

Nuno Filipe Magalhães Guimarães

Detecção Automática da Negação



Universidade da Beira Interior

Departamento de Informática

Agosto 2009

Nuno Filipe Magalhães Guimarães

Detecção Automática da Negação



Tese submetida ao Departamento de Informática para o preenchimento dos requisitos para a concessão do grau de Mestre efectuada sob a supervisão do Doutor Gaël Harry Dias, Professor Assistente no Departamento de Informática da Universidade da Beira Interior, Covilhã, Portugal

Universidade da Beira Interior
Departamento de Informática
Agosto 2009

Agradecimentos

Tendo finalizado a minha dissertação resta-me manifestar a minha gratidão e reconhecimento a todas as pessoas que me ajudaram.

Ao meu orientador, Professor Gaël Harry Dias, por despertar o meu interesse neste projecto e apoio dado.

Aos meus pais, por me permitirem chegar até aqui com o apoio e incentivo.

Finalmente aos meus colegas e amigos pela ajudada dada.

Muito obrigado a todos,

Nuno Guimarães.

Abstract

This dissertation deals with a methodology that has not yet been implemented. The construction of an innovative system able to detect negative sentences without the need to indicate what is the negation and independent of any language. A system capable of alone and without any kind of help, lists of negative words, part-of-speech, can make the identification of the negation.

The system, through a sequence analysis of the corpus, are extracted lists of words that contain a negation. Of each method applied, resulting in a smaller list of words that caused the previous method and be used in the next method, until the list contains only those negations.

Resumo

Esta dissertação trata de uma metodologia que ainda não foi posta em prática. Consiste na construção de um sistema inovador capaz de detectar frases negativas sem a necessidade de se indicar o que é a negação e independente de qualquer língua. Um sistema capaz de sozinho e sem qualquer tipo de ajuda, listas de palavras negativas, etiquetagem morfosintáctica, consiga fazer a identificação da negação.

O sistema, através de uma sequência de análises ao corpus, são extraídas listas de palavras que contêm a negação. De cada método aplicado, resulta uma lista de palavras mais reduzida que a resultante do método anterior e a ser utilizada no método seguinte, até que a lista contenha apenas as negações.

Conteúdo

Lista de Figuras	x
Lista de Tabelas	xii
1 Introdução	1
2 Estado da Arte	3
3 Trabalho desenvolvido	5
3.1 Os termos negativos	5
3.2 Frequência média	6
3.3 Lei de Zipf	8
3.4 Frequência média e Zipf	10
3.5 Co-ocorrência de pares de palavras	10
3.6 Inverse Sentence Frequency	12
3.7 Subjectividade e Objectividade	13
4 Conclusões e trabalho futuro	17
4.1 Conclusões	17
4.2 Trabalho futuro	17
Bibliografia	19
Anexos	21

Lista de Figuras

3.1	Gráfico da frequência média das palavras em português	7
3.2	Gráfico da frequência média das palavras em inglês	7
3.3	Gráfico da lei de Zipf para o português	9
3.4	Gráfico da lei de Zipf para o português	9

Lista de Tabelas

3.1	Número de frases em que aparece pelo menos um par de palavras em português	11
3.2	Número de frases em que aparece pelo menos um par de palavras em inglês	11
3.3	Lista de palavras portuguesas abaixo da média do ISF	13
3.4	Lista de palavras inglesas abaixo da média do ISF	13
3.5	Peso das palavras portuguesas	15
3.6	Peso das palavras inglesas	15
1	Lista de palavras inglesas resultado da intersecção da frequência média com a de Zipf	22
2	(Continuação) Lista de palavras inglesas resultado da intersecção da frequência média com a de Zipf	23
3	(Continuação) Lista de palavras inglesas resultado da intersecção da frequência média com a de Zipf	24
4	Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	25
5	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	26
6	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	27
7	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	28
8	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	29

9	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	30
10	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	31
11	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	32
12	(Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf	33

Capítulo 1

Introdução

A negação faz parte da comunicação humana e constitui uma importante propriedade em todas as línguas. É a capacidade humana para recusar, contradizer, mentir, falar com ironia, etc. Cada língua varia no que diz respeito à posição dos elementos negativos e também na forma desses elementos, o que torna a detecção da negação em Linguagem Natural uma tarefa complexa.

A detecção automática da negação consiste na identificação de frases negativas em textos, isto é, frases que contenham na sua estrutura advérbios ou locuções adverbiais de negação.

Trabalhos existentes relacionados com a detecção da negação fazem uso da etiquetagem morfossintáctica, mas pretende-se que este seja um sistema independente da língua. Esta dissertação tem como objectivo criar um sistema capaz de identificar frases negativas de forma automática para o português e o inglês.

Faz-se aqui a descrição de um sistema capaz de fazer a extracção automática da negação, onde através de uma sequência de análises ao corpus são extraídas listas de palavras que contêm a negação. De cada método aplicado, resulta uma lista de palavras mais reduzida que a resultante do método anterior e a ser utilizada no método seguinte, até que a lista contenha apenas as negações.

Após o presente capítulo introdutório, é abordado o trabalho existente sobre a detecção da negação, no capítulo 2, e é apresentada a abordagem utilizada para a identificação das frases negativas, no capítulo 3.

Por fim, no capítulo 4, são apresentadas as conclusões e trabalho futuro que poderá vir a melhorar os resultados.

Capítulo 2

Estado da Arte

Desenvolver um sistema que permita identificar, para qualquer língua, as palavras ou termos que identificam a negação, sem ser necessário fornecer qualquer tipo de informação sobre o que é a negação, e assim extrair frases negativas de textos, é um trabalho inovador e não existe nada com esta metodologia.

Os métodos que existem capazes de extrair frases negativas fazem utilização da etiquetagem morfossintáctica (Part-of-Speech Tagging). A etiquetagem morfossintáctica é a tarefa de identificar as categorias gramaticais das palavras numa frase, é um processo onde as palavras de um corpus são sequencialmente etiquetadas com etiquetas sintáticas, como “verbo no infinitivo”, “gerúndio” ou “conjunção subordinativa”, por exemplo.

As duas principais abordagens existentes para a implementação da etiquetagem morfossintáctica são: baseada em regras (Constraint Grammar) [4] e probabilística (hidden Markov model (HMM)) [9] [7]. Existe ainda uma terceira abordagem que combina as duas anteriores chamada de híbrida (Brill tagger)[3]. O etiquetador baseado em regras usa uma base de dados escrita à mão que possui regras que especificam, por exemplo, se uma palavra é um nome ou um verbo. O etiquetador probabilístico faz a etiquetagem com a ajuda de um corpus de treino para calcular a probabilidade de uma dada palavra ter uma determinada etiqueta. A abordagem híbrida usa as duas abordagens anteriores para com uma completar o que a outra não conseguiu executar [5]. Depois de todas as palavras estarem etiquetadas, a partir daqui são retiradas todas as frases que contenham na sua estrutura gramatical advérbios ou locuções adverbiais de negação. Mas o problema destas abordagens é serem dependentes da língua, pois exige a existência de alguma informação sobre a negação e não é o que se pretende.

Existe um trabalho já realizado sobre este tema [1] [2] em que se faz um estudo sobre a negação, tentando compreender o que é a negação e como se comporta, de forma a se conseguirem construir técnicas de detecção automática da negação em textos. Este estudo foi efectuado para as línguas portuguesa, inglesa e francesa.

Neste estudo começam por ser averiguados quais os termos negativos para cada língua e realizados uns estudos estatísticos para se perceber qual a percentagem de frases negativas e afirmativas nos corpus de cada língua. De seguida foi feito um teste para a lei de Zipf para compreender a posição das negações no espectro das distribuições dos termos da linguagem. Outro estudo realizado neste trabalho foi a relevância dos termos na vizinhança dos verbos. Foi aplicada a medida TF-IRF, uma adaptação da medida TF-IDF, de 3 termos antes e depois do verbo a que foi chamado de Método Semente. Foi também feito um estudo da distribuição da frequência média dos termos para verificar quais os termos mais atípicos e observar se as negações faziam parte desses termos atípico.

Capítulo 3

Trabalho desenvolvido

Neste capítulo são definidos quais os termos negativos utilizados na língua portuguesa e inglesa, os que permitem transformar frases positivas em frases negativas e são apresentados os testes efectuados para identificar as negações em ambas as línguas.

Para a língua portuguesa, os testes foram realizados usando os 3 primeiros ficheiros do corpus CETEMPúblico (Corpus de Extractos de Textos Electrónicos MCT/Público). O corpus do CETEMPúblico inclui o texto de cerca de 2.600 edições do PÚBLICO, entre os anos de 1991 e 1998, num total de aproximadamente 180 milhões de palavras (versão 1.0). Está dividido em 1.567.625 extractos, cada um estruturado em parágrafos e frases [6].

Para a língua inglesa foram usados os artigos do The Reuters Corpus Volume 1 desde 20/08/1996 até 24/08/1996. O The Reuters Corpus Volume 1 possui 806.791 extractos jornalísticos em língua inglesa produzidos por jornalistas entre 20/08/1996 e 19/8/1997 [8].

3.1 Os termos negativos

Existem várias maneiras de expressar a negação, dependendo da situação, do ênfase dado, entre outros, mas cada língua tem as suas palavras para negar as afirmações.

Na língua portuguesa, uma frase está na forma negativa quando na sua estrutura existem advérbios de negação ou locuções adverbiais de negação. Os advérbios de negação são palavras invariáveis em género e número que modificam verbos, adjectivos, outros advérbios ou frases. Como advérbios de negação temos: *não*, *nem*, *nunca* e *jamais*. Por locuções adverbiais de negação entende-se o conjunto de duas ou mais palavras que têm a função dos advérbios. Como exemplos de locuções adverbiais de negação temos: *de forma alguma*,

de maneira nenhuma, de modo algum. Outras palavras que também podem transformar uma frase afirmativa em negativa são os pronomes indefinidos *nenhum, nenhuma, nenhuns, nenhuma, ninguém e nada* [1].

Na língua inglesa, uma frase está na forma negativa quando a seguir a um verbo auxiliar existe a palavra *not* ou é inserido no verbo um *n't*. Se se tiver, por exemplo, a frase afirmativa “Dan is paying for the meal”, esta torna-se negativa pondo o *not* ou *n't*, “Dan isn't paying for the meal” ou “Dan is not paying for the meal”. Palavras como *no, never e nobody* também transformam frases positivas em negativas [2].

Estes são os termos que o sistema tem de identificar como sendo negativos para que seja possível reconhecer e registar todas as frases negativas do corpus.

3.2 Frequência média

Com esta análise, pretende-se verificar qual a distribuição das palavras por tamanho, em relação ao total de palavras do mesmo tamanho no corpus e qual a posição da negação, ou seja, se a negação se afasta muito ou pouco da média dos termos por tamanho.

Para calcular a frequência média dos termos por tamanho usou-se a seguinte fórmula:

$$FreqMedia(i) = \frac{tokens_i}{types_i}$$

em que para a frequência média de palavras de tamanho *i* se divide todos os tokens de tamanho *i* por todos os types de tamanho *i*. Os resultados obtidos para o português são apresentados no gráfico da figura 3.1 e os resultados para o inglês são apresentados no gráfico da figura 3.2.

Das negações que se verificam acima do desvio padrão, *não, nem, nunca, nenhum e nenhuma* o *não*, é a que mais se afasta da frequência média e do desvio padrão. Logo, conclui-se que o *não* é uma das palavras mais usadas em português, para expressar a negação.

No inglês, a negação que mais relevância tem é o *not*. As outras negações *don't, never e cannot* encontram-se muito perto da linha da frequência média, mas mesmo assim acima da linha do desvio padrão.

Como estas negações são as mais utilizadas, extraiu-se, para cada língua, uma lista das palavras que se situam acima da linha do desvio padrão, onde estão presentes as negações, de forma a remover palavras irrelevantes para o estudo.

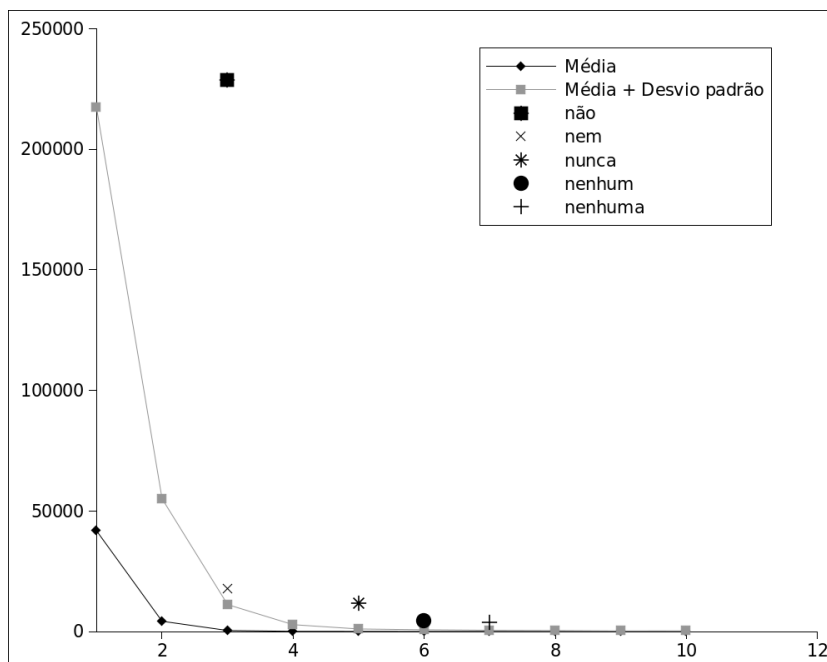


Figura 3.1: Gráfico da frequência média das palavras em português

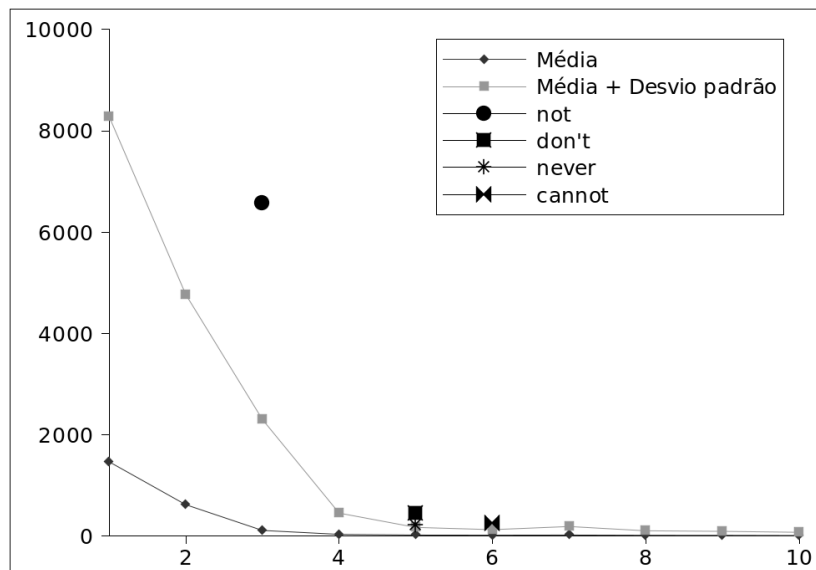


Figura 3.2: Gráfico da frequência média das palavras em inglês

3.3 Lei de Zipf

Para analisar a superfície das palavras de um determinado comprimento em relação a todo o corpus, usou-se a lei de Zipf que nos diz que, num corpus de uma língua, a frequência relativa de um termo é inversamente proporcional ao seu tamanho, isto é, palavras mais longas ocorrem, num corpus, menos vezes que palavras mais curtas.

Utiliza-se, então, a lei de Zipf para calcular a frequência relativa das palavras por tamanho, até tamanho 10, e verificar qual a posição das negações no corpus, indicando se estas são muito utilizadas em texto ou não.

Para calcular a frequência relativa das palavras por tamanho usa-se a seguinte fórmula:

$$Zipf(i) = \frac{tokens_i}{types}$$

onde a frequência relativa das palavras de tamanho i (Zipf), vai ser igual à divisão de todas as ocorrências dos termos de tamanho i (tokens) por todos os termos existentes no corpus (types).

Os resultados obtidos para a língua portuguesa são apresentados no gráfico da figura 3.3 e os resultados obtidos para o inglês são apresentados no gráfico da figura 3.4.

Confirma-se, assim, a lei de Zipf para ambas as línguas, e verifica-se, pela linha de Zipf traçada no gráfico, que quanto mais compridas as palavras menos têm tendência a ser utilizadas, apesar de, no português, nas palavras de tamanho 5 haver uma ligeira subida. Palavras como *e, o, de, da* (português) e *and, the, to, of* (inglês) são mais frequentes do que *flexível* e *undeniable*, respectivamente, que são já umas palavras mais compridas e raras.

No português, a negação com a frequência relativa mais alta é o *não* com 0.0088, seguida do *nem*, com uma frequência relativa de 0.000678. Estas frequências relativas representam já valores muito distantes do valor de Zipf para palavras de tamanho 3, cujo valor da frequência relativa é 0.14.

No inglês, a negação com a frequência relativa mais alta é o *not* com 0.0031, para um valor da frequência relativa de palavras com tamanho 3 de 0.185. A segunda palavra negativa mais frequente é o *no* com um valor de 0.0013, cujo valor da frequência relativa para o tamanho 2 é 0.163.

Pela observação dos gráficos e dos valores da frequência relativa para cada negação, verifica-se também que as negações estão muito abaixo da linha de Zipf em ambas as línguas, o que significa que as negações não ocorrem com tanta frequência, indicando o

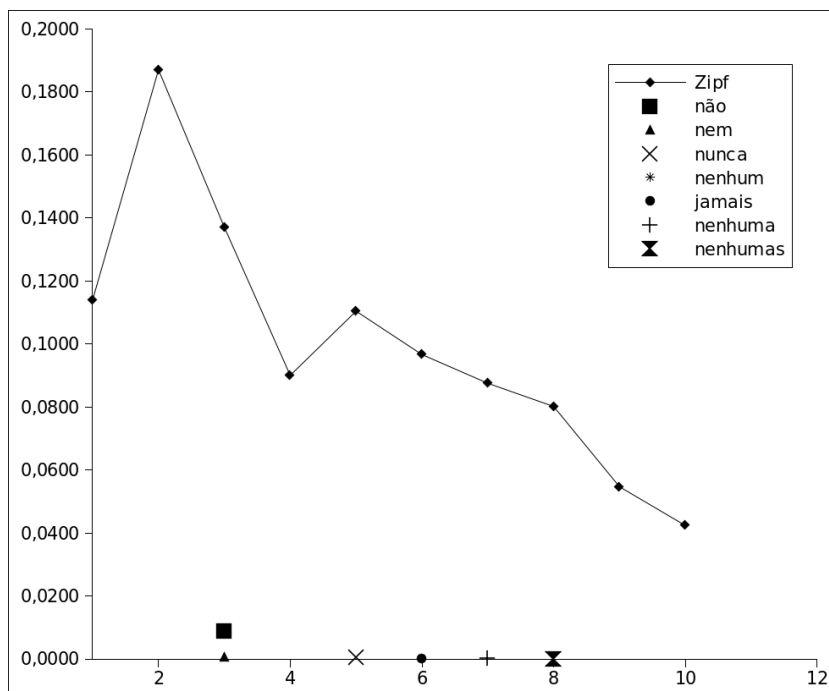


Figura 3.3: Gráfico da lei de Zipf para o português

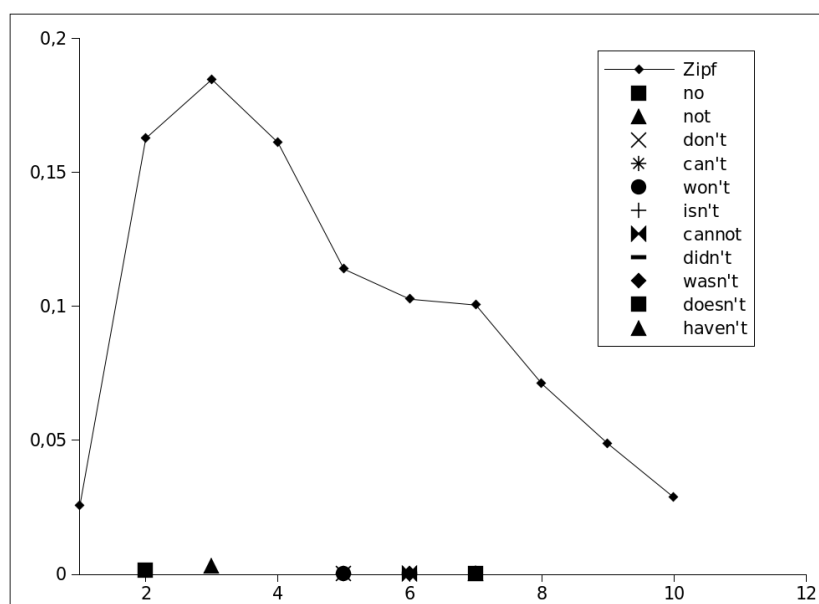


Figura 3.4: Gráfico da lei de Zipf para o português

predomínio de frases afirmativas, tanto no português como no inglês.

Com o cálculo da frequência relativa média para cada tamanho e o cálculo do respectivo

desvio padrão são extraídas todas as palavras acima do desvio padrão para o português e inglês, onde estão presentes as negações.

3.4 Frequência média e Zipf

Destas duas análises ao corpus, foram extraídas duas listas de palavras acima do desvio padrão para ambas as línguas. Para obtermos apenas uma lista com que se possa prosseguir a detecção automática da negação, fez-se a intersecção das duas listas, a da frequência média e a de Zipf, obtendo assim uma única lista de palavras, tendo sido eliminadas muitas palavras sem interesse. Contudo, a lista resultante desta intersecção é ainda muito extensa, com 4107 de 325669 palavras diferentes para o português (tabelas em anexo 4, 5, 6, 7, 8, 9, 11, 12) e 1567 de 83659 palavras diferentes para o inglês (tabelas em anexo 1, 2, 3). Nesta lista final, as negações que estão presentes são o *não*, *nem*, *nunca*, *nenhum*, *nenhuma*, na lista de palavras portuguesas, e o *not*, *don't*, *never* e *cannot*, na lista de palavras inglesas.

3.5 Co-ocorrência de pares de palavras

Partindo do princípio que pode ocorrer dupla negação numa mesma frase, como em “Não acredito que não está a chover”, aplica-se a co-ocorrência de pares de palavras. Pretende-se detectar, na lista resultante da intersecção das listas da frequência média e do Zipf a existência de pares de palavras e quais destes pares ocorrem com mais frequência no corpus, com o intuito de eliminar palavras que não se repetem em frases e que não são importantes.

Calculam-se os valores para cada palavra extraída, obtidos através do somatório de todas as frases em que aparece pelo menos um par da palavra w da lista de palavras que temos.

A fórmula é a seguinte:

$$par(w) = \sum frases(w, w)$$

Mais uma vez, com o cálculo do desvio padrão extraem-se todas as palavras que estão acima do desvio padrão. Os resultados obtidos para o português são apresentados na tabela 3.1 e os resultados obtidos para o inglês são apresentados na tabela 3.2.

Com este método, conseguimos reduzir consideravelmente a lista de palavras, no português para 83 palavras e no inglês para 73 palavras. Contudo, perden-se algumas negações,

1	a	276376	3	num	678
1	o	183655	3	vai	674
1	e	134260	4	para	35888
2	de	363638	4	mais	10306
2	do	101837	4	como	7331
2	da	100253	4	pela	3688
2	em	50639	4	pelo	3589
2	os	45122	4	anos	2688
2	um	38103	4	quer	2241
2	no	24885	4	dois	1442
2	na	21546	4	onde	1207
2	as	21112	4	está	1094
2	se	19675	4	quem	888
2	ao	12365	4	três	862
2	ou	6279	4	seus	729
2	já	1643	4	cada	707
3	que	165774	4	josé	643
3	uma	27872	5	cento	6282
3	com	27303	5	sobre	2741
3	por	25203	5	entre	2367
3	dos	23139	5	muito	1527
3	não	22387	5	ainda	1071
3	das	10707	5	porto	965
3	mil	3908	5	todos	730
3	foi	3510	5	grupo	705
3	ser	3178	5	mesmo	702
3	sua	3129	6	contos	3026
3	nos	2776	6	porque	1162
3	aos	2371	6	contra	1104
3	nem	2311	6	pontos	906
3	são	2138	6	quando	875
3	seu	2120	6	lisboa	828
3	ano	1804	6	estado	819
3	ter	1670	6	teatro	672
3	tem	1657	6	países	642
3	nas	1373	7	milhões	4626
3	dia	1367	7	partido	902
3	sem	1347	7	governo	673
3	mas	1239	8	nacional	731
3	era	1054	8	portugal	643
3	até	964	10	presidente	1167

Tabela 3.1: Número de frases em que aparece pelo menos um par de palavras em português

1	a	8629	4	year	454
2	of	12419	4	bank	445
2	to	12243	4	they	422
2	in	9715	4	this	295
2	on	3483	4	rate	279
2	at	1981	4	been	240
2	is	1484	4	loss	234
2	it	1251	4	1996	229
2	as	1217	4	july	189
2	by	1015	4	more	159
2	be	829	4	week	153
2	he	829	4	corp	148
2	an	371	5	would	547
2	up	283	5	cents	301
3	the	34774	5	which	228
3	and	6997	5	their	209
3	for	2337	5	share	208
3	was	1628	5	rates	190
3	its	823	5	about	187
3	are	601	5	index	174
3	not	541	5	there	165
3	had	532	5	after	161
3	has	522	5	marks	153
3	his	412	5	sales	148
3	new	290	5	party	141
3	per	237	6	market	394
3	pct	198	6	tonnes	344
3	one	177	6	shares	319
3	two	156	6	points	319
3	who	140	6	crowns	183
4	that	1407	6	prices	146
4	said	1315	7	percent	2347
4	from	1069	7	million	1944
4	will	903	7	billion	913
4	with	784	7	company	191
4	were	567	10	government	147
4	have	527			

Tabela 3.2: Número de frases em que aparece pelo menos um par de palavras em inglês

ficando apenas as negações mais utilizadas em cada língua, *não* e *nem* no português e *not* no inglês.

3.6 Inverse Sentence Frequency

O ISF (Inverse Sentence Frequency) é uma técnica usada na extração de informação, que consiste em medir a ocorrência de uma palavra numa coleção de frases com uma relação inversa ao número de frases na coleção, ou seja, a palavras mais comuns como *o*, *a*, *de* para a língua portuguesa e *the*, *a*, *of* para a língua inglesa é lhes atribuído um peso menor em relação a palavras como *teatro* e *impossible* que são palavras menos frequentes.

O ISF é dado pelo logaritmo de base 2 do quociente entre o número total de frases do corpus pelo número pelo número de frases em que a palavra *w* aparece, que se traduz na fórmula:

$$ISF(w) = \log_2 \left(\frac{\text{total de frases do corpus}}{\text{n}^\circ \text{ de frases em que aparece } w} \right)$$

Depois de calculado o ISF para cada palavra da lista, são extraídas as palavras que se verificam abaixo do valor médio do ISF. Os resultados obtidos apresentam-se nas tabelas 3.3 e 3.4, para o português e o inglês respectivamente.

Para o português, obteve-se uma lista com 29 palavras onde se encontra apenas a negação *não*. Para o inglês, obteve-se uma lista de palavras com 36 palavras onde a única negação é o *not*.

Com o ISF, é possível retirar mais palavras sem interesse para a identificação da negação, exceptuando-se a negação *nem*, no português, que acabou por se perder ficando apenas com o *não* que é a mais usada na língua portuguesa. Para o inglês, manteve-se a negação *not* ao mesmo tempo que foram também retiradas palavras sem interesse.

ou	3.091042453358316
pelo	3.044522437723423
sua	3.044522437723423
já	2.995732273553991
ser	2.8903717578961645
foi	2.70805020110221
mas	2.5649493574615367
como	2.4849066497880004
mais	2.3978952727983707
das	2.3978952727983707
ao	2.1972245773362196
se	2.0794415416798357
por	1.9459101490553132
na	1.9459101490553132
as	1.9459101490553132
dos	1.9459101490553132
não	1.791759469228055
no	1.791759469228055
com	1.791759469228055
para	1.6094379124341003
os	1.6094379124341003
uma	1.6094379124341003
um	1.6094379124341003
em	1.3862943611198906
do	1.0986122886681098
da	1.0986122886681098
que	0.6931471805599453
o	0.6931471805599453
e	0.6931471805599453

new	3.1780538303479458
after	3.1780538303479458
this	3.1354942159291497
which	3.044522437723423
would	3.044522437723423
are	2.995732273553991
not	2.995732273553991
have	2.995732273553991
had	2.9444389791664403
million	2.9444389791664403
will	2.8903717578961645
were	2.8903717578961645
an	2.833213344056216
percent	2.833213344056216
he	2.833213344056216
has	2.833213344056216
its	2.772588722239781
as	2.70805020110221
be	2.6390573296152584
it	2.3978952727983707
from	2.3978952727983707
with	2.3978952727983707
is	2.3978952727983707
by	2.3978952727983707
that	2.302585092994046
at	2.1972245773362196
was	2.1972245773362196
for	1.9459101490553132
on	1.6094379124341003
said	1.6094379124341003
in	1.0986122886681098
and	1.0986122886681098
of	1.0986122886681098
to	1.0986122886681098
a	1.0986122886681098
the	0.6931471805599453

Tabela 3.3: Lista de palavras portuguesas abaixo da média do ISF

Tabela 3.4: Lista de palavras inglesas abaixo da média do ISF

3.7 Subjectividade e Objectividade

Nesta fase da detecção automática da negação, o utilizador introduz dois corpus, um subjectivo e outro objectivo. O corpus subjectivo é um corpus onde são expressas as opiniões, pois quando se trata de exprimir opiniões recorre-se frequentemente à negação. O corpus objectivo é um corpus onde são descritos dados factuais como é o caso de jornais.

Aplica-se então a seguinte fórmula:

$$Peso(w) = \frac{w \text{ subjectivo}}{w \text{ objectivo}}$$

O quociente entre o número de ocorrências da palavra w no corpus subjectivo ($w \text{ subjectivo}$) e o número de ocorrências da mesma palavra no corpus objectivo ($w \text{ objectivo}$) permite atribuir maior peso às negações, fazendo com que estas se destaquem.

Os resultados são apresentados na tabela 3.5 e na tabela 3.6, para o português e para o inglês, respectivamente.

Como era pretendido, as negações são das palavras com mais peso da lista de palavras retirada da co-ocorrência. No caso do português, o resultado é excelente podendo-se dizer que a negação *não* foi encontrada, pois esta é a palavra com peso mais alto, 0,95. No caso do inglês, o resultado não ficou muito aquém do desejado, visto a negação *not* ser a terceira palavra com mais peso, 2,95.

não	0,95
ou	0,89
já	0,84
mas	0,82
ser	0,73
se	0,73
que	0,73
mais	0,69
como	0,68
e	0,68
para	0,61
com	0,59
por	0,58
um	0,58
o	0,55
os	0,54
a	0,52
uma	0,5
ao	0,5
as	0,49
foi	0,49
pelo	0,46
no	0,45
de	0,45
em	0,44
na	0,43
do	0,42
dos	0,42
sua	0,41
das	0,39
da	0,38

Tabela 3.5: Peso das palavras portuguesas

its	4,77
it	2,99
not	2,94
would	2,59
this	2,39
that	2,27
be	2,2
have	2,14
which	1,89
as	1,84
were	1,8
was	1,77
are	1,71
is	1,7
had	1,6
the	1,59
of	1,59
has	1,52
by	1,52
will	1,49
and	1,45
an	1,41
to	1,39
for	1,35
in	1,29
at	1,24
on	1,24
from	1,2
a	1,2
with	1,16
said	1,03
new	0,89
after	0,77
he	0,71
percent	0,63
million	0,29

Tabela 3.6: Peso das palavras inglesas

Capítulo 4

Conclusões e trabalho futuro

4.1 Conclusões

O objectivo desta dissertação é construir um sistema que seja capaz de identificar frases negativas, independentemente da língua que se esteja a usar. O sistema construído foi baseado na língua portuguesa e na língua inglesa, onde o objectivo de detectar a negação foi atingido para a língua portuguesa, apesar de apenas se ter detectado a negação *não*, por ser a usada com mais frequência, acabando por ser um bom resultado. Para o inglês, os resultados, apesar de satisfatórios, não corresponderam aos desejados, pois nesta última análise a negação *not*, foi uma das palavras com mais peso, embora tenha ficado na terceira posição da lista final.

4.2 Trabalho futuro

Como neste trabalho não foram encontradas todas as negações, fica ainda trabalho por realizar e concretizar, para serem detectadas todas as negações em várias línguas diferentes, não estivesse esta metodologia de detecção da negação praticamente no início. Futuramente, poderá utilizar-se a abordagem aqui apresentada, mas aplicada de maneira a detectar as palavras negativas com menos frequência, aumentando assim a lista de negações encontradas. Também será bom aplicar esta metodologia a outras línguas estudando o comportamento das negações nessas, tentar métodos alternativos e otimizar os resultados através do conhecimento aqui gerado.

Bibliografia

- [1] Luís Almeida. Detecção automática da negação: Estudos de corpora. Technical report, Universidade da Beira Interior, 2008.
- [2] Luís Almeida. Detecção automática da negação: Extração em línguas diferentes. Technical report, Universidade da Beira Interior, 2008.
- [3] Eric Brill. Transformation-based error-driven learning and natural language processing: a case study in part-of-speech tagging. *Comput. Linguist.*, 21(4):543–565, 1995.
- [4] KARLSSON F., Voutilainen A., and Heikkil J. *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Mouton de Gruyter, Berlin, 1995.
- [5] Ivo Paixão Medeiros. Um etiquetador morfossintático híbrido baseado em modelos de markov escondidos e aprendizado baseado em transformações para o português do brasil. Master's thesis, Universidade Federal do Pará, 2008.
- [6] Rocha P. and Santos D. Cetempúblico: Um corpus de grandes dimensões de linguagem jornalística portuguesa. *Actas do V Encontro para o processamento computacional da língua portuguesa escrita e falada (PROPOR'2000)*.
- [7] Adwait Ratnaparkhi. A maximum entropy model for part-of-speech tagging. In Eric Brill and Kenneth Church, editors, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 133–142. Association for Computational Linguistics, Somerset, New Jersey, 1996.
- [8] T.G. Rose, M. Stevenson, and M. Whitehead. The reuters corpus volume 1—from yesterday's news to tomorrow's language resources. *Proceedings of the Third International Conference on Language Resources and Evaluation*, pages 29–31, 2002.

- [9] Helmut Schmid. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of the International Conference on New Methods in Language Processing*, pages 44–49, 1994.

Anexos

a	last	nana	fall	trade	think	taken	fixed	whole	floor	public	leader	stated
of	also	take	says	rates	north	hours	bills	every	rival	demand	france	raised
to	more	firm	mark	state	costs	terms	coast	trial	heart	period	steady	pounds
in	1996	well	near	years	marks	right	owned	moved	brent	monday	issues	sydney
on	over	bond	repo	index	peace	eight	trust	prior	built	strong	return	rising
at	some	make	case	stock	seven	goods	sampp	tight	cover	likely	rating	bought
is	rate	move	area	since	start	don't	names	filed	dolly	number	issued	street
it	than	both	hold	while	funds	large	hopes	fresh	stood	volume	dealer	county
by	week	good	come	total	later	short	using	calls	ready	saying	fiscal	canada
be	june	such	bill	under	today	given	whose	aimed	sharp	crowns	expect	coming
as	july	same	line	cents	board	sugar	gross	syria	fears	recent	losses	charge
he	into	it's	does	still	cargo	tonne	along	means	eased	latest	energy	rebels
an	down	part	head	ended	range	clear	noted	staff	doing	former	retail	trying
up	told	plan	lead	added	weeks	force	night	peter	giant	trader	signed	year's
the	when	much	cuts	month	yield	times	loans	needs	tough	change	export	across
and	1995	sale	name	close	third	visit	wants	never	asian	policy	rights	indian
for	next	fund	came	bonds	above	worth	break	italy	makes	sector	remain	figure
was	rose	debt	east	lower	crude	white	units	mixed	market	agency	grozny	failed
its	corp	open	left	china	order	grain	judge	quiet	shares	system	killed	africa
has	time	full	meet	south	point	began	miles	black	friday	unless	strike	moscow
but	only	main	must	early	daily	press	david	stage	august	orders	target	survey
had	most	unit	army	party	ahead	april	known	young	prices	levels	mexico	needed
not	half	date	sell	money	japan	taxes	begin	holds	months	amount	senior	labour
are	news	long	want	being	march	light	deals	spain	tonnes	buying	output	groups
new	loss	home	give	power	house	paris	items	lynch	points	quoted	israel	direct
one	high	hong	sept	major	based	radio	named	girls	profit	europe	coupon	except
two	york	held	lost	issue	small	notes	trend	focus	people	credit	tender	school
his	made	west	term	these	least	czech	korea	games	before	traded	assets	gained
who	fell	seen	food	local	final	legal	pesos	study	higher	result	forces	almost
pct	four	like	less	chief	gains	boost	maker	crown	second	taiwan	action	member
per	them	even	crop	world	india	media	reach	franc	growth	office	taking	posted
net	five	each	would	basis	value	areas	round	irish	income	health	better	bank's
all	rise	cost	which	where	found	raise	build	links	should	french	become	francs
said	back	then	after	wheat	again	rules	leave	quite	closed	agreed	bureau	treaty
that	very	past	first	court	firms	allow	death	child	london	common	pretax	impact
from	deal	1994	their	banks	stake	heavy	tokyo	day's	united	called	troops	enough
with	days	help	about	union	below	civil	bring	field	around	future	rupees	copper
will	note	work	there	level	among	swiss	match	signs	states	supply	199697	others
were	many	cash	other	prime	asked	we're	water	storm	german	annual	ruling	effect
have	city	sold	share	plans	paper	dutch	great	dated	police	showed	nearly	behind
year	late	call	could	those	joint	10000	front	fully	during	within	adding	global
this	just	kong	group	going	might	steel	sides	parts	report	record	russia	helped
they	data	took	price	offer	lebed	human	rebel	smith	budget	little	source	spread
been	1997	gold	sales	until	plant	rally	block	hotel	stocks	series	centre	denied
bank	what	used	three	talks	place	texas	labor	broke	dollar	region	making	fourth

Tabela 1: Lista de palavras inglesas resultado da intersecção da frequência média com a de Zipf

199596	middle	that's	effort	traders	members	biggest	systems	japan's	although	republic
active	versus	border	george	because	brokers	defence	brought	cutting	director	activity
deputy	ground	reason	invest	general	without	special	chechen	discuss	currency	proposed
yields	forced	brazil	always	earlier	largest	china's	yearago	jakarta	saturday	maturity
summer	church	accept	larger	reuters	deficit	auction	changed	iranian	japanese	fighting
equity	things	sweden	polish	through	balance	problem	monthly	mexican	november	included
wanted	opened	turned	factor	average	offered	players	airline	boosted	possible	regional
season	stable	played	expand	capital	control	decided	savings	slipped	remained	canadian
parent	family	values	diesel	quarter	clinton	january	ukraine	illegal	chairman	followed
review	either	unions	looked	current	selling	outside	minimum	quickly	delivery	property
having	cattle	option	slight	meeting	lloyd's	changes	certain	world's	research	approved
markka	firmer	houses	nearby	further	planned	officer	initial	ordered	newsdesk	includes
mostly	safety	lowest	scheme	results	comment	weekend	waiting	cargoes	monetary	shipping
access	turkey	damage	living	service	revenue	provide	towards	thursday	slightly	approval
really	merger	martin	toward	markets	details	opening	remains	newsroom	pressure	pakistan
social	ending	matter	choice	october	western	dropped	there's	expected	problems	probably
moving	fields	battle	formed	dealers	nuclear	account	highest	interest	election	discount
island	jumped	itself	prague	another	chinese	zealand	falling	minister	dividend	refinery
though	online	resume	create	federal	african	growing	efforts	national	guilders	february
mainly	claims	seized	status	russian	product	israeli	charged	analysts	treasury	estimate
caused	sunday	priced	trades	support	workers	accused	nations	official	building	offering
weaker	cities	follow	nasdaq	british	reserve	release	england	economic	northern	congress
weekly	plants	hotels	modest	economy	leaders	airport	merrill	business	released	producer
bourse	became	metals	offset	futures	session	closing	reduced	european	measures	software
robert	league	spring	voters	figures	yeltsin	expects	cabinet	exchange	chechnya	computer
coffee	barley	taipei	bodies	already	vessels	numbers	premium	reported	positive	projects
import	chance	ensure	blamed	however	banking	sharply	serious	industry	recovery	standard
metres	100000	letter	kansas	analyst	reached	allowed	lending	earnings	revenues	improved
cannot	silver	zlotys	starts	finance	project	flights	medical	increase	southern	majority
reform	threat	weight	backed	exports	started	housing	revised	forecast	proposal	reserves
attack	mining	decide	center	council	manager	network	improve	ministry	children	involved
easing	offers	moslem	partly	sources	looking	interim	belgian	compared	capacity	launched
actual	korean	appeal	moment	several	include	moody's	thought	services	recently	longterm
reduce	mother	passed	soviet	whether	charges	seeking	running	decision	received	purchase
buyers	stores	boston	accord	tobacco	dollars	limited	swedish	domestic	declined	hospital
flight	single	paying	percent	country	parties	venture	surplus	previous	holdings	millions
barrel	listed	fallen	million	leading	working	forward	michael	security	billions	payments
launch	winter	mutual	billion	profits	process	outlook	italian	products	campaign	strength
rather	bombay	double	company	despite	holding	options	weapons	contract	refugees	planning
normal	morgan	powers	tuesday	private	stories	similar	harvest	american	spending	province
broker	giving	pushed	against	imports	britain	barrels	quality	military	division	finished
warned	prison	issuer	between	chicago	eastern	concern	present	december	internet	stronger
longer	course	sought	central	morning	overall	decline	traffic	consumer	district	personal
senate	10year	warsaw	foreign	reports	believe	beijing	nothing	continue	position	expenses
week's	crisis	planes	trading	germany	america	weather	holiday	turnover	airlines	believed

Tabela 2: (Continuação) Lista de palavras inglesas resultado da intersecção da frequência média com a de Zipf

evidence	russia's	designed	offshore	overnight	coalition	referring	purchases	threemonth	originally
brussels	consider	premiums	tropical	contracts	suspended	brazilian	apartheid	successful	candidates
trillion	combined	maintain	nation's	municipal	expansion	supported	cigarette	equivalent	wellington
supplies	material	brothers	advanced	authority	published	direction	excluding	employment	everything
chemical	totalled	recorded	netscape	forecasts	britain's	passenger	threeyear	considered	foreigners
aircraft	deposits	proceeds	approach	scheduled	beginning	suggested	initially	regulation	aggressive
arrested	commerce	internal	sixmonth	economist	remaining	described	utilities	population	connection
vehicles	prepared	affected	sterling	telephone	effective	materials	northwest	interested	attendance
progress	soldiers	managers	wednesday	interview	reduction	residents	directors	supporters	eventually
movement	policies	19961995	september	confirmed	increases	continues	involving	executives	leadership
provided	accuracy	attorney	president	potential	completed	reporting	consensus	protection	ambassador
negative	response	speaking	statement	situation	resources	candidate	indicator	throughout	associated
accepted	suffered	ordinary	companies	equipment	different	delivered	virtually	3124088787	attractive
violence	costello	minority	officials	operation	predicted	portfolio	advantage	separatist	retirement
hundreds	investor	interfax	financial	relations	interests	wholesale	expecting	accounting	schillings
concerns	malaysia	strongly	investors	petroleum	germany's	certainly	southeast	electronic	regulatory
returned	atlantic	coverage	operating	corporate	generally	treatment	aluminium	immigrants	tournament
weighted	promised	extended	spokesman	important	democrats	expressed	dismissed	guerrillas	associates
takeover	champion	budapest	following	estimates	consumers	preferred	operators	registered	calculated
rejected	verified	shortage	including	specified	alexander	composite	discussed	indicators	apparently
managing	addition	comments	political	technical	worldwide	standards	argentine	yugoslavia	controlled
existing	separate	acquired	announced	customers	ministers	satellite	questions	represents	optimistic
students	complete	governor	agreement	shortterm	civilians	brokerage	reflected	developing	maintained
taiwan's	deutsche	fighters	inflation	institute	netanyahu	emergency	surprised	indonesian	widespread
question	mortgage	people's	unchanged	liquidity	challenge	commander	pollution	commitment	provisions
interior	facility	helsinki	countries	transport	immediate	appointed	diplomats	statements	reductions
fiveyear	anything	hectares	according	positions	education	residence	slaughter	diplomatic	presidency
possibly	announce	mexico's	executive	community	democracy	documents	providing	themselves	seasonally
starting	electric	weakness	australia	firshalf	shipments	commodity	guarantee	currencies	publishing
strategy	required	deadline	newspaper	difficult	sanctions	hurricane	socialist	compuserve	generation
shanghai	straight	directly	increased	sentiment	strategic	exporters	developed	introduced	discovered
partners	conflict	awaiting	committee	benchmark	ceasefire	aggregate	economics	processing	referendum
appeared	pipeline	reaction	singapore	marketing	something	convicted	therefore	philippine	reasonable
assembly	aviation	training	insurance	secretary	microsoft	meanwhile	establish	individual	delegation
gasoline	market's	friday's	company's	quarterly	employees	triggered	influence	tightening	determined
accounts	carrying	tomorrow	estimated	proposals	territory	reference	movements	identified	prudential
overseas	forestry	covering	elections	frankfurt	committed	otherwise	structure	2128591610	
surprise	produced	features	continued	tuesday's	prospects	maskhadov	government	programmes	
declared	officers	chrysler	available	communist	financing	francisco	investment	criticised	
moderate	imported	fullyyear	reporters	amsterdam	yeltsin's	provision	securities	monitoring	
unlikely	benefits	africans	afternoon	yesterday	argentina	producing	production	containers	
alliance	operator	halfyear	currently	condition	chemicals	confident	operations	strengthen	
monday's	slowdown	admitted	country's	concerned	necessary	decisions	commission	passengers	
together	thailand	controls	programme	interbank	producers	attention	australian	particular	
allowing	adjusted	bluechip	thousands	indonesia	indicated	treasurer	bundesbank	properties	

Tabela 3: (Continuação) Lista de palavras inglesas resultado da intersecção da frequência média com a de Zipf

a	ele	isto	data	grupo	nível	época	lopes	zonas	ouvir	pediu	órgão	bases
o	lhe	terá	rede	disse	seria	único	ficar	nomes	greve	dança	custo	rusa
e	diz	todo	sala	forma	norte	série	venda	saída	golos	mudar	apelo	filha
de	fim	após	1991	final	ponto	terra	longe	fácil	baixo	surge	ficam	dupla
do	ver	toda	cabo	antes	plano	havia	taxas	idade	médio	alemã	quero	anual
da	vão	essa	fica	maior	podem	pinto	média	vamos	olhos	torno	pista	jaime
em	para	teve	1994	tinha	pedro	nossa	criar	levar	unita	curto	falou	torna
os	mais	meio	vale	deste	maria	temos	tenho	podia	ciclo	álbum	vídeo	bispo
um	como	área	liga	facto	saúde	falar	risco	china	acaba	troca	avião	preso
no	pelo	seis	cujo	pouco	novos	gente	resto	festa	grave	russo	couto	erros
na	pela	zona	pena	quase	valor	fundo	nesse	temas	viver	évora	deram	major
as	anos	nome	nove	desta	feito	prazo	corpo	japão	breve	igual	suíça	exige
se	está	aqui	alta	mundo	paulo	terem	fazem	porta	caixa	vasco	ronda	chuva
ao	dois	esse	algo	cinco	filme	minha	áreas	praça	banda	graça	bento	rigor
ou	este	deve	veio	lugar	tanto	passa	preço	vindo	belém	amigo	golpe	civis
já	esta	fora	1995	neste	serão	forte	deles	santo	moeda	moura	metro	pedra
que	onde	eles	nuno	poder	nesta	houve	terão	armas	david	devia	subir	redes
uma	três	obra	sede	<i>nunca</i>	ideia	chefe	teria	canal	velho	vinte	viana	serve
com	numa	pois	mãos	estar	união	aliás	haver	estou	clara	etapa	dando	vinha
não	seus	luís	1990	outra	estas	fosse	sejam	feira	trata	natal	cujos	ganha
por	hoje	logo	taça	pelas	coisa	manhã	serem	carro	sinal	dizia	posso	pires
dos	pode	água	país	todas	junto	feita	carta	jogar	abrir	longa	sabia	soube
das	país	tipo	golo	meses	total	civil	curso	deixa	obter	andar	viseu	tirar
mas	quem	modo	sair	então	prova	crise	bolsa	águas	hotel	reino	moral	conde
foi	isso	sete	cima	noite	volta	gomes	pagar	filho	acima	negro	basta	tomou
ser	será	fase	entre	silva	novas	sousa	rádio	visto	essas	somos	vivem	subiu
sua	duas	face	ainda	vezes	ficou	linha	ajuda	verão	vítor	rocha	honra	vento
nos	tudo	eram	sobre	obras	morte	lista	cargo	21h30	posse	dólar	resta	índia
seu	sido	base	ontem	local	casos	autor	ponte	queda	praia	serra	sério	culpa
são	caso	hora	muito	jorge	longo	razão	museu	atrás	marco	largo	sonho	jesus
aos	dias	vila	mesmo	tarde	sendo	disso	tomar	posto	peças	fazia	ferro	bordo
tem	novo	oito	cento	causa	ordem	junho	santa	prevê	chega	casal	fecho	vírus
ter	suas	maio	foram	dizer	saber	porém	devem	crime	macau	legal	sabem	cujas
até	nova	real	todos	quais	papel	certo	comum	esses	ritmo	padre	faria	votar
nas	josé	dado	agora	geral	abril	texto	levou	marca	vigor	delas	leite	verba
ano	lado	sabe	parte	conta	clube	algum	ambos	dessa	menor	visão	actos	irmão
sem	cada	arte	fazer	tendo	acção	paris	verde	casas	pedir	pensa	termo	carga
dia	seja	cuja	porto	horas	tenha	dados	votos	baixa	gosto	êxito	daqui	barco
mil	joão	1992	estão	costa	vista	braga	favor	capaz	opção	velha	custa	canto
num	casa	luta	tempo	falta	jogos	perto	livre	disco	tímor	vinho	perda	monte
vai	vida	taxa	assim	apoio	única	claro	possa	leste	olhar	actor	corte	ópera
era	além	1993	desde	banco	campo	março	nessa	droga	douro	artur	morto	chama
vez	qual	lhes	pelos	estes	livro	fonte	desse	passo	beira	sorte	aérea	aulas
bem	quer	povo	outro	homem	nosso	meios	muita	palco	ramos	artes	salas	busca
<i>nem</i>	<i>nada</i>	alto	menos	mesma	mário	julho	junta	certa	alves	clima	abriu	acaso
têm	jogo	tema	cerca	líder	força	jovem	dizem	cunha	penso	venha	ambas	acusa

Tabela 4: Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

brito	muitos	ensino	manter	função	formas	estilo	fossem	queixa	marido	prazos	básico
unido	europa	século	âmbito	mortos	subida	oferta	textos	livres	tóquio	loures	baliza
nunes	poderá	estudo	ganhar	figura	ganhou	mantém	marcou	secção	possui	france	abrigo
vozes	altura	edição	filmes	antiga	mendes	factos	criado	factor	citado	médica	trazer
matos	parece	sessão	prémio	nossos	torres	isabel	humana	tornou	normas	leitor	haviam
cores	início	acabou	pedido	partes	férias	mesmos	nasceu	rápido	morrer	certas	méxico
lojas	região	cavaco	tentar	língua	venceu	branca	riscos	george	festas	ordens	ouvido
ténis	deverá	metros	prisão	amaral	sinais	género	afonso	listas	urbana	suécia	exigir
james	actual	imagem	efeito	regras	ameaça	correr	quebra	detido	almada	marcha	máxima
nação	música	locais	aquilo	israel	perigo	tornar	sentir	iraque	negros	noites	chinês
parar	último	espera	angola	custos	aveiro	reuter	física	escala	pronto	provar	borges
também	defesa	jornal	poucos	portas	fortes	desejo	seriam	treino	opções	sombra	ligada
depois	vários	tempos	começa	tomada	fiscal	aposta	daniel	terras	marcos	garcia	prazer
quando	teatro	antigo	branco	enorme	classe	abaixo	querer	barros	tensão	actriz	partiu
lisboa	futuro	debate	evitar	quarto	inglês	adesão	núcleo	canais	eficaz	pacote	trinta
contos	espaço	acesso	alemão	índice	haverá	graças	sérgio	sofreu	judeus	comuns	surgem
porque	frança	chegar	alguém	perdeu	entrou	revela	apoiar	cartas	marcas	corpos	ilegal
estado	unidos	razões	aquela	ideias	órgãos	franco	défiçe	lançou	grécia	surgir	legais
grande	dentro	existe	afinal	global	vendas	sérvia	tantos	agenda	pequim	havido	podiam
apenas	última	amanhã	versão	bósnia	finais	esteja	ficará	presos	tutela	alerta	activa
contra	muitas	igreja	querem	máximo	coelho	morreu	rápida	contou	aceite	dispõe	minuto
outros	várias	visita	modelo	estará	açores	ajudar	limite	ligado	editor	ficado	baixos
câmara	coisas	miguel	pensar	doença	níveis	código	tratar	escudo	página	fechar	prédio
sempre	escola	grupos	perder	senhor	linhas	nossas	quinze	jantar	cancro	serviu	flores
alguns	homens	origem	voltou	tenham	verbas	tarefa	nestas	recebe	teriam	sofrer	xavier
acordo	acções	<i>nenhum</i>	provas	artigo	queria	desses	lógica	acerca	aponta	levado	assume
quatro	sector	medida	refere	mostra	planos	lançar	piloto	berlim	irmãos	formal	demais
cidade	santos	alunos	nações	volume	vítima	teresa	franca	velhos	bolsas	doente	simões
apesar	alguma	deixou	madrid	margem	normal	passam	indica	sócios	vieram	lembra	oeiras
países	devido	aquele	diário	castro	vencer	álvaro	robert	letras	propõe	árabes	enviou
melhor	jovens	pessoa	jardim	década	dívida	luanda	mesmas	lucros	chefes	avanço	pensam
quanto	soares	parque	metade	contar	sintra	vender	seguro	tantas	arafat	antena	bandas
equipa	forças	deixar	viagem	mínimo	filipe	apoios	humano	meados	cartão	sexual	helena
semana	termos	sábado	fundos	iorque	inclui	surgiu	passos	obteve	aviões	ficção	fazêlo
outras	chegou	áfrica	destas	voltar	acabar	aldeia	gerais	acusou	causas	drogas	chamou
social	título	itália	amigos	quinta	postos	admite	criada	verdes	ocupar	cabral	formar
carlos	tinham	seguir	clubes	unidas	comité	poucas	juízes	macedo	discos	enviar	coreia
estava	talvez	brasil	vieira	milhão	possam	motivo	paixão	certos	urbano	fátima	faltam
centro	afirma	entrar	livros	membro	atraso	crimes	poesia	chamar	testes	reunir	falava
número	passar	quadro	filhos	guarda	tropas	leiria	nestes	empate	costas	croata	cantar
guerra	cinema	contas	feitas	compra	carros	feitos	cursos	quarta	quinto	choque	actuar
embora	mulher	destes	fontes	aberto	eleito	recusa	russos	dessas	agente	canção	cristo
manuel	regime	preços	missão	médico	sangue	servir	porquê	canadá	almoço	método	vistos
frente	esteve	rússia	bairro	dúvida	graves	cadeia	beleza	apurou	morais	coloca	papéis
partir	gestão	cabeça	ataque	bancos	duarte	tentou	futura	chegam	aborto	rebelo	juntar
pontos	passou	agosto	sequer	campos	aberta	marcar	pobres	chaves	rainha	braços	mestre

Tabela 5: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

exames	pública	agência	receber	aqueles	parecer	ocasião	tráfico	pilotos	eleitos	cuidado
quiser	direito	semanas	marques	mudança	recurso	alemães	galeria	diploma	recorda	andrade
houver	futebol	procura	martins	tratado	musical	acordos	perdido	escrito	japonês	conclui
diante	ninguém	divisão	últimas	estádio	crédito	fórmula	impedir	externa	segredo	séculos
praias	central	poderão	estudos	recebeu	objecto	reforço	entrega	freitas	iniciou	passará
salvar	serviço	madeira	jogador	esperar	mostrar	anúncio	convite	azevedo	círculo	edições
pintor	vitória	pereira	unidade	idades	ricardo	motivos	federal	barbosa	limpeza	passava
atinge	uropeu	terreno	revista	vítimas	líderes	armadas	prémios	tavares	autoria	adjunto
reside	cultura	escolas	funções	revisão	machado	comprar	escrita	miranda	impacto	horário
deixam	benfica	assunto	escolha	autores	espaços	antigos	estoril	promete	tráfego	consigo
colega	criação	prática	clinton	autarca	algarve	holanda	lançado	bancada	telecom	ranking
teoria	período	estamos	deveria	mandato	parecem	figuras	eleição	detidos	recolha	revelar
segundo	opinião	natural	reforma	daquele	aceitar	seguida	descida	moreira	nuclear	radical
governo	própria	vontade	humanos	memória	chegada	feridos	máquina	ligados	moderna	coragem
público	espanha	popular	redução	defende	câmaras	irlanda	termina	músicos	comboio	tesouro
milhões	verdade	maiores	negócio	energia	marcelo	acusado	decorre	ocorreu	aceitou	centena
durante	passada	maneira	efeitos	daquela	campeão	geração	parecia	fonseca	avaliar	sanções
pessoas	dólares	técnico	general	cimeira	animais	vicente	emissão	aparece	conhece	aquelas
passado	começou	decidiu	fizeram	moscovo	avançar	naquela	artista	império	polónia	recinto
partido	próxima	<i>nenhuma</i>	castelo	palácio	atingir	naquele	interno	dizendo	savimbi	adianta
empresa	entrada	simples	permite	jornais	avenida	enviado	elevado	bélgica	aprovou	espinho
mercado	difícil	espécie	alberto	domínio	chamado	privado	mostrou	depende	iniciar	aprovar
antónio	aumento	regiões	técnica	notícia	pintura	decidir	romance	dívidas	resulta	liberal
algumas	janeiro	sociais	deverão	garante	michael	fazendo	atletas	patente	desenho	métodos
grandes	coimbra	contudo	tivesse	pressão	colocar	batalha	estando	estarem	inglesa	integra
questão	perante	pequeno	corrida	cenário	quadros	senhora	limites	manteve	seguido	antigas
relação	membros	atenção	setúbal	actuais	criança	afirmar	admitiu	modelos	doenças	apontam
entanto	família	sucesso	ligação	interna	recorde	balanço	amadora	marinha	duração	inteiro
através	outubro	equipas	tiveram	referiu	comando	desafio	árbitro	estaria	ligadas	ruptura
próximo	militar	recente	médicos	correia	páginas	véspera	barroso	atingiu	excesso	disputa
capital	justiça	tratase	números	produto	certeza	colegas	décadas	capazes	visitar	implica
exemplo	valores	imagens	joaquim	revelou	cascais	estudar	trunfo	viagens	richard	assalto
sistema	francês	almeida	eduardo	crítica	aliança	orienta	cidadão	imposto	crítico	tornase
estados	minutos	existem	américa	estrela	directa	precisa	jacques	afirmam	debates	enviada
momento	medidas	sampaio	títulos	cumprir	actores	torneio	augusto	retirar	queixas	alegria
últimos	solução	emprego	análise	derrota	directo	registo	existir	chegado	croatas	esperam
reunião	preciso	explica	consumo	esforço	chamada	reduzir	levaram	menores	seguros	pelouro
maioria	escudos	lugares	começar	destino	aspecto	votação	começam	escolar	entende	alfredo
segunda	domingo	estação	dezenas	estreia	mercado	todavia	adeptos	editora	funchal	viatura
decisão	poderia	atitude	agentes	podemos	doentes	reserva	manuela	alterar	crescer	prestar
próprio	caminho	estrada	direita	leitura	inicial	reacção	ataques	turquia	escreve	podendo
afirmou	oficial	ribeiro	turismo	centros	ciência	pacheco	definir	fiscais	lembrar	brancos
mundial	estavam	dúvidas	ficaram	ieltsin	coração	santana	supremo	pedidos	sabemos	valente
posição	partida	jornada	palavra	fábrica	assumir	estarão	sessões	jacarta	inverno	nomeado
sentido	peçoal	londres	diálogo	marcada	cardoso	canções	paredes	artigos	veículo	planeta
polícia	matéria	pequena	combate	sérvios	poderes	privada	croácia	noutros	visível	tomadas

Tabela 6: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

situada	estatal	alcance	dirigir	compras	processo	questões	explicou	notícias	boavista	reuniões
ficando	elevada	devendo	fechado	igrejas	enquanto	sporting	espanhol	momentos	domingos	segundos
pudesse	colocou	exercer	cresceu	revelam	projecto	hospital	diversos	defender	impostos	permitiu
angeles	prepara	hungria	activos	caetano	comissão	relações	gabinete	oficiais	recentes	reformas
estejam	menezes	invasão	apostar	maestro	ministro	operação	críticas	famílias	pavilhão	campeões
quantos	situado	holding	insiste	ocorrer	programa	setembro	terceira	garantir	surpresa	objectos
aliados	envolve	jardins	suharto	línguas	empresas	especial	monteiro	permitir	vencedor	salários
recusou	postura	anuncia	abertas	gouveia	conselho	abertura	encontra	percurso	defendeu	fenómeno
bairros	repetir	traçado	comício	vocação	história	partidos	posições	aspectos	inflação	católica
armando	poderem	cotação	barreto	cantora	européia	pretende	realizar	eventual	garantiu	provocou
salazar	comédia	cartões	obrigou	formado	possível	anunciou	deputado	soldados	silêncio	reservas
litoral	riqueza	queirós	cocaína	cadeira	problema	dezembro	fundação	tradição	desporto	garantia
genebra	atrasos	tivemos	prefere	johnson	direcção	interior	evolução	carácter	protesto	agrícola
assumiu	licença	salário	deserto	sucedeu	eleições	bastante	conflito	terminou	passaram	veículos
aplicar	medalha	dossier	rumores	missões	encontro	terceiro	declarou	pequenas	judicial	entidade
estúdio	referir	enormes	fazerem	pimenta	serviços	melhores	mercados	actuação	unidades	utilizar
árvores	admitir	sujeito	virtude	adultos	tribunal	produtos	recursos	sectores	autarcas	internet
utentes	versões	impasse	digital	ligeiro	fernando	palavras	públicos	destaque	acusação	italiana
margens	suporte	amizade	francos	normais	proposta	européus	próprios	chegaram	escritor	preparar
assinar	chinesa	orlando	devemos	guerras	continua	novembro	espírito	excepção	execução	original
coliseu	embargo	incluir	ocupado	quartos	imprensa	respeito	francesa	imediatos	consegue	tornouse
ameaças	cliente	índices	farende	fabrico	dinheiro	controlo	intenção	assuntos	corrente	absoluta
áustria	revolta	gostava	charles	apoiado	superior	guterres	natureza	exterior	sistemas	referido
célebre	capitão	curioso	sevilha	receios	produção	próximos	resolver	conhecer	elemento	contexto
laranja	decreto	bósnios	atlanta	ausente	ambiente	discurso	ausência	entender	explicar	escreveu
urgente	mostram	censura	chamase	sorriso	presença	comércio	vereador	decisões	polícias	universo
renault	futuros	ucránia	rapazes	volumes	ferreira	hipótese	portanto	demissão	estrelas	máquinas
oferece	ligeira	correio	queriam	auxílio	conjunto	finanças	estatuto	despesas	próprias	positivo
deixado	passeio	armados	adriano	futuras	campanha	exército	acontece	decorrer	entregue	escrever
lançada	ribeira	acessos	infante	esgotos	político	presente	pequenos	assinado	provável	incêndio
debaixo	nascido	adoptar	noruega	critica	educação	carreira	portavoz	evidente	governos	naturais
urbanos	antunes	daquilo	extremo	racismo	director	diversas	centenas	montante	habitual	técnicas
realiza	importa	windows	orgulho	mantido	formação	festival	distrito	discutir	conversa	retirada
sozinho	noutras	pensões	células	guardas	direitos	carvalho	previsto	clientes	condição	complexo
lembrou	esquema	tarefas	obrigar	clínica	concelho	seguinte	concerto	soluções	estradas	fatores
símbolo	lectivo	teatral	afastar	alargar	negócios	milhares	bruxelas	contacto	conforme	proceder
narciso	mistura	havendo	estiver	haveria	oposição	edifício	acidente	pergunta	circuito	centrais
titular	parques	criadas	alegada	arlando	anterior	selecção	material	receitas	esforços	resíduos
william	peritos	obtidos	ouvidos	portugal	mulheres	contrato	prevista	próximas	promoção	henrique
receita	visitas	minoria	docente	nacional	economia	cidadãos	vantagem	passando	alentejo	queremos
roberto	regular	adopção	líquido	primeiro	resposta	regresso	nogueira	artistas	recebido	esquecer
pobreza	notável	ensaios	impacte	primeira	regional	públicas	aumentar	mensagem	promover	concluiu
retrato	brigada	inimigo	isolado	qualquer	crianças	passagem	italiano	aprovado	adiantou	decidido
vizinho	assinou	provoca	chicago	política	oliveira	técnicos	dimensão	advogado	capitais	procurar
treinos	argélia	mandela	debater	situação	alemanha	concurso	trânsito	terrenos	ciências	policial
heroína	manobra	permita	herança	trabalho	cultural	esquerda	polémica	teixeira	aprovada	respeita

Tabela 7: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

inferior	concluir	estranho	poluição	adquirir	cenários	redacção	iniciado	salienta	gravação
poderiam	deveriam	episódio	esperava	arguidos	reações	escolheu	divulgar	aguardar	marcação
colecção	loureiro	príncipe	infância	chamadas	defendem	clássico	agências	suportar	vertente
disposto	dirigida	abandono	decisivo	emissões	afastado	liberais	falência	entregou	françaos
melhorar	pressões	combater	sensação	elevados	arranque	williams	fábricas	angolana	começado
sublinha	decorreu	nomeação	turistas	acusados	alcançar	barracas	gerações	procuram	acredito
informou	funciona	violação	investir	suspeita	criminal	olímpico	marcador	carteira	líquidos
provocar	viaturas	entregar	pormenor	amnístia	concreto	razoável	recolher	recordou	rapariga
privados	confusão	angolano	belgrado	cravinho	comentar	aparente	profundo	daquelas	cristãos
animação	pertence	aventura	trabalha	negativa	vitorino	guitarra	partilha	conheceu	trindade
despacho	conteúdo	imediate	intervir	voltaram	noticiou	delegado	perigoso	arquivos	esquerdo
mudanças	deixaram	oriental	resultou	software	perderam	restante	recordar	pianista	vendidos
ocupação	melhoria	semestre	realizou	críticos	montagem	liquidez	exclusão	inúmeras	eliminar
atentado	disseram	extensão	permitem	alegando	composto	memórias	paralelo	suspensão	correcta
consenso	vitórias	hectares	aparecer	convento	panorama	floresta	critério	registos	suicídio
recorrer	ligações	sentença	musicais	referida	tragédia	desafios	dinâmica	gondomar	penafiel
acabaram	dedicado	barragem	tivessem	fernanda	rebeldes	acabaria	agressão	tentando	criticou
assistir	anunciar	caminhos	vermelha	holandês	vaticano	mantémse	fracasso	garantem	derrotas
pinheiro	aparelho	baptista	detenção	relativo	colóquio	africano	colocada	corridas	entradas
deixando	petróleo	vermelho	propinas	passados	reunidos	créditos	plástico	bloqueio	denúncia
exemplos	medicina	jornadas	regressa	potência	liderada	integram	cartazes	assumido	assessor
registou	terminar	marítimo	aumentou	multidão	cinasta	exemplar	salvador	mistério	acabando
daqueles	urgência	consulta	tradução	marrocos	santiago	comboios	atribuir	destruir	aplicada
exibição	prometeu	feminino	lourenço	clássica	estranha	revistas	homólogo	portaria	mediante
sarajevo	completo	espacial	lisboeta	confirma	dedicada	senhores	manterse	internos	permitam
reforçar	carneiro	expresso	ofensiva	chamados	sucessão	arranjar	gasolina	sustenta	vergonha
eficácia	avanzado	leitores	combates	conduzir	liderado	cedência	decisiva	jurídica	sujeitos
pessoais	dirigido	morreram	conjunta	ocorrido	recessão	atitudes	terminal	inovação	caldeira
ministra	explosão	capítulo	maratona	resolveu	condutor	prejuízo	houvesse	explorar	esperase
obrigado	promessa	produzir	mesquita	barcelos	registrar	africana	sucessor	esperado	ocasiões
familiar	estações	vizinhos	academia	massacre	aprender	esquadra	desastre	mantendo	pacífica
integrar	completa	plenário	encargos	docentes	credores	encerrou	efectiva	editores	pesquisa
bandeira	privadas	responde	pretexto	definido	recepção	bernardo	revelado	procurou	divulgou
positiva	difíceis	apareceu	santarém	proposto	assegura	quarteto	ditadura	marítima	vésperas
mundiais	bilhetes	subsídio	negativo	diminuir	japonesa	períodos	tratados	divisões	observar
conceito	novidade	colocado	cuidados	entraram	chineses	surgiram	expulsão	adequada	previsão
telefone	profunda	ingleses	oferecer	religião	idêntica	corredor	islâmica	iniciada	estética
depressa	sondagem	bragança	levantar	práticas	radicais	fundador	barreiro	secretos	ocupados
perceber	autónoma	efectuar	atingido	desenhos	produtor	fornecer	obtenção	empregos	perfeito
acredita	infantil	reduzido	jurídico	culturas	provedor	idêntico	barreira	apontado	conhecem
ocidente	circular	cotações	tornarse	reduzida	pacífico	composta	suspeito	indícios	partidas
reflexão	analisar	satélite	consiste	proteger	absoluto	internas	catálogo	publicou	decidida
sindical	figueira	relativa	análises	feminina	horários	tínhamos	sensível	negocial	avanzada
negociar	expansão	espécies	instalar	condução	matérias	montanha	encerrar	elaborar	inúmeros
alvalade	gostaria	paisagem	valentim	atenções	extinção	parceiro	cristina	bancário	domínios
escolher	opiniões	comentou	aumentos	disputar	árbitros	aparecem	publicar	definida	marcaram

Tabela 8: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

decorrem	república	propostas	incluindo	revolução	anunciado	protestos	igualdade	estivesse
precisam	segurança	realidade	tendência	faculdade	edifícios	austrália	concluído	policiais
camisola	problemas	relatório	sindicato	essencial	israelita	controlar	universal	conseguem
segmento	condições	ministros	operações	constitui	intenções	publicada	adaptação	jerusalém
enviados	sobretudo	sequência	demasiado	distância	soviética	académica	seminário	católicos
inclusão	municipal	comercial	alteração	protocolo	auditório	previstas	descobrir	atletismo
tentação	sociedade	trabalhar	conclusão	fronteira	favorável	intervalo	monetária	exigência
tenciona	televisão	programas	significa	realizada	temporada	italianos	justifica	olímpicos
operador	objectivo	restantes	aplicação	excelente	possíveis	categoria	inspecção	relativos
petrogal	principal	discussão	exercício	critérios	província	acreditar	encenação	incidente
inimigos	instituto	situações	expressão	audiência	seguintes	indivíduo	elaborado	religiosa
terrível	qualidade	rodrigues	franceses	aeroporto	destinado	escolhido	timorense	colocados
sentidos	resultado	fevereiro	liderança	quaisquer	contratos	convenção	dinamarca	estudante
tentaram	população	estrutura	resolução	encontram	existente	respostas	residente	semanário
assistiu	primeiros	diferente	guimarães	histórias	portuense	conflitos	mostrouse	pretendia
integral	exposição	freguesia	entidades	populares	especiais	renovação	dignidade	instrução
reflecte	contrário	aconteceu	americana	espanhóis	crescente	pretendem	margarida	receberam
mostrado	movimento	dirigente	contactos	britânica	garantias	registado	soberania	habituais
agitação	jogadores	conseguir	européias	confirmou	linguagem	convidado	depósitos	delegados
portista	congresso	protecção	construir	analistas	alexandre	encontrou	monetário	promovido
colonial	conseguiu	professor	automóvel	corrupção	ambiental	colectivo	interessa	obrigação
directos	autarquia	inquérito	regionais	moradores	homenagem	advogados	concepção	dispostos
iniciais	políticos	continuam	estiveram	argumento	anunciada	promessas	argentina	horizonte
simpatia	princípio	violência	espanhola	autonomia	trimestre	novamente	colectiva	monumento
salarial	trabalhos	americano	começaram	madrugada	concertos	prestação	permitirá	obrigados
mosteiro	elementos	económico	delegação	magalhães	afirmação	transição	salientou	autocarro
tratar-se	projectos	presentes	realizado	aquisição	prejuízos	integrado	consórcio	editorial
comitiva	considera	primeiras	encontros	hipóteses	acidentes	divulgado	atribuído	dimensões
orgânica	documento	pagamento	habitação	culturais	realmente	atlântico	escândalo	entregues
marginal	económica	confiança	aprovação	conhecida	filosofia	constante	colocação	efectivos
seguindo	nacionais	treinador	gonçalves	toneladas	concessão	prevenção	prossegue	agrícolas
ligeiros	executivo	fernandes	juventude	recuperar	confronto	convicção	potencial	soviético
sucedido	deputados	referendo	esperança	sublinhou	destinada	previstos	estatutos	negativos
governar	interesse	propósito	distrital	abandonar	hospitais	enfrentar	concelhia	optimismo
desfecho	militares	conhecido	avaliação	suspensão	perguntas	afirmando	nucleares	campanhas
variação	federação	comunista	orquestra	recorde-se	sondagens	casamento	integrada	judiciais
imaginar	liberdade	anteontem	histórico	regressar	previsões	embaixada	comissões	utilizada
atração	orçamento	apresenta	responder	materiais	confirmar	quadrados	trabalham	militante
defensor	encontrar	britânico	organismo	eleitores	eventuais	vantagens	japoneses	resolvido
nordeste	indústria	bombeiros	coligação	definição	africanos	conquista	convencer	suspeitas
humberto	companhia	diferença	barcelona	respondeu	cobertura	continuou	instalado	henriques
irlandês	candidato	acontecer	cerimónia	parceiros	condenado	guilherme	famalicão	efectuada
disposta	francisco	processos	publicado	tribunais	decidiram	profissão	apoiantes	concursos
arrancar	eleitoral	indonésia	funcionar	assegurar	histórica	arredores	organizar	discursos
libertar	continuar	tentativa	acusações	reconhece	subsídios	alimentar	regressou	microsoft
português	políticas	município	ocidental	concelhos	utilizado	verificar	novidades	proibição

Tabela 9: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

adiamento	exclusivo	tradições	adquirido	violentos	acentuada	salariais	cidadania	atingindo
gravidade	confessou	francesas	preservar	referente	educativa	definidos	jugoslava	desejável
progresso	abordagem	aventuras	directivo	numerosos	mãodeobra	oferecido	concedido	preencher
prestígio	revelação	repressão	disputada	contornos	aprovadas	vencedora	austriaco	variações
preparado	brilhante	separação	narrativa	solicitou	determina	frequente	pósguerra	contrária
indonésio	incêndios	andamento	terminado	contactar	pretensão	admitindo	completar	simbólico
primavera	inclusive	descobriu	produzido	satélites	obstáculo	excepções	potências	concebido
hollywood	atmosfera	aparelhos	portagens	circuitos	pretender	defendido	coerência	conseguia
relativas	clássicos	indicador	necessita	sugestões	polémicas	centrista	albaneses	reforçada
instalada	península	permitido	digressão	consultas	funcionam	ideologia	vereadora	reservada
respeitar	sublinhar	atribuída	formações	suspender	alcançado	denunciou	islâmicos	incluídos
provocado	frankfurt	profundas	provocada	disputado	referidas	definidas	ocorreram	irregular
barreiras	altamente	artístico	assinalar	reafirmou	muçulmano	femininos	contribui	canadiano
marketing	chanceler	inscritos	preparase	criadores	escolhida	asiáticos	iniciaram	mostrando
correntes	conselhos	capacetes	educativo	holandesa	denunciar	escolares	esperavam	bancários
optimista	permanece	solicitar	comunismo	encerrado	colaborar	dramático	guerreiro	perigosas
impressão	directora	produções	acompanha	utilidade	alexandra	gravações	discutida	inaugurou
garantido	meiocampo	concorrer	finlândia	autónomas	terceiros	recuperou	empregado	inaugurar
episódios	referidos	discoteca	projecção	florestal	lisboetas	desculpas	declarado	actuações
efectuado	escultura	turístico	executiva	reservado	perigosos	procurado	distintas	salientar
abandonou	concretas	instância	perguntou	vestígios	designado	atendendo	numerosas	ciclistas
agostinho	desenhada	mantevese	agradável	distinção	telefones	promovida	bancárias	preferido
sindicais	surpresas	dominante	financiar	académico	afectados	camarária	levantado	prudência
originais	atentados	trabalhou	percorrer	prometido	alcântara	aceitável	caçadores	sinfónica
carvalhas	envolvido	travessia	titulares	distritos	existirem	comunicar	imposição	festivais
masculino	assegurou	facilitar	indicação	conduzido	preparada	defrontar	submetido	monopólio
eléctrica	dinheiros	mecanismo	carreiras	presidido	jugoslavo	esquecido	carências	aproximar
suspeitos	raramente	confiante	associado	violações	contenção	montanhas	literário	produzida
conceição	encenador	discutido	benefício	imigração	apreensão	iraquiano	percursos	protestar
atravessa	afirmaram	paciência	desespero	esclarece	autêntica	negativas	precisava	cometidos
concluída	bandeiras	acrécimo	motorista	fenómenos	vermelhos	celebrado	desmentiu	fotógrafo
alimentos	diplomata	direcções	informado	atingiram	desportos	pavilhões	agressões	inocência
derrotado	edilidade	colocadas	repetição	lembrando	autêntico	garantida	poderemos	concedida
posterior	albufeira	verificou	turística	múltiplas	sensíveis	candidata	gabinetes	convocada
positivos	sebastião	mensagens	eléctrico	meianoite	assinaram	passagens	visitante	percepção
democrata	operários	reflectir	limitouse	destinase	eslovénia	paquistão	consoante	divididos
homicídio	raparigas	dramática	camarário	reparação	associada	alienação	recebidos	despedida
guerrilha	cemitério	portistas	armamento	relevante	contraste	recusouse	florestas	complexos
presidida	selecções	realizase	decorrerá	argentino	muçulmana	estávamos	currículo	ampliação
milosevic	aceitação	participa	excessiva	urbanismo	serralves	vestuário	atlântica	diferendo
conversas	angolanos	montantes	prisional	defendida	exclusiva	excessivo	descrição	compensar
divulgada	perguntar	pertencem	evidência	pareceres	evidentes	salgueiro	pertencer	guineense
artística	economias	inscrição	abstenção	conceitos	demonstra	afectadas	auditoria	aumentado
netanyahu	navegação	religioso	aprovados	correção	revelaram	anonimato	carregado	comemorar
acreditam	registada	mostraram	manifesto	bicicleta	concretos	favoritos	dirigidas	conversar
directiva	avançados	manifesta	cinquenta	masculina	literária	inaugural	injustiça	fantasmas

Tabela 10: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

argumenta	acessível	finalmente	municípios	conseguido	emergência	apreciação	televisões
ambicioso	sustentar	comunicado	sindicatos	refugiados	companhias	justificar	gorbatchov
obrigando	funcionou	património	exploração	assinatura	constituir	preocupado	saneamento
biografia	esperamos	interesses	comerciais	descoberta	diferenças	satisfeito	geralmente
partilhar	arcebispo	candidatos	tecnologia	realizador	elaboração	mecanismos	registaram
emigração	admiração	tratamento	impossível	personagem	sentimento	engenharia	degradação
municípios	prolongar	particular	moçambique	gulbenkian	totalidade	desportiva	dependente
escritora	tranquila	alterações	governador	necessária	diminuição	freguesias	observação
desilusão	simbólica	democracia	permanente	científica	biblioteca	organismos	entusiasmo
terrestre	empenhado	considerou	estruturas	argumentos	comissário	vigilância	registados
portuária	apontados	terça-feira	comandante	exposições	substituir	vereadores	benefícios
livrarias	paisagens	cooperação	autarquias	conclusões	disciplina	confrontos	reportagem
alexander	meramente	entrevista	VERDADEIRO	indivíduos	timorenses	localidade	plataforma
demolição	destacado	utilização	transporte	encontrase	calendário	inevitável	engenheiro
positivas	irlandesa	objectivos	automóveis	familiares	prioridade	eleitorado	componente
presenças	obtiveram	industrial	considerar	convidados	secundário	fortemente	imigrantes
entendido	presidente	americanos	influência	quantidade	semelhança	organizada	favoráveis
plásticas	portuguesa	desemprego	divulgação	sociedades	organizado	afirmações	emigrantes
protegida	ministério	washington	adversário	disponível	obviamente	manifestou	pescadores
cadáveres	associação	participar	autoridade	desportivo	superfície	realizadas	definitiva
incerteza	construção	habitantes	circulação	conhecidos	individual	figueiredo	conquistar
ocupantes	assembleia	inglaterra	princípios	rendimento	brasileira	orçamental	convencido
idênticas	resultados	financeira	claramente	disposição	nascimento	belenenses	secretária
singapura	entretanto	anteriores	municipais	residência	cessar fogo	associados	realizados
motivação	importante	jornalista	publicação	matosinhos	introdução	empréstimo	beneficiar
condenada	informação	instalação	populações	secretaria	accionista	comentário	aplicações
regulares	comunidade	declaração	libertação	computador	encontrado	apuramento	determinar
acontecia	principais	apresentou	militantes	embaixador	reconheceu	pensamento	comparação
atingidos	actividade	representa	fronteiras	constituem	facilmente	utilizados	prosseguir
comprador	iniciativa	juízo	literatura	israelitas	britânicos	frequência	mitterrand
denúncias	parlamento	lançamento	judiciária	produtores	consumidor	escritores	televisiva
resolvida	secretário	legislação	económicas	visitantes	destruição	esclarecer	envolvidas
liverpool	socialista	referência	movimentos	encontrava	envolvidos	respectivo	continuará
idênticos	território	documentos	preparação	muçulmanos	deslocação	esperanças	compositor
praticado	capacidade	competição	verdadeira	arquitecto	eleitorais	arbitragem	designação
amoreiras	diferentes	brasileiro	consideram	obrigações	modalidade	participou	científico
aceitaram	campeonato	velocidade	identidade	juntamente	residentes	tentativas	estabelece
encomenda	realização	existentes	fotografia	equilíbrio	reconhecer	destinados	propaganda
ficheiros	dirigentes	comunistas	semelhante	explicação	exigências	exemplares	desempenho
coleções	necessário	totalmente	continente	acompanhar	aproveitar	diplomacia	demonstrar
considero	igualmente	suficiente	empresário	atribuição	contribuir	incidentes	espectador
aplicadas	sexta-feira	integração	maastricht	cientistas	orientação	audiências	inaugurada
substitui	existência	manutenção	timor-leste	decretolei	ocidentais	repúblicas	simultâneo
frederico	estudantes	certamente	económicos	jugoslávia	aproveitou	respectiva	directores
reduzidos	estratégia	financeiro	operadores	negociação	composição	conhecidas	humanidade
contratar	apresentar	acrescenta	superiores	pormenores	construído	apresentam	luxemburgo

Tabela 11: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf

facilidade	contributo	envolvendo	eléctricos	espiritual	deficiente	produzidos
corredores	sucessivas	compreende	provocando	acompanhou	magistrado	prisionais
contactado	autocarros	investigar	centenário	publicadas	históricas	depoimento
descoberto	justamente	telefónica	garantindo	transferir	concessões	conseguida
justificou	referentes	diplomatas	felicidade	empreitada	praticados	desenvolve
essenciais	constantes	delegações	seguimento	substância	aumentaram	repartição
permitindo	eliminação	facturação	envolvente	catástrofe	escândalos	inesperado
secundária	democratas	escolhidos	frequentes	tempestade	detrimento	autorizado
históricos	inferiores	distribuir	portalegre	inquéritos	instâncias	ignorância
habitações	porventura	contrariar	atribuídos	escrutínio	anunciadas	aspirações
transmitir	importação	continuava	satisfazer	preventiva	proporções	unicamente
empregados	condenação	conjuntura	monumentos	utilizador	britânicas	convocação
satisfação	partidária	madeirense	brevemente	bastidores	assegurada	solicitado
assassínio	bundesbank	espanholas	pressionar	artesanato	assegurado	
efectuadas	divulgados	sucessivos	felizmente	progressão	bilaterais	
ambientais	contribuiu	indignação	legalidade	verificada	exercícios	
destinadas	controlada	televisivo	permitiram	largamente	recuperado	
obstáculos	berlusconi	schumacher	mitsubishi	reportório	divulgadas	
califórnia	promotores	ultrapassa	utilizando	realidades	distinguir	
encontrada	exteriores	exportação	tolerância	seriamente	prevedose	
participam	matemática	hospitalar	economista	felgueiras	realizarse	
publicados	salgueiros	distritais	encomendas	mobiliário	fundadores	
imaginação	procurando	específica	consciente	provocaram	precisamos	
americanas	escritório	permanecer	defendendo	ocorrência	dissolução	
verificado	específico	seguradora	imaginário	convocados	resultaram	
acontecido	inspiração	autárquica	excelência	intérprete	reconhecem	
professora	vencedores	sofrimento	iluminação	significar	condutores	
previsível	incentivos	fermentos	cinemateca	pretensões	corporação	
discussões	pagamentos	manifestar	prejudicar	imediações	burocracia	
anualmente	instalados	constituiu	incidência	liberdades	explicando	
confirmada	limitações	progressos	testemunha	embarcação	cerimónias	
indústrias	procurador	restrições	construtor	atribuídas	permitiria	
conquistou	esmagadora	resultante	complicada	preparados	vendedores	
introduzir	defensores	prestações	renascença	intitulada	recolhidos	
construída	isolamento	florestais	holandeses	tripulação	acumulação	
definitivo	testemunho	rodoviária	utilizadas	fidelidade	provisório	
potenciais	manchester	jornalismo	integrados	categorias	orçamentos	
intitulado	indicações	hierarquia	adiantando	dividendos	apareceram	
tendências	relatórios	confirmado	partidário	trabalhava	reprodução	
terrorismo	sobreviver	socialismo	artificial	fundamento	opositores	
condenados	atravessar	indonésios	inaugurado	masculinos	maisvalias	
qualidades	cruzamento	religiosos	turísticos	cavaleiros	governação	
complicado	governante	abandonado	quilómetro	pronunciar	controlado	
anunciaram	esclareceu	excelentes	monumental	legalmente	províncias	
quotidiano	celebração	instruções	efectuados	terminaram	inesperada	
exministro	religiosas	inquiridos	soviéticos	consensual	congressos	

Tabela 12: (Continuação) Lista de palavras portuguesas resultado da intersecção da frequência média com a de Zipf