

# Documentação UD em português (e para língua portuguesa)

Elvis de Souza

PUC-Rio, Brasil

Tatiana Cavalcanti

Aline Silveira

Wograine Evelyn

Cláudia Freitas

O projeto Universal Dependencies ([mcdonald2013universal](#)) apresenta um tagset e uma gramática. Isso significa dizer que, para além de um conjunto de etiquetas que correspondem às classes da Gramática Tradicional (objeto, sujeito etc.), o UD também faz diversas escolhas que diferem da GT. Nesse documento, apresentamos a documentação detalhadas e as escolhas linguísticas relativas ao processo de revisão do material UD em Português. Considerando que UD funciona como uma espécie de segunda língua gramatical, partimos, sempre que possível, das categorias e análises de GT, e não de UD. Os exemplos de frases e as listas foram retirados do corpus [Bosque-UD](#) ([rademaker2017universal](#)) versão 2.5.



# Conteúdo

## Documentação UD em português

### (e para língua portuguesa)

Elvis de Souza, Tatiana Cavalcanti, Aline Silveira, Wograine Evelyn,  
Cláudia Freitas

1

## 2 Formato UD

5

- 1 Colunas/anotações . . . . . 5
- 2 Manipulação em Python . . . . . 6

## 3 Classes gramaticais (upos)

7

- 1 Verbos auxiliares . . . . . 7
  - 1.1 Verbos de ligação . . . . . 8
  - 1.2 Verbo *ser* como voz passiva . . . . . 8
  - 1.3 Locuções verbais de tempo composto . . . . . 9
  - 1.4 Locuções verbais aspectuais/modais . . . . . 9
  - 1.5 Isso *foi* nos Estados Unidos: verbo *ser* como verbo pleno 10
  - 1.6 Não são locuções verbais: dois verbos plenos . . . . . 12
- 2 Formas participiais: verbo ou adjetivo? . . . . . 12
  - 2.1 Adjetivo . . . . . 12
  - 2.2 Verbo . . . . . 13
  - 2.3 Casos difíceis . . . . . 14
- 3 Numerais . . . . . 19
  - 3.1 Números compostos . . . . . 22
  - 3.2 *R\$ 50*: quem é o pai? . . . . . 22
  - 3.3 *50% das pessoas*: quem é o pai? . . . . . 24
  - 3.4 *Primeiro* lugar: adjetivo ou numeral? . . . . . 25

## 4 Atributos morfológicos (feats)

27

## 5 Dependências (dephead e deprel)

31

- 1 Sujeito da oração . . . . . 31
  - 1.1 Caso seja nominal . . . . . 31

1.2	Caso seja oracional . . . . .	31
2	Adjuntos adverbiais . . . . .	33
2.1	Caso seja advérbio . . . . .	33
2.2	Caso seja nome . . . . .	34
2.3	Caso seja verbo . . . . .	36
2.4	<i>Teve uma relação extraconjugal com Monica Lewinsky</i> . .	37
3	Voz passiva . . . . .	37
4	Estruturas comparativas . . . . .	37
4.1	Frases do Working Group . . . . .	38
5	Particípio . . . . .	38
6	Elementos discursivos . . . . .	39
6.1	É que . . . . .	39
6.2	Tal como / Como . . . . .	40
7	Conjunções . . . . .	42
7.1	Conjunções coordenativas . . . . .	42
7.2	Conjunções subordinativas . . . . .	43
7.3	Advérbios interrogativos . . . . .	43
8	Preposições . . . . .	43
9	Estruturas comparativas . . . . .	43
9.1	Frases do Working Group . . . . .	43
9.2	Frases do Bosque-UD . . . . .	44

## 2 Formato UD

[Ir para tabela de conteúdos](#)

Os treebanks adaptados para a gramática UD são disponibilizados no formato CoNLL, em que há um token por linha. Cada anotação de cada token, por sua vez, é disposta em uma coluna, sendo 10 colunas ao todo. Cada token tem a configuração conforme a [Tabela 1: Colunas do formato UD 2.0](#), com uma tabulação (*Tab*) separando as colunas. Colunas sem nenhum valor devem, necessariamente, ser preenchidas com *underline*.

Tabela 1: Colunas do formato UD 2.0

id	word	lemma	upos	xpos	feats	dephead	deprel	deps	misc
----	------	-------	------	------	-------	---------	--------	------	------

### 1 Colunas/anotações

1. “id” corresponde ao número do token, em ordem crescente;
2. “word”, à palavra tal como aparece na frase (exceto no caso de contração, como “da”, em que a palavra será desmembrada nos tokens “de” e “a”);
3. “lemma” se refere à palavra tal como aparece no dicionário: em no singular e em masculino ou infinitivo;
4. “upos” (classe gramatical “universal”) se refere à classe gramatical;
5. No corpus Bosque-UD, a coluna “xpos” (classe gramatical específica) é preenchida com a saída do sistema PALAVRAS para a mesma frase;
6. “feats” (atributos morfológicos) é preenchida com as características morfológicas do token;
7. “dephead” (dependência sintática), com o id do token de quem é filho;

8. “deprel” (relação de dependência), com a relação sintática que o conecta ao seu pai;
9. “deps” (dependência específica) não é utilizado no Bosque-UD;
10. “misc” (miscelânea) se refere a quaisquer informações extras que desejemos adicionar ao token.

## **2 Manipulação em Python**

Para manipular arquivos no formato UD em Python, com as classes `Corpus`, `Sentence` e `Token` (e suas respectivas anotações), desenvolvemos e utilizamos o `es-estrutura_ud.py`.

### 3 Classes gramaticais (upos)

Ir para tabela de conteúdos

Tabela 2: As classes gramaticais do UD em português

Tabela 2: As classes gramaticais do UD em português

upos	Observações
ADJ	adjetivos e numerais ordinais
ADP	preposições
PUNCT	pontuação
ADV	advérbio
AUX	Seção 1: Verbos auxiliares
SYM	símbolos
INTJ	interjeição
CCONJ	conjunção coordenativa
NOUN	substantivo
DET	determinante - artigos e pronomes adjetivos
PROPN	nomes próprios, apenas se com inicial maiúscula
NUM	numeral - exceto os ordinais, que são adjetivos
PART	partícula
VERB	verbo
PRON	apenas pronomes substantivos
SCONJ	conjunções subordinativas
X	no Bosque-UD, palavras estrangeiras

#### 1 Verbos auxiliares

Verbos auxiliares são classificados como *AUX*. O que conta como um verbo auxiliar é alvo de discussão nas gramáticas do português (elvis2019locverbal). De modo geral, classificamos como verbos auxiliares os verbos de ligação (Subseção 1.1: Verbos de ligação), o verbo auxiliar na voz passiva (Subseção 1.2: Verbo

*ser* como voz passiva), as locuções de tempo composto (Subseção 1.3: Locuções verbais de tempo composto), e as locuções verbais aspectuais e modais (Subseção 1.4: Locuções verbais aspectuais/modais).

Via de regra, verbos auxiliares (*AUX*) não podem ter dependentes, e dependem de um verbo principal (*VERB*). O *deprel* pode ser *cop*, *aux* ou *aux:pass*, de acordo com a função sintática.

Tabela 4: Lista das 23 palavras que ocorrem 3541 vezes como verbo auxiliar (*AUX*) no Bosque-UD

1.1 Verbos de ligação

Apenas os verbos “*ser*” e “*estar*” são considerados verbos de ligação, e portanto serão sempre anotados como *AUX*. Os demais verbos que a GT costuma elencar como verbo de ligação (*parecer*, *permanecer*, etc.) são anotados como *VERB*. Os verbos de ligação *AUX* terão relação sintática *cop*, e nunca poderão ser núcleo de uma oração (*Xx*) nem conter dependentes.

Figura 1: O preço é de US\$ 422

Tabela 5: Lista das 2 palavras que ocorrem 138 vezes como verbo de ligação (*AUX/cop*) no Bosque-UD

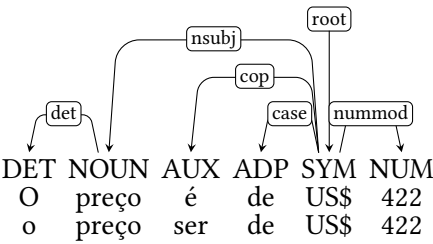


Figura 1: O preço é de US\$ 422

1.2 Verbo *ser* como voz passiva

A anotação de “*ser*” como voz passiva, além do upos *AUX*, deve receber *deprel aux:pass*.

Figura 2: A fotografia foi publicada na imprensa

Tabela 6: Lista das 2 palavras que ocorrem 1115 vezes como voz passiva (*AUX/aux:pass*) no Bosque-UD



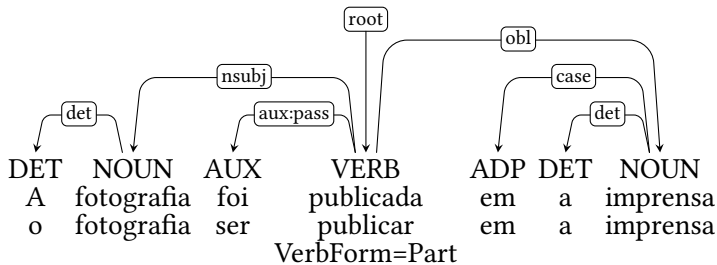


Figura 2: A fotografia *foi* publicada na imprensa

### 1.3 Locuções verbais de tempo composto

Segundo as gramáticas, são locuções verbais de tempo composto aquelas que têm como verbo auxiliar “ter”, “haver” e, para nós, também “ir”.

Figura 3: A Prefeitura não *havia retirado* o golfinho

Tabela 7: Lista das 23 palavras que ocorrem 2256 vezes como locução verbal (AUX/aux) no Bosque-UD

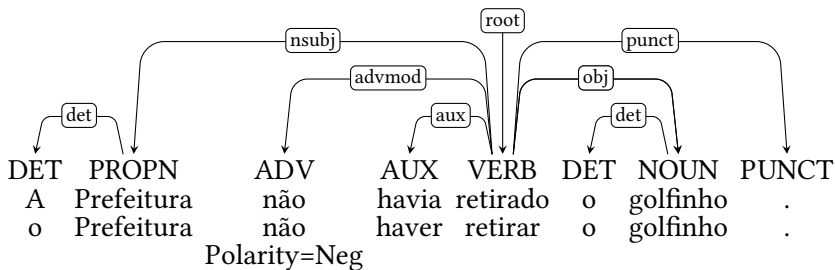


Figura 3: A Prefeitura não *havia retirado* o golfinho

### 1.4 Locuções verbais aspectuais/modais

Para as gramáticas, é um exemplo de locução verbal aspectual a Figura 5: O Tribunal *vai começar a ouvir as testemunhas*, e para algumas gramáticas (e não para outras), é um exemplo de locução modal a Figura 4: A seleção *deve contar* hoje com Giovane.

**Tabela 7: Lista das 23 palavras que ocorrem 2256 vezes como locução verbal (AUX/aux) no Bosque-UD**

Segundo as gramáticas, o que distingue a locução verbal aspectual da modal é que, na primeira, o verbo auxiliar caracteriza a temporalidade da ação do verbo principal, e na segunda, o verbo auxiliar caracteriza um julgamento do enunciador sobre a ação do verbo principal. Alguns gramáticos divergem sobre como a classificação se dá, portanto apenas herdamos, no Bosque-UD, o que já era a anotação originária do PALAVRAS (bick2000parsing), e resolvemos não modificá-la.

Nos casos de locução verbal aspectual, geralmente há uma partícula interveniente no meio da locução, como na **Figura 5: O Tribunal vai começar a ouvir as testemunhas**, em que há um “a” entre “começar” e “ouvir”. Encaramos que estamos diante de um fenômeno de *phrasal verb*, uma MWE do tipo AUX, como sugerido em **elvis2019locverbal**. Na frase, *vai ouvir* é uma locução verbal de tempo composto (**Subseção 1.3: Locuções verbais de tempo composto**), e “começar a ouvir” é uma locução verbal aspectual, sendo “começar a” uma expressão multi-palavra (MWE) do tipo auxiliar.

**Tabela 8: Lista das 31 MWEs que ocorrem 570 vezes como locuções auxiliares (AUX/aux + ADP/compound) no Bosque-UD**

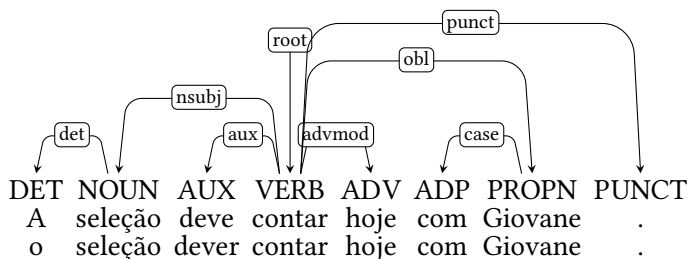


Figura 4: A seleção *deve contar* hoje com Giovane

## 1.5 Isso foi nos Estados Unidos: verbo *ser* como verbo pleno

Atenção para 2 casos em que o “ser” não é verbo auxiliar e deve ser verbo pleno (upos *VERB*).

1) Como na **Figura 6: A expectativa era que chegasse a US\$7 milhões**, o “ser” deve ser núcleo da oração caso o predicado seja uma outra oração (*ccomp*, *xcomp*).

2) “ser” verbo intransitivo (verbo pleno) também deve ter a anotação *VERB*, como na **Figura 7: Isso foi nos Estados Unidos**.

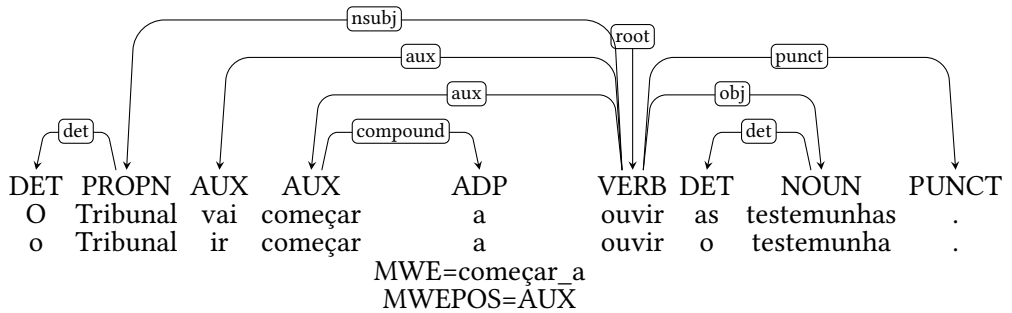


Figura 5: O Tribunal vai *começar a ouvir* as testemunhas

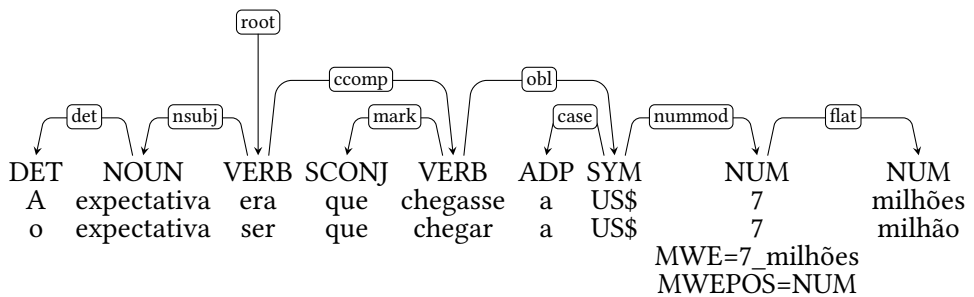


Figura 6: A expectativa *era* que chegasse a US\$7 milhões

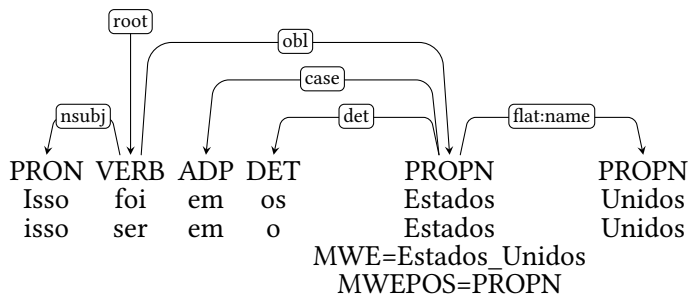


Figura 7: Isso *foi* nos Estados Unidos

## 1.6 Não são locuções verbais: dois verbos plenos

Ainda segundo as decisões do PALAVRAS (**bick2000parsing**), há outras tantas colocações de dois verbos que não são locuções verbais. Portanto, são todos casos de verbos plenos, de upos *VERB*, em que o segundo verbo é dependente do primeiro, como oração substantiva objetiva reduzida de infinitivo (Xx).

Figura 8: Muitos pintores *quiseram aprender a pintar com ele*

Tabela 9: Lista dos 196 verbos plenos que ocorrem 1150 vezes como *pais* de uma colocação verbal no Bosque-UD

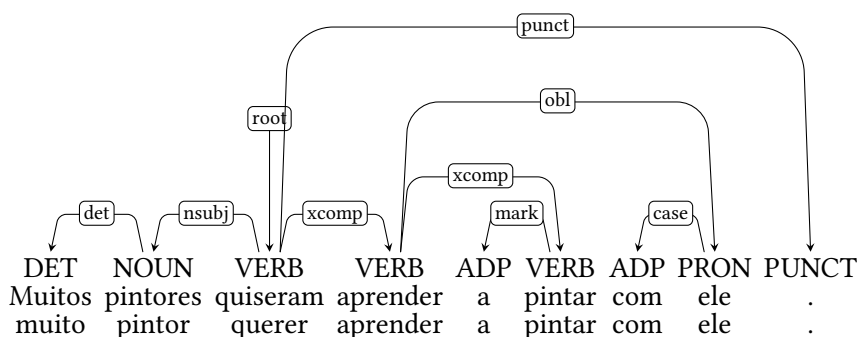


Figura 8: Muitos pintores *quiseram aprender a pintar com ele*

## 2 Formas participiais: verbo ou adjetivo?

A discussão sobre o status das formas participiais, frequentemente ambíguas entre *ADJ* ou *VERB*, é reconhecida na literatura linguística, e não é um fenômeno específico do Português (ver, por exemplo, PROPOR 2018 e Dissert Luiza).

### 2.1 Adjetivo

No processo de anotação do Bosque-UD, anotamos como *ADJ*:

1. Palavras que, ainda que tenham um verbo correlato, possuem um sentido bem demarcado como adjetivo. Como essa decisão nem sempre é clara, listam-se alguns casos considerados *ADJ*.

Figura 9: Cabe ao empresário, ao empreendedor, seja pequeno, médio ou grande, criar o ambiente *adequado* para a valorização de um recurso humano com essas qualificações

Figura 10: A ideia de cinema *engajado* está por baixo

1	Cabe	caber	VERB	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	0	root	-	-
2-3	ao	-	-	-	-	-	-	-	-
2	a	a	ADP	-	-	4	case	-	-
3	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	4	det	-	-
4	empresário	empresário	NOUN	-	Gender=Masc Number=Sing	1	obj	-	SpaceAfter=No
5	,	,	FUNCT	-	-	8	punct	-	-
6-7	ao	-	-	-	-	-	-	-	-
6	a	a	ADP	-	-	8	case	-	-
7	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	8	det	-	-
8	empreendedor	empreendedor	NOUN	-	Gender=Masc Number=Sing	4	nmod	-	SpaceAfter=No
9	,	,	FUNCT	-	-	8	punct	-	-
10	seja	ser	AUX	-	Mood=Sub Number=Sing Person=3 Tense=Pres VerbForm=Fin	-	-	-	-
11	pequeno	pequeno	ADJ	-	Gender=Masc Number=Sing	4	acl:relcl	-	SpaceAfter=No
12	,	,	FUNCT	-	-	13	punct	-	-
13	médio	médio	ADJ	-	Gender=Masc Number=Sing	11	conj	-	-
14	ou	ou	CCONJ	-	-	15	cc	-	-
15	grande	grande	ADJ	-	Gender=Masc Number=Sing	11	conj	-	SpaceAfter=No
16	,	,	FUNCT	-	-	17	punct	-	-
17	criar	criar	VERB	-	VerbForm=Inf	1	csubj	-	-
18	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	19	det	-	-
19	ambiente	ambiente	NOUN	-	Gender=Masc Number=Sing	17	obj	-	-
20	adequado	adequado	ADJ	-	Gender=Masc Number=Sing	19	amod	-	-

Figura 9: Cabe ao empresário, ao empreendedor, seja pequeno, médio ou grande, criar o ambiente *adequado* para a valorização de um recurso humano com essas qualificações

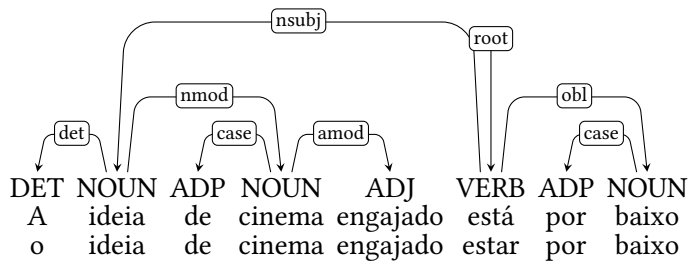


Figura 10: A ideia de cinema *engajado* está por baixo

- 2. Formas participiais que estejam em coordenação com outros adjetivos.

Figura 11: Este futebol *automatizado* e excessivamente defensivo

Figura 12: Foi uma decisão muito *ponderada*, muito difícil e muito amadurecida

2.2 Verbo

No processo de anotação do Bosque-UD, anotamos como *VERB* (deprel *acl*):

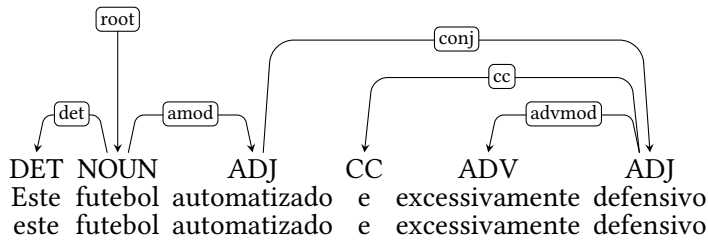


Figura 11: Este futebol *automatizado* e excessivamente defensivo

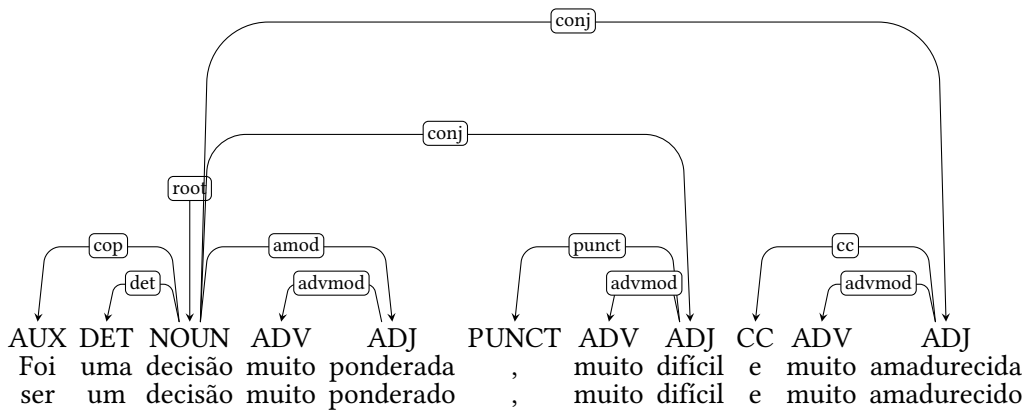


Figura 12: Foi uma decisão muito *ponderada*, muito difícil e muito amadurecida

1. Construções em que se pode incluir um agente da passiva sem prejudicar o sentido da frase.

Figura 13: Resta portanto saber se as medidas *anunciadas*, em tese adequadas, terão na prática o efeito que se espera.

Figura 14: Os serviços mais *procurados* são correio eletrônico e transferência de arquivos

### 2.3 Casos difíceis

1. Os seguintes casos foram anotados como adjetivo:

1	Resta	restar	VERB	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	0	root	-
2	portanto	portanto	-	-	ADV	1	advmod	-
3	saber	saber	VERB	-	VerbForm=Inf	1	csubj	-
4	se	se	SCONJ	-	-	13	mark	-
5	as	o	DET	-	Definite=Def Gender=Fem Number=Plur PronType=Art	6	det	-
6	medidas	medida	NOUN	-	Gender=Fem Number=Plur	13	nsubj	-
7	anunciadas	anunciar	VERB	-	Gender=Fem Number=Plur VerbForm=Part	6	acl	-
8	,	,	PUNCT	-	-	10	punct	-
9	em	em	ADP	-	-	10	case	-
10	tese	tese	NOUN	-	Gender=Masc Number=Sing	6	nmod	-
11	adequadas	adequado	ADJ	-	Gender=Fem Number=Plur	10	amod	-
12	,	,	PUNCT	-	-	10	punct	-
13	terão	ter	VERB	-	Mood=Ind Number=Plur Person=3 Tense=Fut VerbForm=Fin	3	ccomp	-
14-15	na	-	-	-	-	-	-	-
14	em	em	ADP	-	-	16	case	-
15	a	o	DET	-	Definite=Def Gender=Fem Number=Sing PronType=Art	16	det	-
16	prática	prática	NOUN	-	Gender=Fem Number=Sing	13	obl	-
17	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	18	det	-
18	efeito	efeito	NOUN	-	Gender=Masc Number=Sing	13	obj	-
19	que	que	PRON	-	Gender=Masc Number=Sing PronType=Rel	21	obj	-
20	se	se	PRON	-	Case=Acc Gender=Unsp PronType=Prs	21	nsubj	-
21	espera	esperar	VERB	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	18	acl:reicl	-
22	.	.	PUNCT	-	-	1	punct	-

Figura 13: Resta portanto saber se as medidas *anunciadas*, em tese adequadas, terão na prática o efeito que se espera.

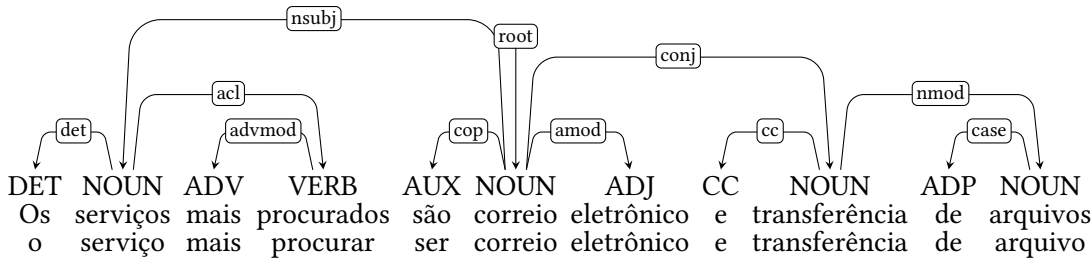


Figura 14: Os serviços mais *procurados* são correio eletrônico e transferência de arquivos

Figura 15: Opinião diferente têm por vezes as autoridades monetárias nacionais, *empenhadas* em fazer passar a ideia de que o escudo não tem sempre de seguir a evolução da moeda espanhola.

Figura 16: Joilson, *casado*, pedia informações sobre o abrigo

Figura 17: O procurador de Justiça *aposentado*

A presença do “mais” e pode auxiliar a análise como um adjetivo, ainda que esse critério nem sempre seja definitivo ou preciso. Exemplos:

Figura 19: Divulgou as acusações sem uma checagem mais *aprofundada*

Figura 20: O mais *cotado* na bolsa de apostas é o brasileiro Raul Boesel, que ontem não conseguiu terminar a corrida em Detroit.

[illegible]

Figura 15: Opinião diferente têm por vezes as autoridades monetárias nacionais, *empenhadas* em fazer passar a ideia de que o escudo não tem sempre de seguir a evolução da moeda espanhola.

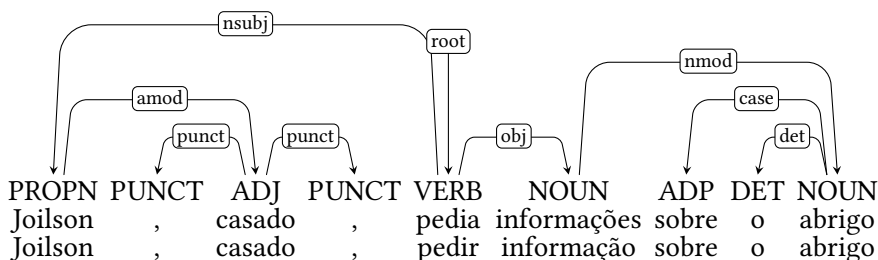


Figura 16: Joilson, *casado*, pedia informações sobre o abrigo

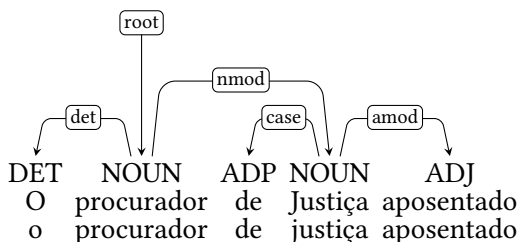


Figura 17: O procurador de Justiça *aposentado*



No exemplo encontrado na **Figura 20: O mais cotado na bolsa de apostas é o brasileiro Raul Boesel, que ontem não conseguiu terminar a corrida em Detroit.**, temos um caso de *ADJ/acl* e não de *ADJ/amod*. Aqui, “cotado” é o predicativo do sujeito de uma oração predicativa na ordem invertida: O brasileiro Raul Boesel é o mais cotado na bolsa de apostas. Como o verbo de ligação nunca é o núcleo da oração, o predicativo recebe a etiqueta de *root*. Anotação similar se encontra na **Figura 18: Chegou atrasado.**

Em frases como a **Figura 21: Um coletivo tão *alargado* e dispersivo**, “tão” funciona como “mais”. Além disso, há o critério da coordenação.

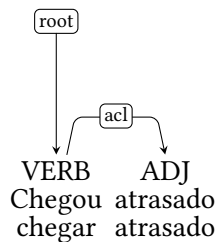


Figura 18: Chegou *atrasado*

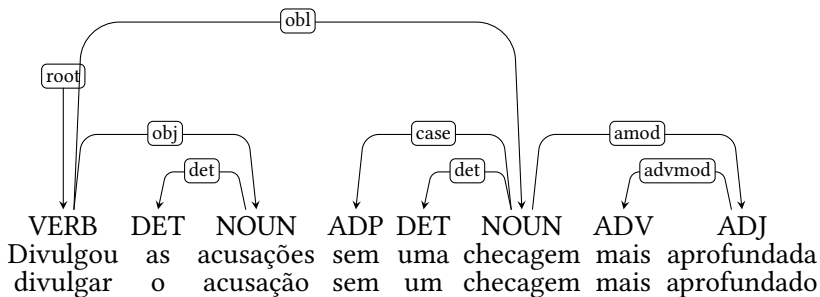


Figura 19: Divulgou as acusações sem uma checagem mais *aprofundada*

2. Os seguintes casos foram analisados como verbo:

Figura 22: O técnico, *irritado* com sua atuação, reclamou.

Figura 23: Utilizaram bactérias conhecidas por devorarem material orgânico

1	O	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	3	det	-	-
2	mais	mais	ADV	-	-	3	advmod	-	-
3	cotado	cotado	ADJ	-	Gender=Masc Number=Sing	0	root	-	-
4-5	na	-	-	-	-	-	-	-	-
4	em	em	ADP	-	-	6	case	-	-
5	a	o	DET	-	Definite=Def Gender=Fem Number=Sing PronType=Art	6	det	-	-
6	bolsa	bolsa	NOUN	-	Gender=Fem Number=Sing	3	obl	-	-
7	de	de	ADP	-	-	8	case	-	-
8	apostas	aposta	NOUN	-	Gender=Fem Number=Plur	6	nmod	-	-
9	é	ser	AUX	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	3	cop	-	-
10	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	11	det	-	-
11	brasileiro	brasileiro	NOUN	-	Gender=Masc Number=Sing	3	nsubj	-	-
12	Raul	Raul	PROPN	-	Gender=Masc Number=Sing	11	appos	-	-
13	Boesel	Boesel	PROPN	-	Number=Sing	12	flat:name	-	SpaceAfter=No
14	,	,	PUNCT	-	-	18	punct	-	-
15	que	que	PRON	-	Gender=Masc Number=Sing PronType=Rel	18	nsubj	-	-
16	ontem	ontem	ADV	-	-	18	advmod	-	-
17	não	não	ADV	-	Polarity=Neg	18	advmod	-	-
18	consegui	conseguir	VERB	-	Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin	12	acl:reicl	-	-
19	terminar	terminar	VERB	-	VerbForm=Inf	18	xcomp	-	-
20	a	o	DET	-	Definite=Def Gender=Fem Number=Sing PronType=Art	21	det	-	-
21	corrida	corrida	NOUN	-	Gender=Fem Number=Sing	19	obj	-	-
22	em	em	ADP	-	-	23	case	-	-
23	Detroit	Detroit	PROPN	-	Gender=Masc Number=Sing	19	obl	-	SpaceAfter=No

Figura 20: O mais *cotado* na bolsa de apostas é o brasileiro Raul Boesel, que ontem não consegui terminar a corrida em Detroit.

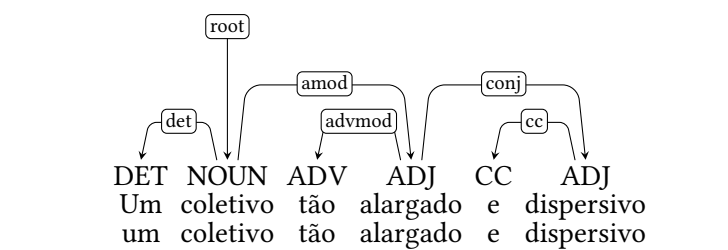


Figura 21: Um coletivo tão *alargado* e dispersivo

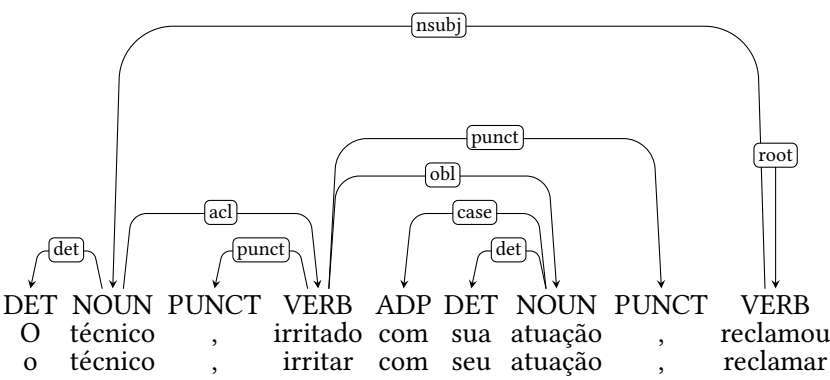


Figura 22: O técnico, *irritado* com sua atuação, reclamou

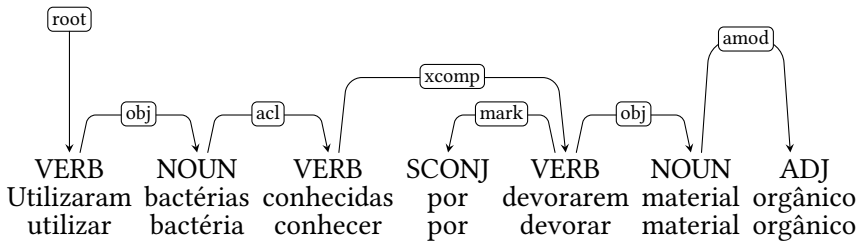


Figura 23: Utilizaram bactérias conhecidas por devorarem material orgânico

3. Há casos em que a palavra pode ser anotada como verbo ou adjetivo, dependendo do contexto. Alguns exemplos são “passado/a”, “elevado/a” e “disposto/a”.

Figura 24: O consumidor poderia optar pelo envio de um cheque sobre o estrangeiro *passado* pelo banco do comprador

Figura 25: Esta senhora senta-se junto dela e começa a recordar-lhe o tempo *passado*

Figura 26: Citam a generosidade dos subsídios de desemprego, o muito *elevado* salário mínimo e as regras de protecção de emprego como factores desmotivadores da criação de postos de trabalho

Figura 27: Os países Ibéricos são os que mostram mais *elevada* taxa de crescimento, 4,5 por cento ao ano

Figura 28: Uma placa de mármore *disposta* sobre um soclo

Figura 29: Tem já muitos candidatos nas suas costas *dispostos* a roubar-lhe o terceiro lugar

### 3 Numerais

Numerais cardinais (360) e fracionários (dois terços) devem ser anotados como de upos *NUM*, sejam por extenso ou em algarismos. Numerais multiplicativos (dobro, triplo), por outro lado, devem ser anotados como de upos *NUM*.

Elvis de Souza, Tatiana Cavalcanti, Aline Silveira, Wograin Evelyn, Cláudia Freitas

1	O	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	2	det	—
2	consumidor	consumidor	NOUN	—	Gender=Masc Number=Sing 4	nsubj	—	—
3	poderia	poder	VERB	—	Mood=Cnd Number=Sing Person=3 VerbForm=Fin	0	root	—
4	optar	optar	VERB	—	VerbForm=Inf 2	xcomp	—	—
5-6	pelo	—	—	—	—	—	—	—
5	por	por	ADP	—	7	case	—	—
6	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	7	det	—
7	envio	envio	NOUN	—	Gender=Masc Number=Sing 4	obj	—	—
8	de	de	ADP	—	10	case	—	—
9	um	um	DET	—	Definite=Ind Gender=Masc Number=Sing PronType=Art	10	det	—
10	cheque	cheque	NOUN	—	Gender=Masc Number=Sing 7	nmod	—	—
11	sobre	sobre	ADP	—	13	case	—	—
12	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	13	det	—
13	estrangeiro	estrangeiro	NOUN	—	Gender=Masc Number=Sing 10	nmod	—	—
14	passado	passar	VERB	—	Gender=Masc Number=Sing 10	acl	—	—
15-16	pelo	—	—	—	—	—	—	—
15	por	por	ADP	—	17	case	—	—
16	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	17	det	—
17	banco	banco	NOUN	—	Gender=Masc Number=Sing 14	obl:agent	—	—
18-19	do	—	—	—	—	—	—	—
18	de	de	ADP	—	20	case	—	—
19	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	20	det	—
20	comprador	comprador	NOUN	—	Gender=Masc Number=Sing 17	nmod	—	—

Figura 24: O consumidor poderia optar pelo envio de um cheque sobre o estrangeiro *passado* pelo banco do comprador

1	Esta	este	DET	—	Gender=Fem Number=Sing PronType=Dem	2	det	—
2	senhora	senhora	NOUN	—	Gender=Fem Number=Sing 3	nsubj	—	—
3-4	senta-se	—	—	—	—	—	—	—
3	senta	sentar	VERB	—	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	0	root	—
4	se	se	PRON	—	Case=Acc Gender=Fem Number=Sing Person=3 PronType=Prs	3	obj	—
5	junto	junto	ADV	—	3	advmod	—	—
6-7	dela	—	—	—	—	—	—	—
6	de	de	ADP	—	7	case	—	—
7	ela	ela	PRON	—	Gender=Fem Number=Sing Person=3 PronType=Prs	5	obl	—
8	e	e	CCONJ	—	11	cc	—	—
9	começa	começar	VERB	—	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	3	conj	—
10	a	a	SCONJ	—	11	mark	—	—
11-12	recordar-lhe	—	—	—	—	—	—	—
11	recordar	recordar	VERB	—	VerbForm=Inf 9	xcomp	—	—
12	lhe	ela	PRON	—	Case=Dat Gender=Masc Number=Sing Person=3 PronType=Prs	11	iobj	—
13	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	14	det	—
14	tempo	tempo	NOUN	—	Gender=Masc Number=Sing 11	obj	—	—
15	passado	passado	ADJ	—	Gender=Masc Number=Sing 14	amod	—	SpaceAfter=No

Figura 25: Esta senhora senta-se junto dela e começa a recordar-lhe o tempo *passado*

1	Citam	citar	VERB	—	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	0	root	—
2	a	o	DET	—	Definite=Def Gender=Fem Number=Sing PronType=Art	3	det	—
3	generosidade	generosidade	NOUN	—	Gender=Fem Number=Sing 1	obj	—	—
4-5	dos	—	—	—	—	—	—	—
4	de	de	ADP	—	6	case	—	—
5	os	o	DET	—	Definite=Def Gender=Masc Number=Plur PronType=Art	6	det	—
6	subsídios	subsídio	NOUN	—	Gender=Masc Number=Plur 3	nmod	—	—
7	de	de	ADP	—	8	case	—	—
8	desemprego	desemprego	NOUN	—	Gender=Masc Number=Sing 6	nmod	—	SpaceAfter=No
9	,	,	PUNCT	—	13	punct	—	—
10	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	13	det	—
11	muito	muito	ADV	—	12	advmod	—	—
12	elevado	elevado	VERB	—	Gender=Masc Number=Sing 13	acl	—	—
13	salário	salário	NOUN	—	Gender=Masc Number=Sing 3	conj	—	—
14	mínimo	mínimo	ADJ	—	Gender=Masc Number=Sing 13	amod	—	—
15	e	e	CCONJ	—	17	cc	—	—
16	as	o	DET	—	Definite=Def Gender=Fem Number=Plur PronType=Art	17	det	—
17	regras	regra	NOUN	—	Gender=Fem Number=Plur 3	conj	—	—
18	de	de	ADP	—	19	case	—	—
19	protecção	protecção	NOUN	—	Gender=Fem Number=Sing 17	nmod	—	—
20	de	de	ADP	—	21	case	—	—
21	emprego	emprego	NOUN	—	Gender=Masc Number=Sing 19	nmod	—	—
22	como	como	ADP	—	23	case	—	—
23	factores	factore	NOUN	—	Gender=Masc Number=Plur 17	nmod	—	—
24	desmotivadores	desmotivador	ADJ	—	Gender=Masc Number=Plur 23	amod	—	—

Figura 26: Citam a generosidade dos subsídios de desemprego, o muito *elevado* salário mínimo e as regras de protecção de emprego como factores desmotivadores da criação de postos de trabalho

## Documentação UD em português (e para língua portuguesa)

1	Os	o	DET	-	Definite=Def Gender=Masc Number=Plur PronType=Art	2	det	-
2	países	país	NOUN	-	Gender=Masc Number=Plur 5 nsubj	-	-	-
3	Ibéricos	ibérico	ADJ	-	Gender=Masc Number=Plur 2 amod	-	-	-
4	são	ser	AUX	-	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	5	cop	-
5	os	o	PRON	-	Gender=Masc Number=Plur PronType=Dem 0 root	-	-	-
6	que	que	PRON	-	Gender=Masc Number=Plur PronType=Rel 7 nsubj	-	-	-
7	mostram	mostrar	VERB	-	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	5	acl:relcl	-
8	mais	mais	ADV	-	9 advmod	-	-	-
9	elevada	elevado	ADJ	-	Gender=Fem Number=Sing 10 amod	-	-	-
10	taxa	taxa	NOUN	-	Gender=Fem Number=Sing 7 obj	-	-	-
11	de	de	ADP	-	12 case	-	-	-
12	crescimento	crescimento	NOUN	-	Gender=Masc Number=Sing 10 nmod	-	-	SpaceAfter=No
13	,	,	PUNCT	-	15 punct	-	-	-
14	4,5	4,5	NUM	-	NumType=Card 15 nummod	-	-	-
15	por	por	ADP	-	Gender=Masc Number=Sing 12 appos	-	-	MWE=por_cento MWEPOS=NOUN
16	cento	cento	NUM	-	Gender=Masc Number=Sing NumType=Mult 15 compound	-	-	-
17-18	ao	-	-	-	-	-	-	-
17	a	a	ADP	-	19 case	-	-	-
18	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	19	det	-
19	ano	ano	NOUN	-	Gender=Masc Number=Sing 12 nmod	-	-	SpaceAfter=No
20	.	.	PUNCT	-	5 punct	-	-	SpaceAfter=No

Figura 27: Os países Ibéricos são os que mostram mais *elevada* taxa de crescimento, 4,5 por cento ao ano

1	Uma	um	DET	-	Definite=Ind Gender=Fem Number=Sing PronType=Art	2	det	-
2	placa	placa	NOUN	-	Gender=Fem Number=Sing 0 root	-	-	-
3	de	de	ADP	-	4 case	-	-	-
4	mármore	mármore	NOUN	-	Gender=Fem Number=Sing 2 nmod	-	-	-
5	disposta	dispor	VERB	-	Gender=Fem Number=Sing 2 acl	-	-	-
6	sobre	sobre	ADP	-	8 case	-	-	-
7	um	um	DET	-	Definite=Ind Gender=Masc Number=Sing PronType=Art	8	det	-
8	soclo	soclo	NOUN	-	Gender=Masc Number=Sing 5 obl	-	-	-

Figura 28: Uma placa de mármore *disposta* sobre um soclo

1	Tem	ter	VERB	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	0	root	-
2	já	já	ADV	-	1 advmod	-	-	-
3	muitos	muito	DET	-	Gender=Masc Number=Plur PronType=Ind 4 det	-	-	-
4	candidatos	candidato	NOUN	-	Gender=Masc Number=Plur 1 obj	-	-	-
5-6	nas	-	-	-	-	-	-	-
5	em	em	ADP	-	8 case	-	-	-
6	as	o	DET	-	Definite=Def Gender=Fem Number=Plur PronType=Art	8	det	-
7	suas	seu	DET	-	Gender=Fem Number=Plur PronType=Prs 8 det	-	-	-
8	costas	costa	NOUN	-	Gender=Fem Number=Plur 1 obl	-	-	-
9	dispostos	disposto	ADJ	-	Gender=Masc Number=Plur 4 amod	-	-	-
10	a	a	ADP	-	11 mark	-	-	-
11-12	roubar-lhe	-	-	-	-	-	-	-
11	roubar	roubar	VERB	-	VerbForm=Inf 9 ccomp	-	-	-
12	lhe	ele	PRON	-	Case=Dat Gender=Masc Number=Sing Person=3 PronType=Prs 11 iobj	-	-	-
13	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	15	det	-
14	terceiro	terceiro	ADJ	-	Gender=Masc Number=Sing NumType=Ord 15 amod	-	-	-
15	lugar	lugar	NOUN	-	Gender=Masc Number=Sing 11 obj	-	-	SpaceAfter=No

Figura 29: Tem já muitos candidatos nas suas costas *dispostos* a roubar-lhe o terceiro lugar

Numerais coletivos (como “dezenas”, “centenas”, “milhares”, etc.), no corpus Bosque-UD, podem estar anotados de duas formas diferentes: como *NUM*, caso seja um número exato (“duas centenas”, “dois bilhões”), ou como *NOUN*, quando é indefinido (“centenas de pessoas”).

Numerais ordinais devem ser anotados como *ADJ* (Subseção 3.4: *Primeiro lugar: adjetivo ou numeral?*).

Ver também: Subseção 3.1: *Números compostos*

### 3.1 Números compostos

Números compostos podem conter upos *NUM*, *CCONJ* e *ADP*. Quando o número composto modifica um substantivo, o primeiro número deve se subordinar ao substantivo como “nummod”; quando há um sintagma preposicionado, o primeiro número do número composto será o head, e o substantivo, *nmod* do número. Além disso, deve-se notar que o primeiro número deve receber no misc *MWE* e *MWEPOS*.

Figura 30: *Trinta e sete*

Figura 31: *Montante de 1300 milhões de dólares*

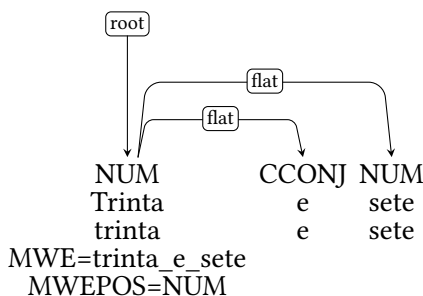


Figura 30: *Trinta e sete*

### 3.2 R\$ 50: quem é o pai?

Quando há a inserção de símbolos como Figura 32: *R\$ 50*, Figura 33: *U\$ 50* ou Figura 34: *50%*, o símbolo deve ser o head. No caso de números coletivos (Figura 35: *2 centenas*), o coletivo será o head.

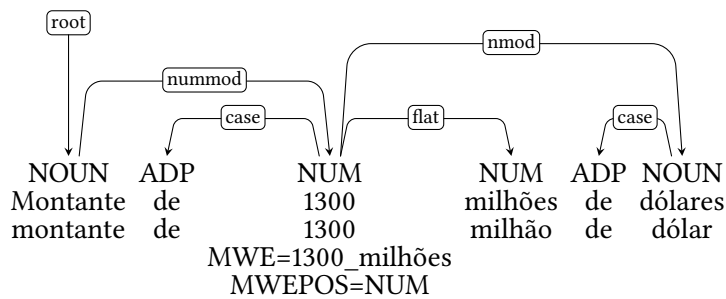


Figura 31: Montante de 1300 milhões de dólares

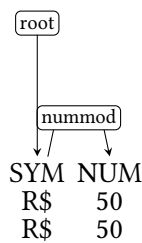


Figura 32: R\$ 50

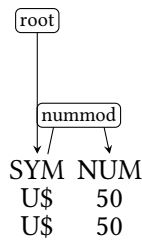


Figura 33: U\$ 50

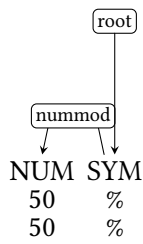


Figura 34: 50%

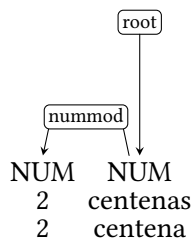


Figura 35: 2 centenas

3.3 50% das pessoas: quem é o pai?

Sempre que houver sintagma preposicionado, o sintagma anterior será o head, como em **Figura 36: 50% das pessoas** (onde o símbolo é o head, segundo **Subseção 3.2: R\$ 50: quem é o pai?**).

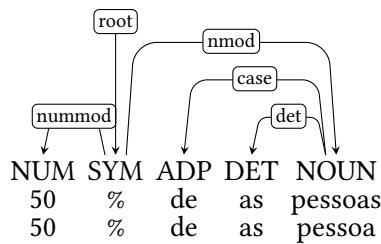


Figura 36: 50% das pessoas



### 3.4 *Primeiro* lugar: adjetivo ou numeral?

Numerais ordinais escritos por extenso devem ser anotados como *ADJ*, e recebem a feature *NumType=Ord*, como na **Figura 37: Primeira tentativa**.

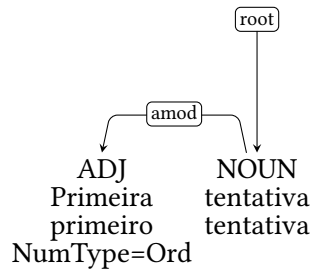


Figura 37: *Primeira tentativa*



# 4 Atributos morfológicos (feats)

Ir para tabela de conteúdos

Temos a seguinte distribuição de atributos morfológicos por classe gramatical (**Tabela 3: Atributos morfológicos (feats)**). É importante notar que os atributos morfológicos devem constar em ordem alfabética e são separados por uma barra reta.

upos	features
ADJ	Gender=[Fem, Masc, Unsp] NumType=[Ord] Number=[Plur, Sing]
ADP	–
ADV	Polarity=[Neg] –
AUX	Gender=[Fem, Masc] Mood=[Cnd, Imp, Ind, Sub] Number=[Plur, Sing] Person=[1, 2, 3] Tense=[Fut, Imp, Past, Pqp, Pres] VerbForm=[Fin, Ger, Inf, Part]
CCONJ	–
DET	Definite=[Def, Ind] Gender=[Fem, Masc, Unsp] Number=[Plur, Sing, Unsp] PronType=[Art, Dem, Emp, Ind, Int, Neg, Prs, Rel, Tot]

INTJ	–
NOUN	Foreign=[Yes] Gender=[Fem, Masc, Unsp] NumType=[Ord] Number=[Plur, Sing, Unsp]
NUM	Gender=[Fem, Masc, Unsp] NumType=[Card, Frac, Mult, Ord, Range, Sets] Number=[Plur, Sing]
PART	Gender=[Masc] Number=[Sing]
PRON	Case=[Acc, Dat, Nom] Definite=[Def, Ind] Gender=[Fem, Masc, Unsp] Number=[Plur, Sing, Unsp] Person=[1, 2, 3] PronType=[Art, Dem, Ind, Int, Neg, Prs, Rel, Tot] Reflex=[Yes] VerbForm=[Ger]
PROPN	Gender=[Fem, Masc, Unsp] Number=[Plur, Sing]
PUNCT	–
SCONJ	Gender=[Fem, Masc] Number=[Plur, Sing] PronType=[Ind, Rel]
SYM	–

VERB	Gender=[Fem, Masc] Mood=[Cnd, Imp, Ind, Sub] Number=[Plur, Sing] Person=[1, 2, 3] Tense=[Fut, Imp, Past, Pqp, Pres] VerbForm=[Fin, Ger, Inf, Part] Voice=[Pass]
------	---

X	—
---	---



# 5 Dependências (dephead e deprel)

Ir para tabela de conteúdos

## 1 Sujeito da oração

### 1.1 Caso seja nominal

O sujeito será anotado com a deprel *nsubj*.

Figura 38: O *McDonald's* proibiu

Figura 39: Surgem *problemas* diariamente

Quando em uma oração na voz passiva, o sujeito deve receber a deprel *nsubj:pass*, como na Figura 52: A URV é produzida por três índices.

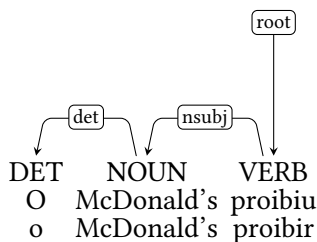


Figura 38: O *McDonald's* proibiu

### 1.2 Caso seja oracional

O sujeito será anotado com a deprel *csubj*, que corresponde a uma oração subordinada substantiva subjetiva.

Figura 40: Será prematuro *falar*

Figura 41: Pareceu estranho nunca ter *lido*

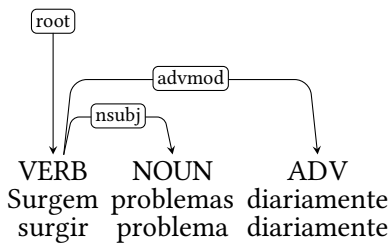


Figura 39: Surgem *problemas* diariamente

Em construções como **Figura 6: A expectativa *era* que chegasse a US\$7 milhões**, “chegasse” não é *csubj*, mas *ccomp*, uma vez que o verbo “ser” é pleno (**Subseção 1.5: Isso *foi* nos Estados Unidos: verbo *ser* como verbo pleno**).

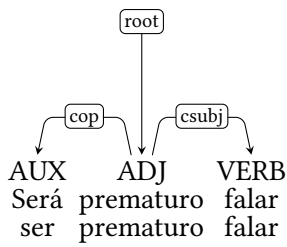


Figura 40: Será prematuro *falar*

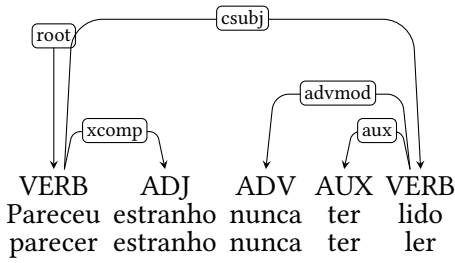


Figura 41: Pareceu estranho nunca ter *lido*



## 2 Adjuntos adverbiais

### 2.1 Caso seja advérbio

O advérbio, que recebe upos *ADV*, será anotado com a deprel *advmod*. Poderá modificar *VERB*, *ADJ* ou *ADV*.

Como exemplo de modificador de *VERB*, observa-se a **Figura 42: A limpeza é feita *diariamente*.**

Como exemplo de modificador de *ADJ*, observa-se a **Figura 43: Sempre tivemos canções *mais lentas*.**

Como exemplo de modificador de *ADV*, observa-se a **Figura 44: *Só assim* se evitará a estratificação da legislação.**

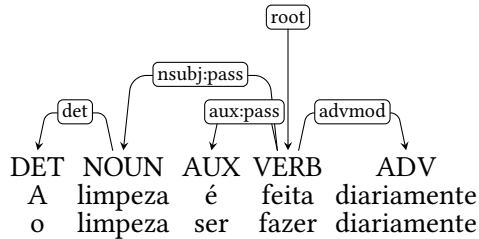


Figura 42: A limpeza é feita *diariamente*

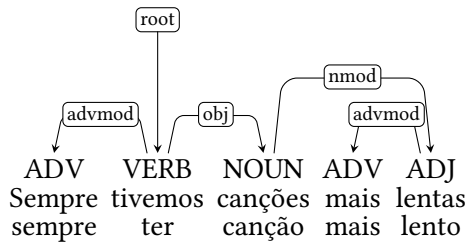


Figura 43: Sempre tivemos canções *mais lentas*

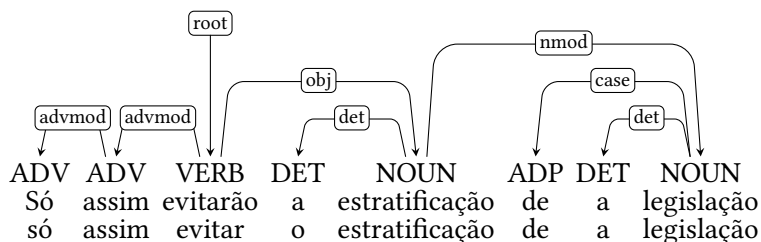


Figura 44: Só assim se evitará a estratificação da legislação

## 2.2 Caso seja nome

O núcleo nominal de um adjunto adverbial (substantivo, pronome, sintagma nominal) será anotado como *obl*. Poderá modificar *VERB*, *ADJ* ou *ADV*.

Como exemplo de modificador de *VERB*, observam-se a **Figura 45: A ideia para o livro surgiu após uma entrevista** e a **Figura 48: Terrenos entregues como pagamento de uma dívida**.

Como exemplo de modificador de *ADJ*, observa-se a **Figura 47: Passaram a exigir funcionários mais abertos ao diálogo**.

Como exemplo de modificador de *ADV*, observa-se a **Figura 46: Independentemente do poder civil**.

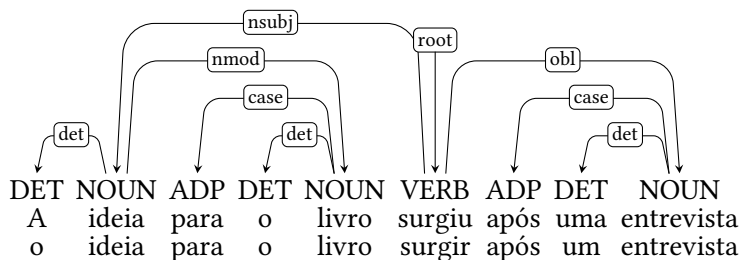


Figura 45: A ideia para o livro surgiu após uma entrevista

A relação de *obl*, como é vista na **Figura 47: Passaram a exigir funcionários mais abertos ao diálogo**, pode ser considerada um argumento de adjetivo. (Ver Xx)

A subrelação *obl:agent* é utilizada para indicar o agente de um verbo na voz

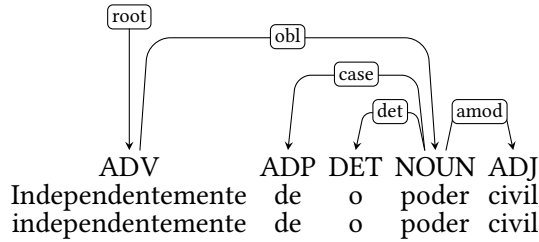


Figura 46: Independentemente do *poder* civil

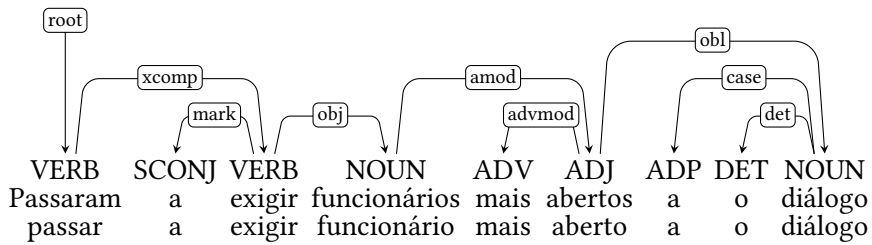


Figura 47: Passaram a exigir funcionários mais abertos ao *diálogo*

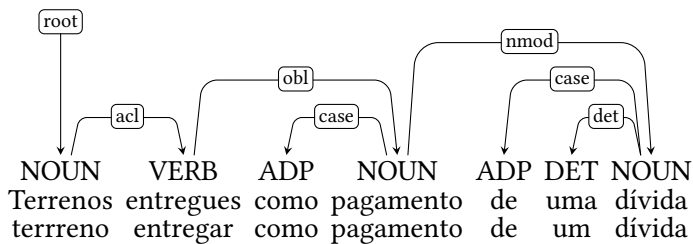


Figura 48: Terrenos entregues como *pagamento* de uma dívida

passiva. (ver **Seção 3: Voz passiva**)

2.3 Caso seja verbo

A oração subordinada adverbial será anotada como *advcl*, não havendo diferenciação semântica (temporal, causal, concessiva, final, etc). Poderá modificar *VERB*.

**Figura 49: Como *notou* o jornalista, é uma boa grana**

**Figura 50: Quando Sarney *começou* a escrever, chegaram protestos**

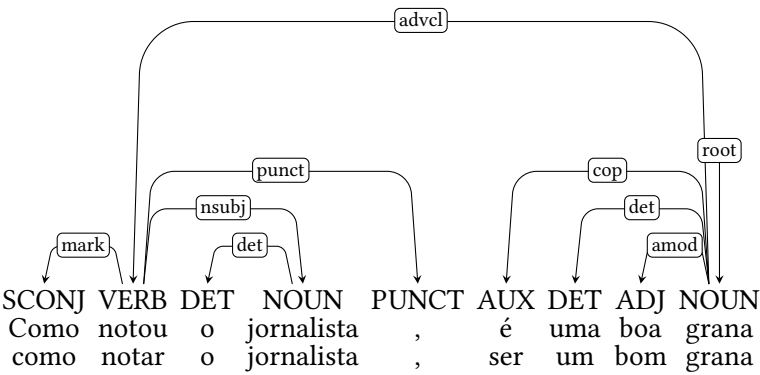


Figura 49: Como *notou* o jornalista, é uma boa grana

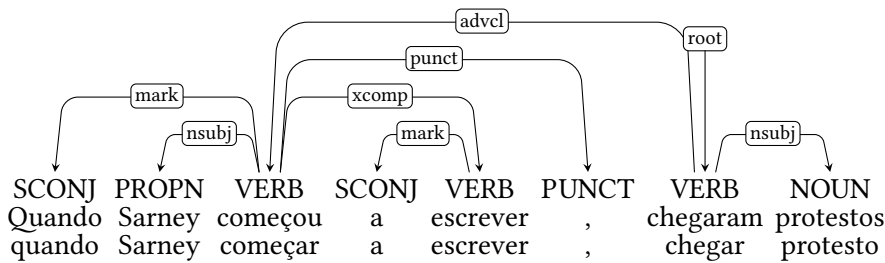


Figura 50: Quando Sarney *começou* a escrever, chegaram protestos

## 2.4 Teve uma relação extraconjugal com Monica Lewinsky

Consideramos que “Monica” apontará para o verbo, e não para o nome. Sendo assim, o pai de “Monica” será “teve” e sua deprel será *obl*.

Figura 51: Teve uma relação extraconjugal com Monica Lewinsky

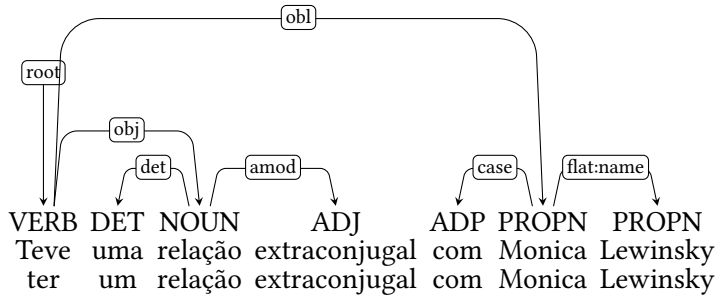


Figura 51: Teve uma relação extraconjugal com Monica Lewinsky

Nas estruturas comparativas, independente de o núcleo do fragmento ser um verbo ou um nome, ele será anotado com a deprel *advcl*. (ver Xx)

## 3 Voz passiva

Na voz passiva, anota-se o sujeito com a deprel *nsubj:pass*, e o agente da passiva, com a deprel *obl:agent*. O verbo “ser”, como explicitado na [Subseção 1.2: Verbo ser como voz passiva](#), recebe upos *AUX* e deprel *aux:pass*.

Figura 52: A URV é produzida por três índices

Figura 53: O trabalho pode ser feito por qualquer pessoa

Tabela 6: Lista das 2 palavras que ocorrem 1115 vezes como voz passiva (*AUX/aux:pass*) no Bosque-UD

## 4 Estruturas comparativas

Estruturas comparativas são de anotação complexa, o que se verifica pela existência de um **working group (WG) em UD** dedicado especialmente a elas. A seguir, listamos as frases utilizadas no WG, traduzidas em português, e com a anotação adequada, além de algumas frases de anotação complexa no Bosque-UD.

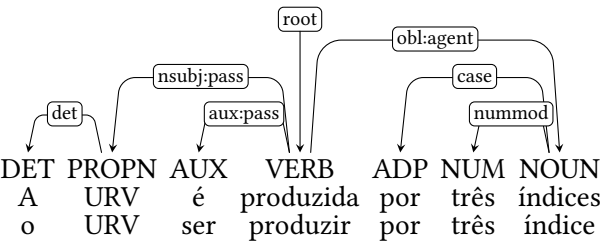


Figura 52: A URV é produzida por três índices

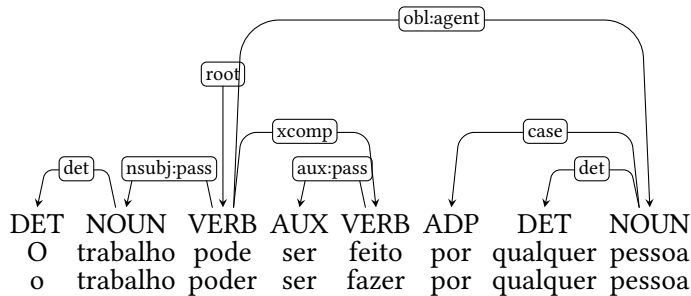


Figura 53: O trabalho pode ser feito por qualquer pessoa

4.1 Frases do Working Group

Eu coloquei tanta farinha quanto a receita pedia.									
1	Eu	eu	PRON	-	Case=Nom Gender=Fem Number=Sing Person=1 Prontype=Prs	2	nsubj	-	-
2	coloquei	colocar	VERB	-	Mood=Ind Number=Sing Person=1 Tense=Past VerbForm=Fin	0	root	-	-
3	tanta	tanto	DET	-	Gender=Fem Number=Sing Prontype=Ind	4	det	-	-
4	farinha	farinha	NOUN	-	Gender=Fem Number=Sing	2	obj	-	-
5	quanto	quanto	SCONJ	-	mark	3	-	-	-
6	a	a	DET	-	Definite=Def Gender=Fem Number=Sing Prontype=Art	7	det	-	-
7	receita	receita	NOUN	-	Gender=Fem Number=Sing	8	nsubj	-	-
8	pedia	pedia	VERB	-	Mood=Ind Number=Sing Person=3 Tense=Imp VerbForm=Fin	2	advcl	-	SpaceAfter=no
9	.	.	PUNCT	-	punct	-	-	-	-

Figura 54: Eu coloquei *tanta farinha quanto* a receita pedia

5 Participípio

Ver **Seção 2: Formas participiais: verbo ou adjetivo?**.

Martin é o cara mais inteligente de todos.									
1	Martin	Martin	PROPN	—	Gender=Masc Number=Sing	4	nsubj	—	—
2	é	ser	AUX	—	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	4	cop	—	—
3	o	o	DET	—	Definite=Def Gender=Masc Number=Sing PronType=Art	4	det	—	—
4	cara	cara	NOUN	—	Gender=Fem Number=Sing	0	root	—	—
5	mais	mais	ADV	—	6	advmod	—	—	—
6	inteligente	inteligente	ADJ	—	Gender=Fem Number=Sing	4	amod	—	—
7	de	de	ADP	—	8	case	—	—	—
8	todos	todo	PRON	—	Gender=Masc Number=Plur PronType=Tot	6	obl	—	SpaceAfter=No
9	.	.	PUNCT	—	4	punct	—	—	—

Figura 55: Martin é o cara *mais inteligente de todos*

## 6 Elementos discursivos

### 6.1 É que

A construção “é que” pode aparecer de duas maneiras: marcada ou não como MWE. Como MWE, “é que” é uma ocorrência típica do discurso oral, sem ligação clara com a estrutura da oração. Por outro lado, “é” e “que” são independentes entre si quando introduzem uma oração subordinada substantiva predicativa ou subjetiva. (ver Xx)

Quando “é que” tem esse efeito discursivo, anota-se da seguinte maneira:

Ê: tem upos *AUX* e deprel *discourse*, e aponta para a head da oração;

Que: tem upos *SCONJ* e deprel *fixed*, e aponta para “é”.

Figura 56: Só depois *é que* levanto a cabeça para fazer um lançamento

Figura 57: A melhor equipe *é que* deve jogar

Figura 58: *Ê que* são uns antipáticos que nunca oferecem copos

Igualmente, consideramos “Não é que” como MWE.

“*Não é que* o sábio das matrizes encontrou, no PSD, entusiastas seguidores?”

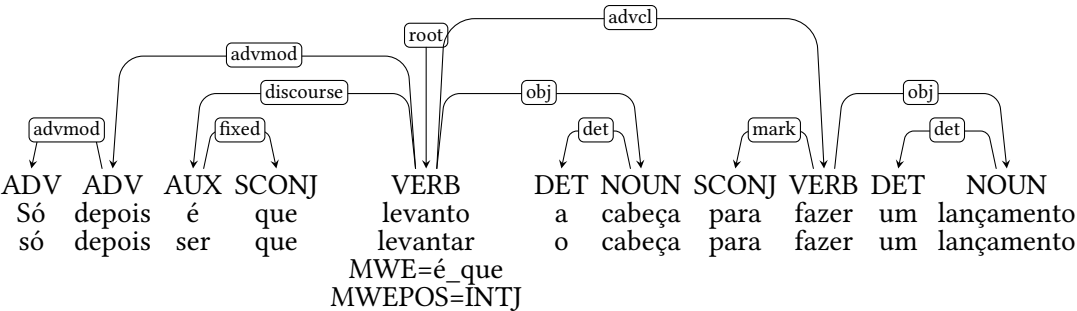


Figura 56: Só depois *é que* levanto a cabeça para fazer um lançamento

1	text =	ê	que	são	uns	antipáticos	que	nunca	oferecem	copos.
2	{	ê	que	é	AUX	discourse				
3		são	que	SCONJ	-	Gender=Mascl Number=Sing	1	fixed	MWE=ê_que	
4		são	ser	AUX	-	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	5	cop	-	-
5		uns	um	DET	-	Gender=Mascl Number=Plur PronType=Art	5	det	-	-
6		antipáticos	antipático	ADJ	-	Gender=Mascl Number=Plur	0	root	-	-
7		que	que	SCONJ	-	8	mark	-	-	-
8		nunca	nunca	ADV	-	8	advmod	-	-	-
9		oferecem	oferecer	VERB	-	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	5	acl:reicl	-	-
10		copos	copo	NOUN	-	Gender=Mascl Number=Plur	8	obj	SpaceAfter=No	SpaceAfter=No
11		.	.	PUNCT	-	5	punct	SpaceAfter=No		

2	O	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	2	det	-
3	fax	fax	NOUN	-	Gender=Masc Number=Sing 4	sub:pass	-	-
4	fui	ser	AUX	-	Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin	4	aux:pass	-
5	enviado	enviar	VERB	-	Gender=Masc Number=Sing VerbForm=Part Voice=Pass	0	root	-
6	tal	tal	PUNCT	-	4	punct	-	-
7	como	como	ADV	-	8	mark	-	MWEPOS:SCONJ
8	pedido	pedir	SCONJ	-	6	fixed	-	-
9	pedido	pedir	VERB	-	Gender=Masc Number=Sing VerbForm=Part	4	advcl	-
10-11	á	á	PUNCT	-	8	punct	-	SpaceAfter=No
12	a	a	ADP	-	-	-	-	-
13	a	a	DET	-	-	-	-	-
14	presidente	presidente	NOUN	-	Definite=Def Gender=Fem Number=Sing PronType=Art	12	det	-
15	do	de	ADP	-	NOUN	Gender=Fem Number=Sing 4	obj	-
16	do	de	DET	-	-	-	-	-
17	Instituto	instituto	PUNCT	-	-	-	-	-
18	Zita	Zita	PROPN	-	Definite=Def Gender=Masc Number=Sing PronType=Art	15	det	-
19	Seabra	Seabra	PROPN	-	Gender=Masc Number=Sing 12	nmod	-	SpaceAfter=No
20	Seabra	Seabra	PUNCT	-	17	punct	-	-
21	.	.	PROPN	-	Gender=Fem Number=Sing 12	appos	-	SpaceAfter=No
22	.	.	PUNCT	-	Number=Sing 17	flat:name	-	SpaceAfter=No
23	.	.	PUNCT	-	4	punct	-	SpaceAfter=No

40



[illegible]

Figura 60: A zona lúdica, com os divertimentos, áreas comerciais de «souvenirs» e de restauração, compreende espaços distintos para os vários temas, ainda em análise, *tais como* Portugal, Japão, Brasil, África e Far-West.

“conforme”, “tal como” foi anotado como MWE com valor de conjunção subordinativa.

Essa MWE com valor de conjunção subordinativa inicia uma oração adverbial; logo, o núcleo do fragmento, independente de ser verbo ou não, teve deprel anotado como *advcl*. (Figura 61: *Tal como a rainha Isabel II, John Major limitou-se a acompanhar os factos.*)

1	Tal	tal	ADV	-	-	4	mark	-	MWEPOS=SCOMJ		
2	como	como	SCONJ	-	-	1	fixed	-			
3	a	o	DET	-	-			Definite=Def Gender=Fem Number=Sing PronType=Art		4	det
4	rainha	rainha	NOUN	-	-			Gender=Fem Number=Sing	10	advcl	
5	Isabel	Isabel	PROPN	-	-			Gender=Masc Number=Sing	4	appos	
6	II	II	PROPN	-	-			Number=Sing	5	flat:name	
7	,	,	PUNCT	-	-	1	punct				
8	John	John	PROPN	-	-			Gender=Masc Number=Sing	10	nsubj	
9	Major	Major	PROPN	-	-			Number=Sing	8	flat:name	
10-11	limitou-se										
10	limitou	limitar	VERB	-	-			Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin	0		root
11	se	se	PRON	-	-			Case=Acc Gender=Masc Number=Sing Person=3 PronType=Prs	10		expl
12	a	a	ADP	-	-	13	mark				
13	acompanhar	acompanhar						VerbForm=Inf	10	xcomp	
14	os	o	DET	-	-			Definite=Def Gender=Masc Number=Plur PronType=Art		15	det
15	factos	facto	NOUN	-	-			Gender=Masc Number=Plur	13	obj	
16	.	.	PUNCT	-	-	10	punct			SpaceAfter=No	

Figura 61: *Tal como a rainha Isabel II, John Major limitou-se a acompanhar os factos.*

Para os casos em que “tal como” indica hiperonímia, criamos uma MWE com valor prepositivo (Figura 60: A zona lúdica, com os divertimentos, áreas comerciais de «souvenirs» e de restauração, compreende espaços distintos para os vários temas, ainda em análise, *tais como* Portugal, Japão, Brasil, África e Far-West.).

Além disso, essa estrutura de hiperonímia também pode aparecer em frases com apenas “como”, em vez de “tal como”. Essas frases, em termos semânticos, têm valores semelhantes; no entanto, não há MWE nas frases de “como”. (Figura 62: Abstratos entre medianos e medíocres, *como* Fukushima, Pêrsio, Raimo e Douchez, têm o mesmo ou maior destaque que Volpi e nada que se possa chamar de autonomia para oferecer como lenitivo.)

1	Abstratos	abstrato	NOUN	3	Gender=Masc Number=Plur 16	nsubj	-	-
2	entre	ADP	-	3	case	-	-	-
3	medianos	mediano NOUN	-	1	Gender=Masc Number=Plur 1	nmod	-	-
4	e	CCONJ	-	5	cc	-	-	-
5	medíocres	mediocre	NOUN	3	Gender=Masc Number=Plur 3	conj	-	SpaceAfter=No
6	,	PUNCT	-	8	punct	-	-	-
7	como	ADP	-	8	case	-	-	-
8	Fukushima	Fukushima	PROPN	1	Gender=Masc Number=Sing 1	appos:parataxis	-	SpaceAfter=No
9	,	PUNCT	-	10	punct	-	-	-
10	Pêrsio	Pêrsio	PROPN	8	Gender=Masc Number=Sing 8	conj	-	SpaceAfter=No
11	,	PUNCT	-	12	punct	-	-	-
12	Raimo	Raimo	PROPN	8	Gender=Masc Number=Sing 8	conj	-	-
13	e	CCONJ	-	14	cc	-	-	-
14	Douchez	Douchez	PROPN	8	Gender=Masc Number=Sing 8	conj	-	SpaceAfter=No
15	,	PUNCT	-	8	punct	-	-	-
16	têm	ter	VERB	0	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin 0	root	-	-
17	o	DET	-	21	Definite=Def Gender=Masc Number=Sing PronType=Art 21	det	-	-
18	mesmo	mesmo	ADJ	21	Gender=Masc Number=Sing 21	amod	-	-
19	ou	CCONJ	-	20	cc	-	-	-
20	maior	grande	ADJ	18	Gender=Masc Number=Sing 18	conj	-	-
21	destaque	destaque	NOUN	16	Gender=Masc Number=Sing 16	obj	-	-

Figura 62: Abstratos entre medianos e medíocres, *como* Fukushima, Pêrsio, Raimo e Douchez, têm o mesmo ou maior destaque que Volpi e nada que se possa chamar de autonomia para oferecer como lenitivo.

## 7 Conjunções

Há dois tipos de conjunções no Bosque-UD: conjunção coordenativa (CCONJ) e conjunção subordinativa (SCONJ). As funções que apresentam são cc (coordenador) e mark (subordinador). As locuções coordenativas e subordinativas frequentemente aparecem como expressões multi palavras (MWE) e têm a mesma função que as conjunções coordenativas (cc/mark), além da mesma classe gramatical (CCONJ/SCONJ).

### 7.1 Conjunções coordenativas

As conjunções coordenativas que estão marcadas como tal no corpus são *e*, *mas*, *ou* *nem*, *porém*, *entretanto*, *quer*, *portanto*, *contudo*, *todavia*, *tampouco*. Além disso, existem as conjunções que aparecem como expressões multi palavras (MWE), são elas: *no entanto*, *ou seja*, *além disso*, *por outro lado*, *ainda assim*, *isto é*, *em vez de*, *desse modo*, *apesar disso*, *sendo assim*.

Além disso, também são consideradas conjunções casos em que não há uma oração. Toda vez que houver a coordenação entre dois ou mais termos, sejam eles

oracionais ou nominais, o conectivo entre eles será classificado como conjunção. Exemplo:

## 7.2 Conjunções subordinativas

As conjunções subordinativas que estão marcadas como tal no corpus são . Além disso, existem as conjunções que aparecem como expressões multi palavras(MWE), são elas: *a fim de que, à medida que, a partir de, a favor de, a não ser que, ao passo que, ainda que, sem que, já que.*

É importante lembrar que palavras tidas usualmente como preposição, ao introduzir orações, são classificadas como conjunções subordinativas e recebem a deprel mark. (Ver seção Preposições)

## 7.3 Advérbios interrogativos

Advérbios são palavras que costumam modificar verbos, adjetivos e outros advérbios. Além disso, em português, algumas dessas palavras podem operar como conjunções subordinativas quando aparecem na introdução de orações subordinadas adverbiais. Exemplo:

# 8 Preposições

As preposições são palavras que operam como conectores, geralmente, de termos nominais. Entretanto, no Bosque-UD, palavras tidas usualmente como preposições podem operar como conjunções subordinativas quando em contexto de introdução de oração. Exemplo:

# 9 Estruturas comparativas

Estruturas comparativas são de anotação complexa, o que se verifica pela existência de um **working group (WG) em UD** dedicado especialmente a elas. A seguir, listamos as frases utilizadas no WG, traduzidas em português, e com a anotação adequada, além de algumas frases de anotação complexa no Bosque-UD.

## 9.1 Frases do Working Group

### 9.1.1 Estruturas como "tanto...quanto"

TANTO será DET/amod QUANTO será SCONJ/mark

8	é	ser	AUX	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	9	cop	-	-
9	essencial	essencial	ADJ	-	Gender=Masc Number=Sing	1	ccomp	-	-
10	para	para	ADP	-	11	case	-	-	-
11	funções	função	NOUN	-	Gender=Fem Number=Plur	9	obl	-	-
12	mentais	mental	ADJ	-	Gender=Fem Number=Plur	11	amod	-	-
13	de	de	ADP	-	15	case	-	-	-
14	altíssimo	alto	ADJ	-	Gender=Masc Number=Sing	15	amod	-	-
15	nível	nível	NOUN	-	Gender=Masc Number=Sing	11	nmod	-	SpaceAfter=No
16	,	,	PUNCT	-	20	punct	-	-	-
17	tais	tal	DET	-	Gender=Fem Number=Plur PronType=Dem	20	case	-	MNEPOS=ADP
18	como	como	ADP	-	17	fixed	-	-	-
19	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	20	det	-	-
20	raciocínio	raciocínio	NOUN	-	Gender=Masc Number=Sing	11	appos:parataxis	-	SpaceAfter=No
21	,	,	PUNCT	-	23	punct	-	-	-
22	a	o	DET	-	Definite=Def Gender=Fem Number=Sing PronType=Art	23	det	-	-
23	memória	memória	NOUN	-	Gender=Fem Number=Sing	20	conj	-	SpaceAfter=No
24	,	,	PUNCT	-	26	punct	-	-	-
25	o	o	DET	-	Definite=Def Gender=Masc Number=Sing PronType=Art	26	det	-	-
26	reconhecimento	reconhecimento	NOUN	-	Gender=Masc Number=Sing	20	conj	-	-
27-28	dos	de	ADP	-	29	case	-	-	-
27	de	de	ADP	-	29	case	-	-	-
28	os	o	DET	-	Definite=Def Gender=Masc Number=Plur PronType=Art	29	det	-	-
29	rostos	rosto	NOUN	-	Gender=Masc Number=Plur	26	nmod	-	SpaceAfter=No
30	,	,	PUNCT	-	33	punct	-	-	-
31-32	das	de	ADP	-	33	case	-	-	-
31	de	de	ADP	-	33	case	-	-	-
32	as	o	DET	-	Definite=Def Gender=Fem Number=Plur PronType=Art	33	det	-	-
33	melodias	melodia	NOUN	-	Gender=Fem Number=Plur	29	conj	-	-
34	musicais	musical	ADJ	-	Gender=Fem Number=Plur	33	amod	-	SpaceAfter=No
35	,	,	PUNCT	-	29	punct	-	-	-
36	etc.	etc.	ADV	-	20	conj	-	-	SpaceAfter=No

Figura 63: Eu coloquei *tanta farinha quanto* a receita pedia

9.1.2 Estruturas como "mais(...) do que" ou "menos(...) do que"

Casos distintos:

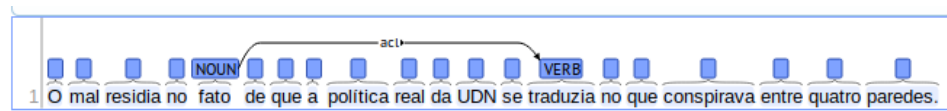


Figura 64: Martin é o cara *mais inteligente de todos*

9.2 Frases do Bosque-UD

# Apêndice

Todas as listas são retiradas do corpus Bosque-UD em sua versão 2.5.

Tabela 4: Lista das 23 palavras que ocorrem 3541 vezes como verbo auxiliar (*AUX*) no Bosque-UD

Verbo auxiliar	Frequência
acabar	56
andar	4
chegar	19
começar	62
continuar	57
costumar	10
deixar	24
dever	217
estar	282
faltar	1
ficar	11
haver	37
ir	358
parar	3
parecer	17
passar	39
poder	394
procurar	1
quer	1
ser	1304
ter	536
vir	68
voltar	40

Verbo auxiliar	Frequência
ser	1304
ter	536
poder	394
ir	358
estar	282
dever	217
vir	68
começar	62
continuar	57
acabar	56
voltar	40
passar	39
haver	37
deixar	24
chegar	19
parecer	17
ficar	11
costumar	10
andar	4
parar	3
faltar	1
procurar	1
quer	1

<b>Primeiro verbo (ordem alfabética)</b>	<b>#</b>	<b>Primeiro verbo (ordem de frequência)</b>	<b>#</b>
abster	2	querer	108
acabar	3	conseguir	83
aceder	1	tentar	64
aceitar	6	fazer	45
achar	5	pretender	39
aconselhar	5	permitir	37
acreditar	2	decidir	33
acusar	30	acusar	30
admitir	12	deixar	27
afirmar	13	procurar	21
aguentar	1	levar	20
ajudar	9	ver	19
alegar	3	dar	18
ambicionar	1	gostar	17
ameaçar	14	obrigar	17
andar	2	ser	17
anunciar	2	saber	16
aperceber	1	ameaçar	14
apetecer	1	precisar	14
aplicar	1	preferir	14
apostar	1	resolver	14
aprender	5	afirmar	13
apresentar	2	admitir	12
apressar	1	considerar	12
aprestar	1	encontrar	12
aproveitar	1	ter	12
arriscar	2	dizer	11
atender	1	destinar	10
atrever	1	limitar	10
autorizar	3	ajudar	9
avisar	1	ficar	9
bastar	3	preparar	9
caber	1	esperar	8

*Documentação UD em português  
(e para língua portuguesa)*

cansar	1	pensar	8
cessar	1	comprometer	7
chamar	4	consistir	7
citar	1	continuar	7
começar	1	convidar	7
comprometer	7	recusar	7
concordar	3	aceitar	6
condenar	1	convencer	6
conduzir	1	estar	6
confidenciar	1	manter	6
confirmar	1	prometer	6
conseguir	83	achar	5
considerar	12	aconselhar	5
consistir	7	aprender	5
consultar	1	forçar	5
contar	1	mandar	5
continuar	7	parecer	5
contribuir	3	passar	5
convencer	6	propor	5
convencionar	1	tender	5
convidar	7	visar	5
credenciar	1	chamar	4
criticar	1	desejar	4
culpar	1	encarregar	4
dar	18	evitar	4
decidir	33	impedir	4
declarar	3	interessar	4
declinar	1	proibir	4
dedicar	2	reconhecer	4
deixar	27	sentir	4
depender	1	tencionar	4
desafiar	2	tratar	4
desejar	4	acabar	3
desistir	1	alegar	3
destinar	10	autorizar	3
dever	3	bastar	3

dispor	3	concordar	3
dizer	11	contribuir	3
duvidar	1	declarar	3
empenhar	2	dever	3
encarregar	4	dispor	3
encontrar	12	insistir	3
ensinar	1	ir	3
equivaler	1	optar	3
escolher	1	ousar	3
escusar	2	pôr	3
esperar	8	vir	3
estar	6	abster	2
estimar	1	acreditar	2
estimular	1	andar	2
estudar	1	anunciar	2
evitar	4	apresentar	2
exigir	1	arriscar	2
exortar	1	dedicar	2
experimentar	1	desafiar	2
falar	2	empenhar	2
fartar	1	escusar	2
fazer	45	falar	2
ficar	9	garantir	2
fingir	1	haver	2
foi	1	imaginar	2
forçar	5	impor	2
garantir	2	importar	2
gostar	17	incentivar	2
habituar	1	mostrar	2
haver	2	necessitar	2
imaginar	2	negar	2
impedir	4	ouvir	2
impor	2	parar	2
importar	2	queixar	2
incentivar	2	receber	2
incluir	1	referir	2



*Documentação UD em português  
(e para língua portuguesa)*

indagar	1	resistir	2
influenciar	1	sentar	2
informar	1	tornar	2
insinuar	1	aceder	1
insistir	3	aguentar	1
instar	1	ambicionar	1
interessar	4	aperceber	1
ir	3	apetecer	1
lembrar	1	aplicar	1
levar	20	apostar	1
limitar	10	apressar	1
mandar	5	aprestar	1
manter	6	aproveitar	1
merecer	1	atender	1
mostrar	2	atrever	1
necessitar	2	avisar	1
negar	2	caber	1
obrigar	17	cansar	1
optar	3	cessar	1
orgulhar	1	citar	1
ousar	3	começar	1
ouvir	2	condenar	1
parar	2	conduzir	1
parecer	5	confidenciar	1
passado	1	confirmar	1
passar	5	consultar	1
pedir	1	contar	1
pensar	8	convencionar	1
perder	1	credenciar	1
permanecer	1	criticar	1
permitir	37	culpar	1
persuadir	1	declinar	1
planear	1	depende	1
poder	1	desistir	1
precisar	14	duvidar	1
preferir	14	ensinar	1

preparar	9	equivaler	1
pretender	39	escolher	1
prevenir	1	estimar	1
prever	1	estimular	1
procurar	21	estudar	1
proibir	4	exigir	1
projectar	1	exortar	1
prometer	6	experimentar	1
propiciar	1	fartar	1
propor	5	fingir	1
provocar	1	foi	1
pôr	3	habituar	1
queixar	2	incluir	1
querer	108	indagar	1
receber	2	influenciar	1
reclamar	1	informar	1
recomeçar	1	insinuar	1
recompensar	1	instar	1
reconhecer	4	lembrar	1
recordar	1	merecer	1
recusar	7	orgulhar	1
reduzir	1	passado	1
referir	2	pedir	1
resistir	2	perder	1
resolver	14	permanecer	1
respeitar	1	persuadir	1
restar	1	planear	1
retirar	1	poder	1
saber	16	prevenir	1
salientar	1	prever	1
sentar	2	projectar	1
sentir	4	propiciar	1
ser	17	provocar	1
soar	1	reclamar	1
sonhar	1	recomeçar	1
sublinhar	1	recompensar	1

*Documentação UD em português  
(e para língua portuguesa)*

sujeitar	1	recordar	1
suportar	1	reduzir	1
surgir	1	respeitar	1
suspeitar	1	restar	1
tardar	1	retirar	1
teimar	1	salientar	1
temer	1	soar	1
tencionar	4	sonhar	1
tender	5	sublinhar	1
tentar	64	sujeitar	1
ter	12	suportar	1
tornar	2	surgir	1
tratar	4	suspeitar	1
trazer	1	tardar	1
ver	19	teimar	1
vir	3	temer	1
virar	1	trazer	1
visar	5	virar	1
voltar	1	voltar	1

Tabela 9: Lista dos 196 verbos plenos que ocorrem 1150 vezes como *pais* de uma colocação verbal no Bosque-UD

Tabela 5: Lista das 2 palavras que ocorrem 138 vezes como verbo de ligação (*AUX/cop*) no Bosque-UD

Verbo de ligação	Frequência
estar	33
ser	105

Tabela 6: Lista das 2 palavras que ocorrem 1115 vezes como voz passiva (*AUX/aux:pass*) no Bosque-UD

Voz passiva	Frequência
ficar	1
ser	1114

Tabela 7: Lista das 23 palavras que ocorrem 2256 vezes como locução verbal (*AUX/aux*) no Bosque-UD

Locução verbal	Frequência
acabar	56
andar	4
chegar	19
começar	62
continuar	57
costumar	10
deixar	24
dever	217
estar	249
faltar	1
ficar	10
haver	37
ir	358
parar	3
parecer	17
passar	39
poder	394
procurar	1
quer	1
ser	53
ter	536
vir	68
voltar	40

Locução verbal	Frequência
ter	536
poder	394
ir	358
estar	249
dever	217
vir	68
começar	62
continuar	57
acabar	56
ser	53
voltar	40
passar	39
haver	37
deixar	24
chegar	19
parecer	17
costumar	10
ficar	10
andar	4
parar	3
faltar	1
procurar	1
quer	1

Tabela 8: Lista das 31 MWEs que ocorrem 570 vezes como locuções auxiliares (*AUX/aux + ADP/compound*) no Bosque-UD

MWE auxiliar	Frequência
acabar de	11
acabar por	30
andar a	3
chegar a	22
começar a	58
começar por	6
continuar a	57
continuar por	1
deixar de	30
dever a	1
estar a	124
estar para	1
estar por	1
ficar a	4
ficar de	1
haver a	1
haver de	2
haver que	1
ir a	3
ir de	1
parar de	3
passar a	43
poder a	3
ser de	3
tender a	1
ter a	9
ter de	62
ter que	3
tornar a	1
vir a	42
voltar a	42

MWE auxiliar	Frequência
estar a	124
ter de	62
começar a	58
continuar a	57
passar a	43
vir a	42
voltar a	42
acabar por	30
deixar de	30
chegar a	22
acabar de	11
ter a	9
começar por	6
ficar a	4
andar a	3
ir a	3
parar de	3
poder a	3
ser de	3
ter que	3
haver de	2
continuar por	1
dever a	1
estar para	1
estar por	1
ficar de	1
haver a	1
haver que	1
ir de	1
tender a	1
tornar a	1

# Abreviações

[Ir para tabela de conteúdos](#)





# Agradecimentos

[Ir para tabela de conteúdos](#)