \_\_\_ Indicações sobre \_\_\_ são fornecidas. O \_\_\_ permite o \_\_\_ usando o \_\_\_

### EXEMPLO 2.115 - IDENTIFICAÇÃO DA LÍNGUA NA QUAL UM DADO DOCUMENTO SE ENCONTRA ESCRITO

Se executarmos na linha de comandos o comando `./text_cat -l "Isto é uma experiência."` obtemos o seguinte resultado:

latin or portuguese or spanish or rumantsch or romanian or italian or catalan.

### EXEMPLO 2.116 - FUNCIONAMENTO DO CONCORDANCIADOR DO PROJECTO TEXTQUIM



Ao introduzirmos o texto do exemplo 2.110 e após a introdução da palavra 'carruagem', obtivemos os resultados mostrados em seguida:

**Concordanciador**  
**Total de concordâncias: 3**  
 e sentei-me a meio da **carruagem** foi quase das primeiras pessoas Como é que esta a **carruagem?** Tem muita gente? - A Tem muita gente? - A **carruagem** esta com muita gente. Mas

## EXEMPLO 2.117 – FUNCIONAMENTO DO LISTADOR DE PALAVRAS DO PACKAGE

### TEXTQUIM

Considerando o texto de *input* do exemplo 2.110, e pedindo a listagem por ordem alfabética o listador de palavras forneceu-nos o seguinte resultado:

Lista de palavras							
Tokens: 205							
Types: 124							
18	a	1	dificuldades	1	levar	1	primeiras
1	agosto	1	dossier	1	mais	3	quando
1	andando	1	do	1	marcada	1	quase
1	anos	7	e	2	mas	3	que
1	ao	4	é	1	me	1	quente
1	até	1	entrar	1	meio	1	região
2	barcelos	1	entrei	1	minha	1	se
2	beatriz	1	era	1	minutos	1	seguiu
1	beira	1	espera	2	muita	2	senhora
1	bem	2	esta	3	muito	1	sentei-me
1	bilhete	2	estação	1	mulher	1	sentou-se
2	boa	1	estamos	1	na	1	seu

1	bocado	1	estas	4	não	3	sou
1	bom	3	estava	1	nascença	3	sua
1	bonita	2	eu	1	nasci	1	também
2	cada	1	falou	1	nine	3	tarde
3	carruagem	1	fantástica	1	no	1	tem
1	cega	1	fazer	1	nos	2	tinha
1	cegueira	1	fim	1	numa	1	tinha-mos
1	chama	1	foi	4	o	1	tinhamos
1	chegado	1	foram	2	onde	2	um
1	chegou	1	frança	1	os	3	uma
1	cinco	1	fronteira	1	palavras	1	vai
1	com	2	gente	2	para	1	ver
3	comboio	1	hélter	4	paris	1	vez
2	como	1	hora	1	passados	1	via
1	conseguia	1	incrível	1	passando	3	viagem
1	cura	1	infelizmente	1	passar	1	viajar
2	da	1	iniciar	1	percurso	2	vou
1	das	1	iria	1	prestes	1	
7	de	1	já	1	prevista		

**EXEMPLO 2.118 - FUNCIONAMENTO DA FERRAMENTA N-GRAMAS**

Considerando o texto do exemplo 2.110 como *input* e pedindo a listagem 2-gramas o n-gramas forneceu-nos o seguinte resultado:

2	o comboio	1	incrível a	1	sua cegueira
2	a carruagem	1	iniciar o	1	seguiu a
2	a senhora	1	região muito	1	uma região
2	de barcelos	1	cegueira é	1	um seguiu
2	mas não	1	cada vez	1	já nos
2	muita gente	1	como se	1	uma viagem
2	e cada	1	cega sentou-se	1	um bom
2	e vou	1	carruagem foi	1	sua viagem
2	estação de	1	cinco minutos	1	uma mulher
2	sou de	1	comboio estava	1	paris e
2	a sua	1	conseguia ver	1	palavras tinha-mos
2	barcelos e	1	é que	1	bem tinha
2	boa tarde	1	cada um	1	hélter e
1	vou para	1	iria levar	1	chegou a
1	vou fazer	1	o seu	1	hora prevista
1	vai viajar	1	chegado a	1	os anos
1	infelizmente não	1	com muita	1	bom bocado
1	levar a	1	como é	1	onde vai
1	viagem fantástica	1	carruagem esta	1	prevista que
1	viagem incrível	1	é de	1	primeiras pessoas
1	vai viajar	1	chegado a	1	os anos
1	infelizmente não	1	com muita	1	bom bocado
1	levar a	1	como é	1	onde vai
1	viagem fantástica	1	carruagem esta	1	prevista que
1	viagem incrível	1	é de	1	primeiras pessoas
1	fazer uma	1	senhora falou	1	pessoas a
1	fantástica até	1	sentei-me a	1	passados cinco
1	é uma	1	seu percurso	1	o hélter
1	foi quase	1	sentou-se a	1	esta a
1	de paris	1	nine a	1	numa tarde
1	agosto muito	1	no bilhete	1	era muito
1	a espera	1	tarde de	1	eu também

1	disser estas	1	tínhamos andando	1	estamos a
1	da beatriz	1	tinha-mos chegado	1	quando nasci
1	das primeiras	1	de agosto	1	quando a
1	de nascença	1	tinha cura	1	meio da
1	a viagem	1	tem muita	1	mulher cega
1	até paris	1	também sou	1	me iria
1	andando um	1	tinha dificuldades	1	muito quente
1	a beatriz	1	a passar	1	anos foram
1	a paris	1	a fronteira	1	ao fim
1	a minha	1	entrei e	1	a estação
1	a iniciar	1	e sentei-me	1	minha beira
1	a hora	1	estava prestes	1	muito bonita
1	a meio	1	estava marcada	1	que estava
1	a entrar	1	espera do	1	quase das
1	da carruagem	1	que esta	1	quando o
1	do comboio	1	estava eu	1	que me
1	não era	1	eu na	1	quente estava
1	não conseguia	1	e a	1	muito bem
1	não tinha	1	e para	1	mais não
1	nasci via	1	esta com	1	comboio chegou
1	nos tínhamos	1	estas palavras	1	
1	na estação	1	comboio que	1	

**EXEMPLO 2.119 - FUNCIONAMENTO DA FERRAMENTA CONCORDANCIADOR ALINHADO**





## **ANEXO VII – TRADUÇÃO AUTOMÁTICA**



## **EXEMPLO 2.120 - TRADUÇÃO DE UM TEXTO DE PORTUGUÊS PARA ESPANHOL**

### **UTILIZADO O APERTIUM**

Considere-se o seguinte texto escrito em Português:

“Numa tarde de Agosto muito quente estava eu na estação de Nine a espera do comboio que me iria levar a viagem fantástica até Paris (França). O comboio chegou a hora prevista que estava marcada no bilhete. Entrei e sentei-me a meio da carruagem foi quase das primeiras pessoas a entrar. Passados cinco minutos, quando o comboio estava prestes a iniciar o seu percurso, uma mulher cega sentou-se a minha beira. Já nos tínhamos andando um bom bocado quando a senhora falou.”

Ao seleccionarmos em Direction: Português → Espanhol obtemos o seguinte resultado:

"En una tarde de Agosto muy caliente estaba yo en la estación de Nine la espera del tren que iría a llevarme el viaje fantástico hasta París (Francia). El tren llegó la hora prevista que estaba marcada en el pasaje. Entré y me senté a medio de la carruagem fue casi de las primeras personas a entrar. Pasados cinco minutos, cuando el tren estaba prestes a iniciar su recorrido, una mujer ciega se sentó mi riba. Ya nos teníamos andando un bueno bocado cuando la señora habló. "

## EXEMPLO 2.121 - TRADUÇÃO DE UM TEXTO DE PORTUGUÊS PARA ESPANHOL

### UTILIZADO O GALICIAN-PORTUGUESE TRANSLATOR

Considere-se como *input* o texto do exemplo 2.120.

Depois de executarmos o comando `cat input-file | port2gal.x > output-file` na linha de comandos, o ficheiro `output-file` fica com o seguinte conteúdo:

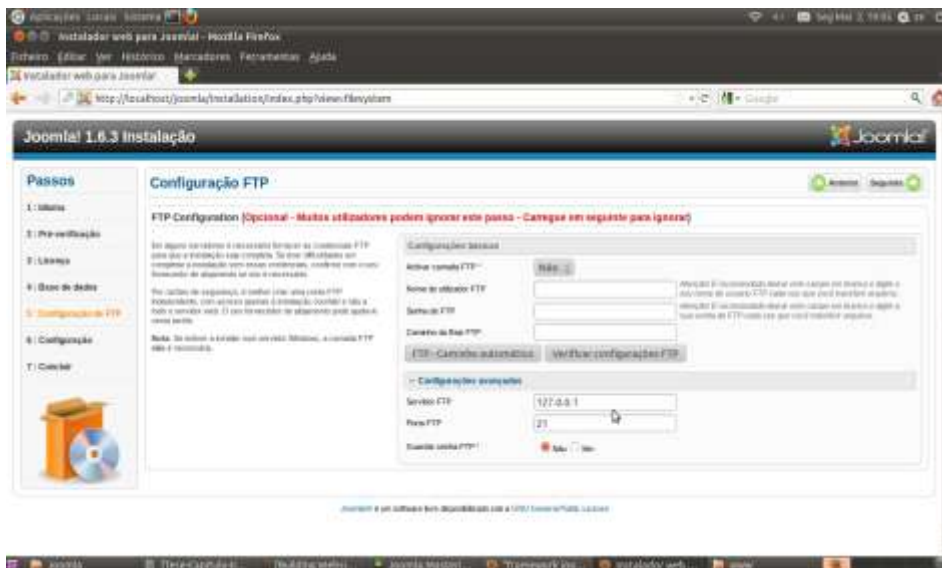
*output-file*

En una tarde de Agosto muy caliente estaba yo en la estaci3n de Nine la espera del tren que ir3a a llevarme el viaxe fant3stico hasta Par3s (Franc3a). El tren lleg3 la hora prevista que estaba marcada en el pasaxe. Entr3 y me sent3 a med3o de la carruaxe f3e cas3 de las primeras personas a entrar. Pasados cinco minutos, cuando el tren estaba prestes a iniciar su recorrido, una muxer ciega se sent3 mi riba. Ya nos ten3amos andando un bueno bocado cuando la se3ora habla3.

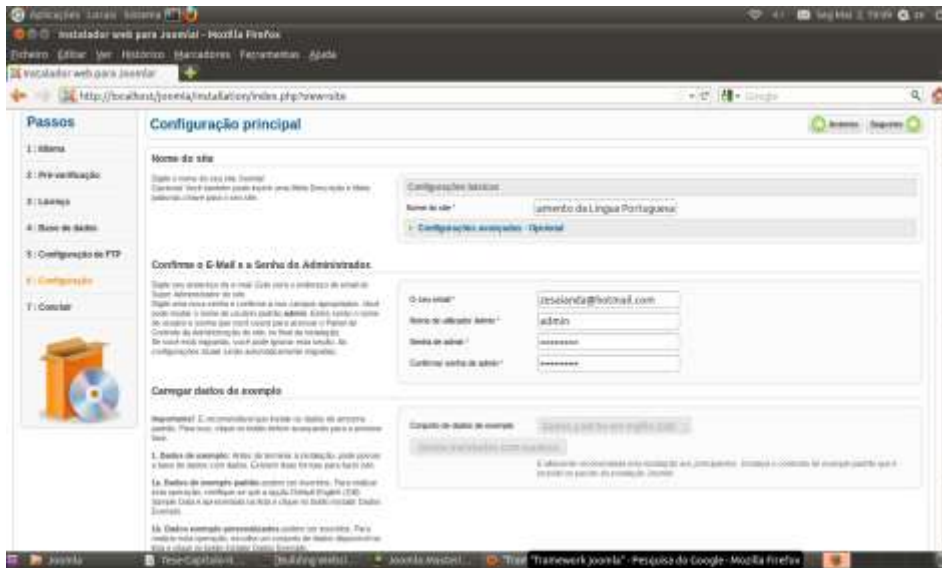
## **ANEXO VIII – PROCESSO DE INSTALAÇÃO DO CMS JOOMLA!**













**ANEXO IX – FICHEIRO DE ESTILIZAÇÃO DO MENU PRINCIPAL DO *SITE***



## menuprincipal\_css.css

```
.moduletable_mainmenu{
width:200px;
background: #eee;
}

.moduletable_mainmenu ul {
/* remove bullets and list indents */
list-style: none;
margin: 0;
padding: 0;
}

/* style, color and size links and headings to suit */
.moduletable_mainmenu a, .moduletable_mainmenu h2 {
font: bold 11px/16px arial, helvetica, sans-serif;
display: block;
border-width: 1px;
border-style: solid;
border-color: #ccc #888 #555 #bbb;
margin: 0;
padding: 2px 3px;
}

.moduletable_mainmenu h2 {
color: #fff;
background: #000;
text-transform: uppercase;
}

.moduletable_mainmenu a {
color: #000;
background: #efefef;
text-decoration: none;
}

.moduletable_mainmenu a:hover {
color: #a00;
background: #fff;
}

/* Positioning the Pop Outs */
```

```
.moduletable_mainmenu li {
/* make the list elements a containing block for the nested lists */
position: relative;
}

.moduletable_mainmenu ul ul ul {
position: absolute;
top: 0;
left: 100%; /* to position them to the right of their containing block */
width: 100%; /* width is based on the containing block */
}

/* Hiding and Revealing using :hover */

div.moduletable_mainmenu ul ul ul,
div.moduletable_mainmenu ul ul li:hover ul ul
{display: none;}

div.moduletable_mainmenu ul ul li:hover ul,
div.moduletable_mainmenu ul ul ul li:hover ul
{display: block;}
```

**ANEXO X – FICHEIRO DE ESTILIZAÇÃO DA *TEMPLATE***





```
template_css.css
```

```
body{  
    text-align:center;  
}  
  
#wrap{  
    min-width:760px;  
    max-width:960px;  
    width:auto !important; /* IE6 hack */  
    width:960px; /* IE6 hack */  
    margin:0 auto; /* center hack */  
    text-align:left; /* center hack */  
}  
  
#header{  
  
#sidebar{float:left;width:20%;overflow:hidden;}  
  
#content{float:left;width:60%;overflow:hidden;}  
  
#sidebar-2{float:left;width:20%;overflow:hidden;}  
  
#footer{clear:both;}  
  
.inside{padding:10px;}
```

```
/*Compass Design typography css*/

|
*{margin : 0 ;
padding:0;
}

h1,h2,h3,h4,h5,h6,p,blockquote,form,label,ul,ol,dl,fieldset,address{
margin:0.5em 0}

li,dd{margin-left:1em;}
fieldset{padding:.5cm;}

body{font-size:76%;
font-family:Verdana,Arial,Helvetica,sans-serif;
line-height:1.3;
margin:1em 0;
}

#wrap{
border:1px solid #999;
}

#header{
border-bottom:1px solid #999;
```

```
}

#fotter{
    border-top:1px solid #999;
}

a{text-decoration:none;
}

a:hover{text-decoration:underline;
}

h1,.componentheading{
    font-size:1.7em;
}

h2,.contentheading{
    font-size:1.5em;
}

h3{
    font-size:1.3em;
}

h4{
    font-size:1.2em;
}
```

```
h5{
    font-size:1.1em;
}

h6{
    font-size:1em;
    font-weight:bold;
}

#footer,.small,.createdate,.modifydate,.mosimage_caption{
    font:0.8em Arial,Helvetica,sans-serif;
    color:#999;
}

#title{
    text-align: center;
}
```