

## Expressões Multipalavra

### Questões lexicais e sintáticas

Elisabete Marques Ranchhod  
Faculdade de Letras da Universidade de Lisboa  
LabEL / Onset  
<http://label.ist.utl.pt>  
Email: [elisabet@label.ist.utl.pt](mailto:elisabet@label.ist.utl.pt)

Paula Carvalho  
Faculdade de Letras da Universidade de Lisboa  
Email: [pfcarvalho@yahoo.com](mailto:pfcarvalho@yahoo.com)

## Expressões Multipalavra – Ponto de Situação

- As unidades lexicais multipalavra (ULM) foram tradicionalmente consideradas como objectos linguísticos idiossincráticos e marginais.
  - A linguística teórica dedicou-lhes pouca atenção.
  - Nas gramáticas, as referências são escassas e pouco consistentes.
  - Nos dicionários, só as ULM grafadas com hífen constituem entradas lexicais (**bem-estar, guarda-costas, livre-arbítrio**).

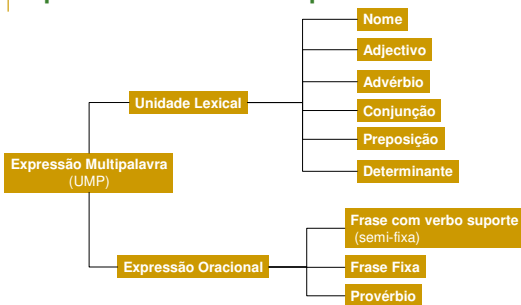
## Expressões Multipalavra – Ponto de Situação

- Porém, na área do Processamento das Línguas Naturais (PLN), tem-se vindo a assistir a um interesse crescente pelas expressões multipalavra e pelos problemas que elas colocam.
  - ACL'03, ACL'04, COLING/ACL'06, EAACL'06, ...
  - *Computer Speech and Language*, 19: 4, 2005, ...
- O não processamento ou o processamento inadequado deste tipo de expressões pode pôr em causa a análise sintáctica e semântica das construções em que elas ocorrem, dificultando ou impedindo o bom desempenho da maioria das aplicações de PLN.
  - tradução automática, resposta a perguntas, elaboração de resumos, extracção de informação, ...

## Na ordem do dia

- Tipos (classes) ...
- Propriedades (variabilidade morfológica, sintáctica) ...
- Interpretação (composicional, não-composicional) ...
- Métodos e técnicas (formalização, processamento) ...
- Ambiguidade

## Tipos de Unidades Multipalavra



À noite, todos os gatos são pardos – De noite, todos os gatos são pardos

## Nomes compostos

- Constituem a parte mais numerosa do léxico nominal das línguas.
- São geralmente formados a partir de palavras simples por meio de regras gerais de combinação de palavras. O seu significado é, na maior parte das vezes, não composicional.
- Ocorrem em contextos tipicamente nominais, mas não devem ser tratados como SN livres!

O tronco é fabricado em **betão** e **fibra de vidro**, enquanto os ramos e as folhas são de plástico.

É uma questão que **quem de direito** deve resolver. A **administração** deve pronunciar-se sobre isto.

## Nomes compostos - Propriedades

O Zé gosta de **algodão doce**

Para identificar os nomes compostos, pode ser utilizado um conjunto de critérios linguísticos, que vão desde a verificação do **comportamento morfológico** dos seus constituintes até à verificação da sua, total ou parcial, **perda de composicionalidade** (lexical, sintáctica e semântica)

### Predicatividade do adjetivo

\*O Zé gosta de **algodão** que (seja + esteja) **doce**

### Variação do adjetivo em grau

\*O Zé gosta de **algodão muito doce**

### Coordenação adjetival

\*O Zé gosta de **algodão doce e saboroso**

### Elição do Adjetivo

\*O Zé gosta de **algodão**

### Ruptura paradigmática

O Zé gosta de algodão (**doce** + \*amargo + \*ácido)

## Nomes compostos - Propriedades

Mantém-se o **braço de ferro** entre os dois sindicatos

Para identificar os nomes compostos, pode ser utilizado um conjunto de critérios linguísticos, que vão desde a verificação do **comportamento morfológico** dos seus constituintes até à verificação da sua, total ou parcial, **perda de composicionalidade** (lexical, sintáctica e semântica)

### Inserção de elementos no GN

\*Mantém-se o **braço forte de ferro** entre os dois ...

### Elição de elementos no GN

\*Mantém-se o **braço** entre os dois sindicatos

### Variação do Det de N2

\*Mantém-se o **braço do ferro** entre os dois sindicatos

### Coordenação de GN

\*Mantém-se o **braço de ferro e de aço** entre os dois..

### Ruptura paradigmática

Mantém-se o braço de (**ferro** + \*aço + \*ferro forjado)

## Classes de nomes compostos

Classe	Estrutura	Exemplos
NA	Nome Adjectivo	vía verde; batata-doce
NDN	Nome De Nome	efeito de estufa; braço-de-ferro
AN	Adjectivo Nome	falsa modéstia; mau-olhado
NPN	Nome Prep Nome	barco a remos; voto em branco
NPV	Nome Prep Verbo	canção de embalar; ferro de engomar
VN	Verbo Nome	coca-bichinhos; ganha-pão
PN	Prep Nome	sem-abrigo; à-vontade
NN	Nome Nome	cara-metade; raio laser
NCN	Nome Conj Nome	saia e casaco; prós e contras
XX	---	Habeas corpus; modus vivendi

## Advérbios compostos (idiomáticos)

- São o caso mais simples de formas compostas. Ocupam as posições sintáticas características dos advérbios e complementos circunstanciais, e, em geral, não são interpretáveis composicionalmente.
- Embora muitas vezes sejam introduzidos por preposição, não devem ser tratados como SP livres!
- São **adunções facultativas** que **comutam com** (e têm o valor de) advérbios simples.

O Zé fez isso (a **contragosto** + constringidamente)

O Zé expôs a questão (de **viva voz** + pessoalmente)

O Zé contou isso à Ana (**tintim-por-tintim** + pormenorizadamente)

## Advérbios compostos - Propriedades

Não permitem

### Inserções

\*O Zé fez isso a **esse contragosto**

\*O Zé expôs a questão **de muita viva voz**

### Reduções

\*O Zé contou isso à Ana **tintim**

\*O Zé expôs a questão **de voz**

### Comutações

O Zé expôs a questão **de viva** (voz + \*presença)

### Alterações morfológicas

\*O Zé expôs a questão **de vivas vozes**

## Classes de advérbios compostos

Classe	Estrutura	Exemplos
P-PADV	Adv	tão-somente
P-PC	Prep C	de rompante
P-PDETC	Prep Det C	à pressa
P-PAC	Prep Adj C	de bom grado
P-PCA	Prep C Adj	a olhos vistos
P-PCDC	Prep C de C	com pezinhos de lá
P-PCPC	Prep C Prep C	de alto a baixo
P-PCDN	Prep C de N	em matéria de N
P-PCPN	Prep C Prep N	no tocante a N
P-PCONJ	Prep C Conj C	contra ventos e marés
P-PV	Prep V W	a bem dizer
P-ACO	(Adj) como C	como uma porta
P-PVCO	(V) como C	como sopa no mel
P-PCO	(V) como Prep C	como do dia para a noite
P-PJC	Conj C	e assim por diante
P-PF	F	sem tigrir nem mugir

## Adjectivos compostos

- A estrutura interna dos adjectivos compostos é variada, e pode ser complexa. Como acontece com todas as expressões não composicionais, observa-se uma fixidez total ou parcial entre os elementos constituintes. Os tipos estruturais mais frequentes são os seguintes:

Classe	Estrutura	Exemplos
AdvA	Adv Adj	bem-parecido; mal-agrado; muito visto
APrepC	Adj Prep C	duro de ouvido; novo em folha; baço para espelho
AConjA	Adj Conj Adj	certo e sabido; impávido e sereno; pobre e mal-agrado
PrepN	Prep Nome	de luxo; de renome; sem retorno

## Classes de adjectivos predicativos compostos

- Tal como os adjectivos simples, os adjectivos compostos podem ser integrados em classes sintácticas, de acordo com os seguintes critérios principais:
- Construção dos adjectivos com **ser** e **estar** ou apenas com um deles;
  - Existência ou não de complementos livres;
  - Aceitação de uma completiva na posição de sujeito e/ou complemento.

Classe	Estrutura	Exemplos
SA	N <sub>0</sub> ser Adj	O Zé é maior e vacinado
EA	N <sub>0</sub> estar Adj	O bife está mal-passado
SEA	N <sub>0</sub> (ser + estar) Adj	O Zé (é + está) doído varrido
QSA	(Que F <sub>0</sub> ) ser Adj	É certo e sabido que vai haver problemas
SAPN	N <sub>0</sub> ser Adj Prep N	A Ana é mal-empregada para o Zé
EAPN	N <sub>0</sub> estar Adj Prep N	O Zé está bem visto junto do eleitorado

## Construções com verbos suporte

- **Frases** em que um verbo sem conteúdo lexical específico aparece combinado com um nome que tem propriedades verbais (i.e. tem estrutura argumental).

### Equivalência entre V e V - N

Eu **concordo** com essa afirmação

Eu **estou de acordo** com essa afirmação

### Equivalência entre Vaux - Adj e Vsup - N

O PR **está atento** a essa situação

O PR **está com atenção** a essa situação

### Sem equivalência (entre Vsup - N e V ou Vaux-Adj)

F = O Zé **tem um fraquinho** pela Ana

## Variabilidade lexical e sintáctica

### Substituição por outros verbos sem alteração do significado

Nós (**temos** + estamos com + **atravessamos**) dificuldades

A Europa (**tem** + desempenha + **assume**) um papel importante

### Frases semanticamente equivalentes a SN

As dificuldades que **temos** > **As nossas dificuldades**

O papel que a Europa **tem** > **O papel da Europa**

### Restrições sintácticas

O Zé **deu** uma rosa à Ana > O Zé deu-a à Ana

O Zé **deu** um estalo à Ana > \*? O Zé deu-o à Ana

## Frases fixas

- Pertencem a registos variados e possuem uma característica comum: contêm combinações *verbo-nome* que **não são distribucionalmente produtivas** nem são interpretáveis composicionalmente.

O Zé **esticou** o pernil

O Zé **salvou** a honra do convento

O Zé **meteu** a mão na consciência

O Zé **chamou** a Ana à pedra

O Zé **esticou** (a perna + o braço)

O Zé **salvou** o património da empresa

O Zé **meteu** a mão no bolso

O Zé **chamou** a Ana ao gabinete

## Frases fixas – Propriedades gerais

Não permite

O Zé **esticou** o pernil

**Apassivação**

\*O pernil foi **esticado** pelo Zé

**Pronominalização**

\*O Zé **esticou-o**

Mas permite

**Inserções** entre o verbo e os SN complemento

O Zé **esticou, inesperadamente, o pernil**

## Frases fixas - classificação

Classe	Estrutura	Exemplos
P- VC0	C <sub>0</sub> V W	O Senhor chamou o Zé à sua presença
P- VC1	N <sub>0</sub> V C <sub>1</sub>	O Zé perdeu a cabeça
P- VC2	N <sub>0</sub> V (C de C) <sub>1</sub>	O Zé salvou a honra do convento
P- VC3	N <sub>0</sub> V C <sub>1</sub> A N <sub>1</sub>	O Zé deu carta-branca à Ana
P- VC4	N <sub>0</sub> V C <sub>1</sub> Prep N <sub>2</sub>	O Zé passou uma esponja sobre o assunto
P- VC5	N <sub>0</sub> V Prep C <sub>1</sub>	O Zé rema contra a maré
P- VC6	N <sub>0</sub> V Prep (C de N) <sub>1</sub>	O Zé tocou na corda sensível da Ana
P- VC7	N <sub>0</sub> V N <sub>1</sub> Prep C <sub>2</sub>	O Zé meteu a Ana num chinelo
P- VC8	N <sub>0</sub> V Prep C <sub>1</sub> C <sub>2</sub>	O Zé fez das tripas coração
P- VC9	N <sub>0</sub> V C <sub>1</sub> Prep C <sub>2</sub>	O Zé entregou a alma ao Criador
P- VC10	N <sub>0</sub> V Prep C <sub>1</sub> Prep C <sub>2</sub>	A novidade voou de boca em boca
P- VC11	N <sub>0</sub> V que F Prep C <sub>2</sub>	Ele soube isso de fonte segura
P- VC12	N <sub>0</sub> V C <sub>1</sub> Prep que F	O Zé pensa duas vezes antes de falar
P- VC13	N <sub>0</sub> V C <sub>1</sub> Prep C <sub>2</sub> Prep N <sub>3</sub>	O Zé tirou as palavras da boca à Ana

## Formalização (LABEL-LEX-mw)

alto (A001) **comissário** (N001), N+NA +Hum+Cargo  
 alto (A216) mar., N+AN+Geo  
 altos (N292) e baixos., N+NConjN +Abst  
 antes de mais., ADV+PDC  
 antes que., CONJ  
 ataque (N201) de asma, N+NP+Med.ms  
 ataque (N201) pessoal (A211) alta comissária, alto comissário. N+NA+Hum+Cargo:fs  
 altas comissárias, alto comissário. N+NA+Hum+Cargo:fp  
 chapéu (N201) alto (A201), N alto comissário, alto comissário. N+NA+Hum+Cargo:ms  
 saxofone (N201) alto (A201), alto mar, alto mar. N+NA + Geo:ms  
 altos comissários, alto comissário. N+NA+Hum+Cargo:mp  
 altos e baixos, altos e baixos. N+NConjN+Abst:mp  
 antes de mais, antes de mais. ADV+PDC  
 antes que, antes que. CONJ  
 ataque de asma, ataque de asma. N+NDN+Med:ms  
 ataque pessoal, ataque pessoal. N+NA+Abst:ms  
 ataques de asma, ataque de asma. N+NDN+Med:mp  
 ataques pessoais, ataque pessoal. N+NA+Abst:mp  
 chapéu alto, chapéu alto. N+NA+Vest:ms  
 chapéus altos, chapéu alto. N+NA+Vest:mp  
 saxofone alto, saxofone alto. N+NA+Mus:ms  
 saxofones altos, saxofone alto. N+NA+Mus:mp

## ULM – Restrições sobre a flexão

As ULM exibem **restrições flexionais** que não podem ser descritas através das propriedades morfológicas dos seus constituintes.

### Palavras simples

alto, A001_ss001 -alto, alto, Adv -altos, alto, Advp -altíssima, alto, A:fs [...] alto, alto, A:ms altos, alto, A:mp	alto, N201_dh201 -alto, alto, N:ms altos, alto, N:mp -altinho, alto, N:Dms [...] baixo, N201 -baixo, baixo, N:ms baixos, baixo, N:mp	chapéu, N201_dh238 chapéu, chapéu, N:ms chapéus, chapéu, N:mp -chapelinho, chapéu, N:Dms [...]
--	---	--

### Palavras Compostas

altos (N292) e baixos., N+NConjN+Abst  
 altos e baixos, altos e baixos. N+NConjN+Abst:mp  
 chapéu (N201) alto (A201), N+NA+Vest  
 chapéu alto, chapéu alto. N+NA+Vest:ms  
 chapéus altos, chapéu alto. N+NA+Vest:mp

Em altos e baixos, N292 indica que se trata de um nome invariável, masculino plural;  
 Chapéu alto é um nome masculino com flexão em número; 'chapéu' mantém a flexão; 'alto' apenas flexiona em número (A201 e não A001).

## Etiquetagem

De um ponto de vista computacional (e linguístico), a correcta identificação dos compostos lexicais evita a geração de uma multiplicidade de análises incorrectas, resultantes da homografia das palavras simples.

8 análises lexicais sw → 1 análise lexical mw

até., ADV  
 até.,PREP  
 prova,prova. N:fs  
 prova,provar.V:P2s:P3s:Y2s  
 em.,PREP  
 contrário,contrário.A:ms  
 contrário,contrário.N:ms

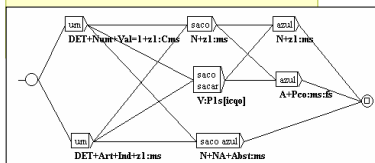
até prova em contrário., ADV

Estes resultados, além de serem desadequados, manteriam ambiguidades que podem ser imediatamente resolvidas ao nível da análise lexical.

## Ambiguidade

As unidades lexicais multipalavra são menos ambíguas do que as palavras simples. Porém, existem combinações de palavras que podem ser analisadas quer como **unidades lexicais compostas** quer como **seqüências livres de palavras simples**.

A empresa **constituiu** um **saco azul**  
 A Maria **comprou** um **saco azul**



## Ambiguidade lexical

Em casos mais isolados, uma mesma palavra composta pode ter associados valores linguísticos distintos.

O Zé sempre viveu em **casas alugadas**

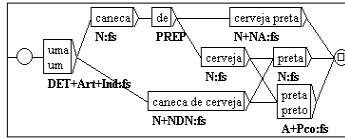
Os **casas alugadas** costumam alajar-se em conchas de moluscos gastrópodes

ZOOLOGIA crustáceo que se aloja, em regra, em conchas de moluscos gastrópodes, protegendo assim o abdómen, que é mole, também conhecido por bernardo-eremita, casa-roubada, paguro, etc. (Dicionário Pro de Língua Portuguesa 2005 Porto Editora, Lda.)

## Ambiguidade – Sobreposição de compostos

Por vezes, uma mesma sequência de palavras pode ser segmentada em dois compostos cujos constituintes se sobrepõem parcialmente.

- Zé **desenhou** uma **caneca de cerveja preta**
- Zé **bebeu** uma **caneca de cerveja preta**
- Zé **queria** uma **caneca de cerveja preta**
- Zé **queria** uma **caneca de cerveja preta**

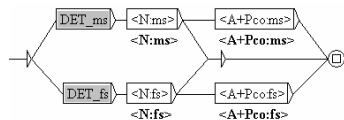


## Ambiguidade

- Como resolver estas ambiguidades?
  - Poderão ser resolvidas localmente? Ou só poderão ser resolvidas no âmbito da análise sintáctica?
- Será adequado etiquetar todas as unidades lexicais multipalavra no momento da análise ou etiquetagem lexical dos textos? Ou só apenas algumas? Quais?
- Que tipo de regras deverão ser utilizadas para distinguir compostos ambíguos de estruturas sintáticas livres?

## Um exemplo de resolução de ambiguidades

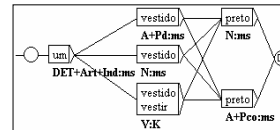
A resolução de ambiguidades, no sistema Intex (M. Silberztein, 1993; 1997), faz-se mediante a aplicação de gramáticas locais, representadas em transdutores de estados finitos.



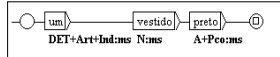
## Análise de SNs que contêm palavras simples

### um vestido preto

Representação em FST das hipóteses de análise lexical:



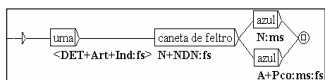
Resultado da resolução de ambiguidades:



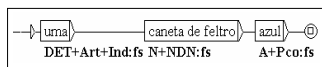
## Análise de SNs que contêm ULM não ambíguas

### uma caneta de feltro azul

Representação em FST das hipóteses de análise lexical:



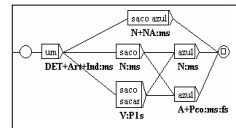
Resultado da resolução de ambiguidades:



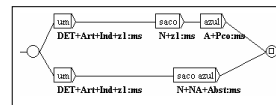
## Análise de SNs que contêm ULM ambíguas

### um saco azul

Representação em FST das hipóteses de análise lexical:



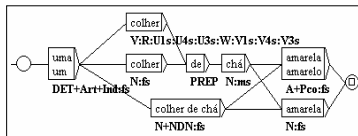
Resultado da resolução de ambiguidades:



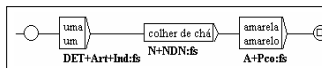
## Análise de SNs que contém ULM ambíguas

### uma chávena de chá amarela

Representação em FST das hipóteses de análise lexical:



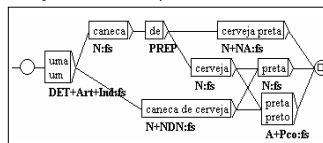
Resultado da resolução de ambiguidades:



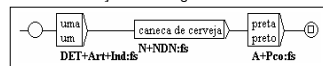
## Análise de SNs onde se observa sobreposição entre compostos

### uma caneca de cerveja preta (a black beer mug / a mug of stout)

Representação em FST das hipóteses de análise lexical:



Resultado da resolução de ambiguidades lexicais:



## A identificação de compostos e a TA

### O Zé só bebe **cerveja preta**

- The Zé alone drinks **black beer**. (Google)
- THE ZÁ© sÃ³ drinks **black beer**. (Amikai)
- THE Zé alone drinks **black beer**. (FreeTranslation)
- The Zé alone drinks **black beer**. (Babelfish)
- The Zé alone drinks **black beer**. (Systran)

### Mantém-se o **braço de ferro** entre os dois sindicatos

- The **arm of iron** between the two unions is remained. (Google)
- Mantém-Itself the **arm of iron** between the two unions. (FreeTranslation)
- MantÃ©m-if the **braço de iron** among the two unions. (Amikai)
- The **arm of iron** between the two unions is remained. (Babelfish)
- The **arm of iron** between the two unions is remained. (Systran)

## A identificação de compostos e a TA

### O Zé expôs a questão de **viva voz**

- The Zé displayed the question of **alive voice**. (Google)
- THE Zé exposed the question **orally**. (FreeTranslation)
- THE ZÁ© expÃ³ s the questÃ£o of **alive voice**. (Amikai)
- The Zé displayed the question of **alive voice**. (Systran)
- The Zé displayed the question of **alive voice**. (Babelfish)

### O Zé é **duro de ouvido** (hard of hearing)

- The Zé is **hard of ear**. (Google)
- ZÁ© Ã© **hard of ear**. (Amikai)
- The Zé is **hard of ear**. (Babelfish)
- The Zé is **hard of ear**. (Systran)
- THE Zé is **hard of ear**. (FreeTranslation)

## A identificação de compostos e a TA

### O Zé **está de acordo com** essa afirmação.

- The Zé is in accordance with this affirmation. (Google)
- THE ZÁ© estÃ¡ in agreement with that afirmaÃ§Ã£o. (Amikai)
- THE Zé is according to that assertion. (FreeTranslation)
- The Zé is in accordance with this affirmation. (Babelfish)
- The Zé is in accordance with this affirmation. (Systran)

### O Zé **concorda com** essa afirmação.

- The Zé agrees to this affirmation. (Google)
- THE Zé complies with that assertion. (FreeTranslation)
- ZÁ© agrees with that afirmaÃ§Ã£o. (Amikai)
- The Zé agrees to this affirmation. (Systran)
- The Zé agrees to this affirmation. (Babelfish)

## A identificação de compostos e a TA

### METRA3

Tradução de "A tarde foi **de novo** passada no mar, **desta vez** na companhia de baleias e golfinhos"

The afternoon **of new** was passed in the sea, **of this time** in the company of whales and dolphins (Google)

The afternoon **of new** was passed in the sea, **of this time** in the company of whales and dolphins (Babelfish)

THE afternoon was **again** last in the sea, **this time** in the company of whales and dolphins (Amikai)

TO late was **of past news** in the sea, **this time** in the company of whales and little gulfs (FreeTranslation)

The afternoon **of new** was passed in the sea, **of this time** in the company of whales and dolphins (Systran)

## A identificação de compostos e a TA

### METRA3

**Tradução de** "No Verão, sempre passa **todos os dias** e traz **caras novas**, que **dão animação** à ilha. "

In the Summer, it passes **every day** and it brings **new faces**, that **give animation** to the island.

**Tradução de** " ... não tivemos, **até hoje**, qualquer acidente grave, que obrigue a uma evacuação **por via aérea**. "

we did not have, **until today**, any serious accident, that compels to an evacuation **for by airmail**.

**Tradução de** "os lares para a terceira idade com fins lucrativos estão a proliferar. "

**the homes for the third age with lucrative ends** are to proliferate.

## A identificação de compostos e a TA

**De vez em quando** falta às aulas  
**Of time in when** it lacks to the lessons

A ponte ruiu **num abrir e fechar de olhos**  
The bridge ruiu **in one to open and to close of eyes**

Desmoronou-se **como um castelo de cartas**  
It was pulled down **as a castle of letters**

Não se pode fumar nos **locais de trabalho**  
If it cannot smoke in the **workstations**

Esse quadro é uma **obra prima** da pintura romântica  
This picture is a **workmanship cousin** of the romantic painting

Tinha uma enorme **dor de cabeça**  
It had an enormous **migraine** 0.K

## Etiquetagem

- Unidades lexicais multipalavra: **nomes, advérbios, preposições.**
- Construções com verbos-suporte

Sem **falsas modéstias**, o leitor que me acompanha sabe que não poderia **estar** mais de **acordo com** a afirmação de que nem todas as **dificuldades** que **atrevessamos têm** a sua **origem** ou a sua **solução no Estado**, entendido como **administração pública**. O **PR chamou**, justamente, **a atenção** para alguns tipos de comportamentos que radicam ou na (falta de) **educação cívica** ou numa atitude laxista e demissionista **por parte** das famílias e dos cidadãos e que, nessa sede, podem encontrar solução. Infelizmente, os exemplos poderiam multiplicar-se: a maior **taxa de incidência** de sida na **Europa** mas a maior percentagem de recurso à **pílula do dia seguinte**; o maior incremento na percentagem de jovens do **sexo feminino** que fumam; dos maiores aumentos de **obesidade infantil**; maior taxa de **abandono escolar**; etc. **Ao contrário** do que nos querem fazer crer alguns discursos, tudo isto não tem só a ver com a **estrutura envolvente** ou com o **Estado** e, portanto, não se sanam apenas nessa instância. São **problemas políticos** cuja resolução depende, **em muito**, da forma como cada um de nós se situa e envolve na promoção do **bem comum**.

Alberto Castro, <http://jn.sapo.pt/2006/06/13>

## Etiquetagem

- Unidades lexicais multipalavra (**nomes, advérbios, preposições, conjunções, ...**)

Mas o tema não é pacífico. É complicado romper **a curto prazo** com um modelo tecnológico, dominante desde o século **XIX**, assente no uso maciço de carvão, petróleo e gás. Estes representam actualmente **80%** do consumo total de energia e são responsáveis pela libertação de **seis mil milhões** de toneladas de **dióxido de carbono** para a atmosfera, conduzindo aos problemas na **camada de ozono** e consequentes **alterações climáticas**. Aposta nas **energias alternativas**. As outras **fontes de energia** actualmente utilizadas no mundo são a **nuclear** (16% da electricidade gerada) e as modernas **energias renováveis**, como a **hidroeléctrica, eólica, solar e geotérmica**, que ocupam apenas **4,5%** da produção total de energia. Daí que um dos objectivos estabelecidos para o **desenvolvimento sustentável** seja a diversificação das **fontes de energia** mais limpas e o aumento da quota de **fontes de energia renováveis** para, **pelo menos**, **5%** em todos os países, até **2010**. Neste caso, a meta da **União Europeia** é mais ambiciosa. A **UE** pretende passar da actual média de **6%** de consumo proveniente das **energias renováveis** para **15%**. Para Portugal não será difícil, **uma vez que já conta com** uma fatia de **13%** **graças à energia hidroeléctrica de acordo com o Eurostat**. Porém, **mais uma vez**, os **Estados Unidos** são os principais opositores à definição de objectivos quantificáveis para o uso das **energias renováveis**, tentando esquecer-se de que são responsáveis por **um quarto** da poluição mundial.

Expresso, Edição online, 31/08/2002

A detenção de [Fátima Felgueiras] na **semana passada** acabou por levar a uma declaração de [Ferro Rodrigues] onde este retirou a **"confiança política"** à autarca e realizou aquele difícil **número de malabarismo** que consiste em **meter os pés pelas mãos**: o **PS** acha que [Fátima Felgueiras] não tem condições para continuar como autarca, já não tem **"confiança política"** na autarca, acha que ela se deve demitir se for acusada formalmente, mas se se provar que houve uma cabala, lá estará para a ajudar a **subir ao pódio** e desmascarar os cabalistas.

Esta posição é incompreensível porque o que os dirigentes do **PS** sempre disseram foi que [Fátima Felgueiras] era inocente e seria considerada como tal pelo partido **até prova em contrário**. Como os tribunais ainda não concluíram nada, não se percebe a mudança de posição.

Para além de todo o circo que envolve a questão e do evidente **mal-estar** dos dirigentes do **PS** (e do **PSD**: veja-se [Durão Barroso], com a prudência de um vidraceiro, a dizer que [Felgueiras] é uma questão interna do **PS...**), há uma questão que **vale a pena** reter: a facilidade com que os políticos **em geral** invocam a independência da justiça e a **presunção de inocência** para evitar comprometer-se no julgamento desta ou daquela figura.

Que a independência da justiça e a **presunção de inocência** são princípios fundamentais já se sabe. Mas isso não pode **servir de cobertura para que** os partidos **fujam à responsabilidade** de avaliar os homens e mulheres que propõem ao país nas suas listas e muito menos à responsabilidade de os sancionar (e afastar) quando a sua prática é inaceitável.

O que gostaríamos de saber, no caso de [Fátima Felgueiras] e dos outros autarcas suspelitos, é se os partidos que os fizeram eleger continuam a **pôr a mão no fogo** e a **cabeça no cepo** por eles ou se, **pelo contrário**, lhes retiraram realmente a sua confiança - o que é, em política, uma posição **sem retorno**. É importante saber isto porque, se um dado autarca pode vir a reconquistar a confiança do seu partido se se safar das **telas da lei**, isso pode significar que esse partido deve deixar de merecer a nossa.

Dizer que só aos tribunais cabe julgar está muito bem, mas isso não pode significar que os partidos se desvinculam dos actos dos seus eleitos **na hora H**, em que a porcaria chega à ventoinha - mas não antes.

José Vítor Malheiros: *A Mão no Fogo Público*, Edição online, 2003

## Em síntese

Tentou demonstrar-se que:

- As unidades lexicais multipalavra podem ser adequadamente identificadas mediante a utilização de um conjunto de critérios sintácticos;
- As unidades lexicais multipalavra não ambíguas podem ser automaticamente etiquetadas no momento da análise lexical dos textos, evitando que cada um dos seus constituintes seja identificado como palavra simples;
- As unidades lexicais multipalavra que também podem corresponder a uma sequência livre de palavras simples, na maioria dos casos, só podem ser desambiguadas por analisadores sintácticos robustos e precisos.

## Bibliografia (1)

- Baptista, J. (1995), *Estabelecimento e formalização de classes de nomes compostos*. Tese de Mestrado, Universidade de Lisboa.
- Carvalho, P. (2001), *Gramáticas de Resolução de Ambiguidades Resultantes da Homografia de Nomes e Adjectivos*. Tese de Mestrado, Universidade de Lisboa.
- Carvalho, P.; Mota C.; Ranchhod, E. (2002), Complex Lexical Units and Automata. In Ranchhod, Elisabete M.; N. Mamede (eds.) *Advances in Natural Language Processing*, LNAI 2389 (Heidelberg: Springer), 229-238.
- Carvalho, P. and Ranchhod, E. (2003), Analysis and Disambiguation of Nouns and Adjectives in Portuguese by FST. In *Proceedings of the Workshop on Finite State Methods in Natural Language Processing*, EAACL'03, 105-112 .
- Gross, M. (1986), Lexicon-grammar. The representation of compound words. In *Proceedings of COLING '86* (Bonn: University of Bonn).
- Mota, C.; Carvalho, P.; Ranchhod, E. (2004) Multiword Lexical Acquisition and Dictionary Formalization. In *Proceedings of the Workshop Enhancing and Using Electronic Dictionaries, Coling 2004*, Geneva, Switzerland, 73-77.

## Bibliografia (2)

- Ranchhod, Elisabete M. (1991), Frozen Adverbs. Comparative Forms como C in Portuguese, *Linguisticae Investigationes*, XV: 1, 141-170.
- Ranchhod, Elisabete M. (2001), O uso de dicionários e de autómatos finitos na representação lexical das línguas naturais. In Ranchhod, Elisabete M. (org.), *Tratamento das Línguas por Computador. Uma Introdução à Linguística Computacional e suas Aplicações* (Lisboa: Caminho), 13-47.
- Ranchhod, Elisabete M. M. (2003), O Lugar das Expressões Fixas na Gramática do Português. In Castro, Ivo e I. Duarte (Orgs.), *Razões e Emoção. Miscelânea de estudos oferecida a Maria Helena Mira Mateus*, pp. 239-254, Lisboa: Imprensa Nacional - Casa da Moeda.
- Ranchhod, Elisabete M.; Carvalho, P.; Mota, C.; Barreiro, A. (2004), Portuguese Large-scale Language Resources for NLP Applications. In *Proceedings of the IV Conference on Language Resources and Evaluation, LREC*, Lisboa, 1755-1759.
- Ranchhod, Elisabete M. (2005), Using Corpora to Increase Portuguese MWU Dictionaries. Tagging MWU in a Portuguese Corpus. In *Proceedings from the Corpus Linguistics Conference Series, Vol. 1, no. 1, ISSN 1747-9398*.  
[<http://www.corpus.bham.ac.uk/PCLC/RanchhodCorpora.pdf>]