

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 detalhada 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 Numerais 12
 - 2.1 Números compostos 13
 - 2.2 *R\$ 50*: quem é o pai? 14
 - 2.3 *50% das pessoas*: quem é o pai? 14
 - 2.4 *Primeiro* lugar: adjetivo ou numeral? 14

4 Atributos morfológicos (feats)

17

5 Dependências (dephead e deprel)

21

- 1 Estruturas comparativas 21
 - 1.1 Frases do Working Group 21
 - 1.2 Frases do Bosque-UD 22
- 2 Elementos discursivos 22
 - 2.1 É que 22

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 `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 (X_s) 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

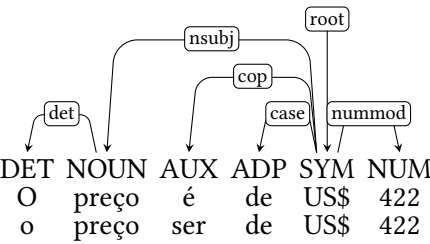


Figura 1: O preço é de US\$ 422

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

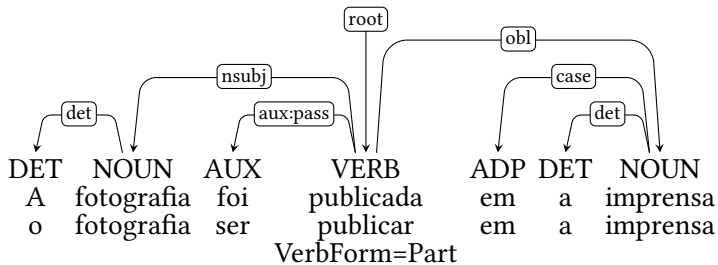


Figura 2: A fotografia *foi* publicada na imprensa

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

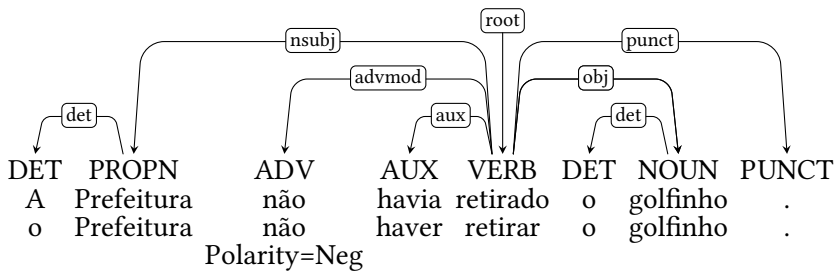


Figura 3: A Prefeitura não *havia retirado* o golfinho

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

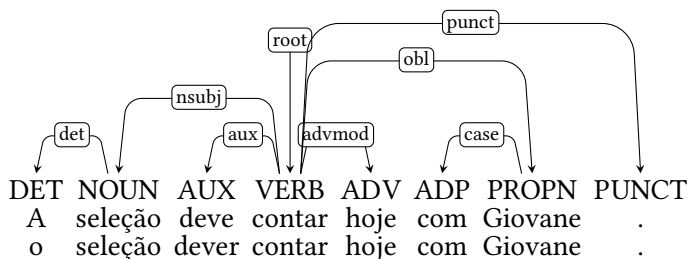


Figura 4: A seleção *deve contar* hoje com Giovane

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**.

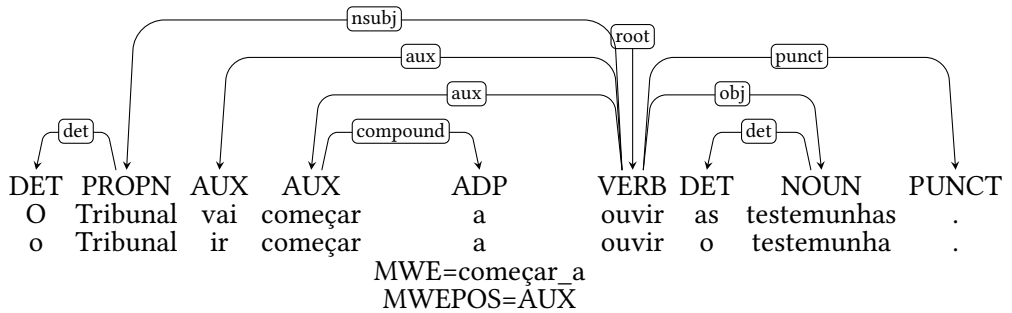


Figura 5: O Tribunal vai *começar a ouvir* as testemunhas

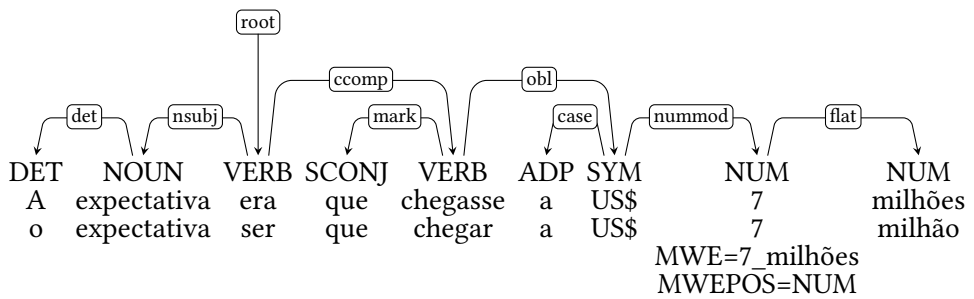


Figura 6: A expectativa *era* que chegasse a US\$7 milhões

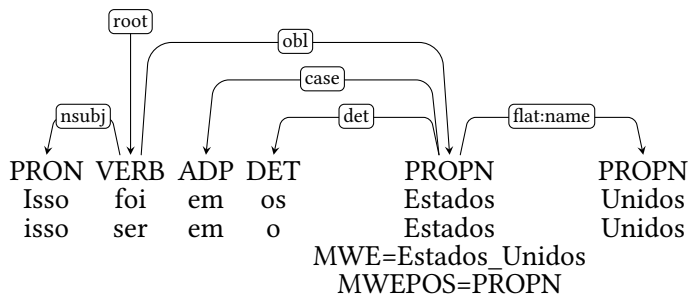


Figura 7: Isso *foi* nos Estados Unidos

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

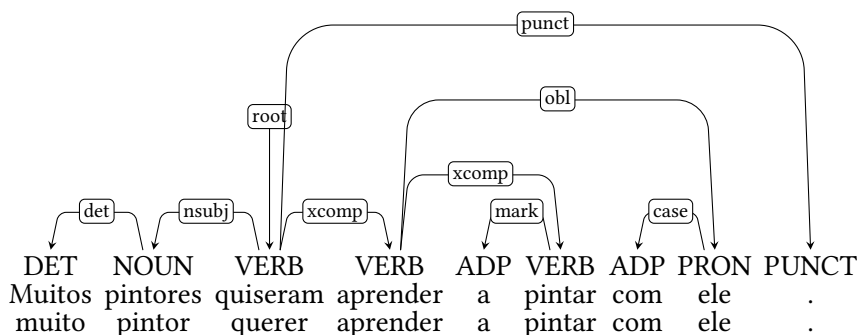


Figura 8: Muitos pintores *quiseram aprender a pintar com ele*

2 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*.

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 2.4: Primeiro lugar: adjetivo ou numeral?**).

Ver também: **Subseção 2.1: Números compostos**

2.1 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 9: *Trinta e sete*

Figura 10: *Montante de 1300 milhões de dólares*

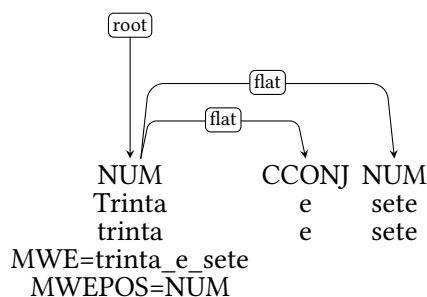


Figura 9: *Trinta e sete*

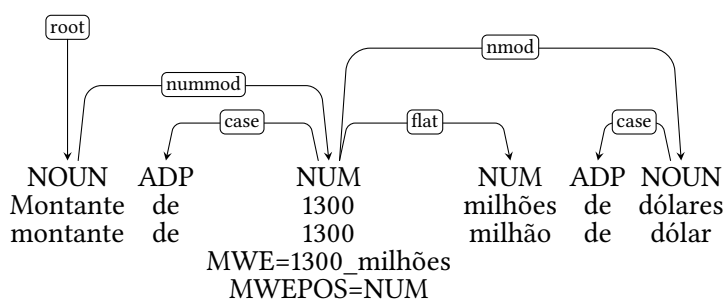


Figura 10: *Montante de 1300 milhões de dólares*

2.2 R\$ 50: quem é o pai?

Quando há a inserção de símbolos como **Figura 11: R\$ 50**, **Figura 12: U\$ 50** ou **Figura 13: 50%**, o símbolo deve ser o head. No caso de números coletivos (**Figura 14: 2 centenas**), o coletivo será o head.

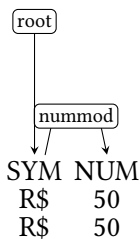


Figura 11: R\$ 50

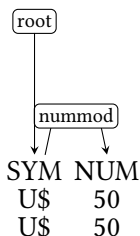


Figura 12: U\$ 50

2.3 50% das pessoas: quem é o pai?

Sempre que houver sintagma preposicionado, o sintagma anterior será o head, como em **Figura 15: 50% das pessoas** (onde o símbolo é o head, segundo **Subseção 2.2: R\$ 50: quem é o pai?**).

2.4 Primeiro lugar: adjetivo ou numeral?

Numerais ordinais escritos por extenso devem ser anotados como *ADJ*, e recebem a feature *NumType=Ord*, como na **Figura 16: Primeira tentativa**.

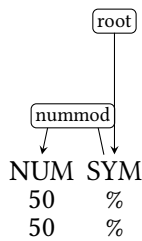


Figura 13: 50%

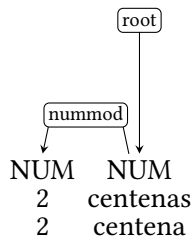


Figura 14: 2 centenas

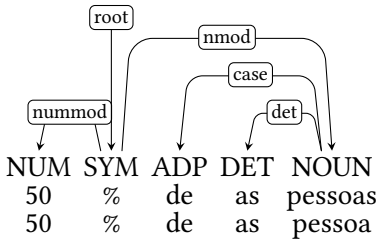


Figura 15: 50% das pessoas

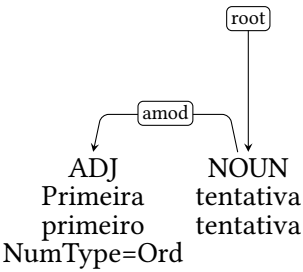


Figura 16: *Primeira tentativa*

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 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.

1.1 Frases do Working Group

8	é	ser	AUX	-	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	9	cop	-	-
9	essencial	para	ADP	-	ADJ	-	11	case	Gender=Masc Number=Sing 1 ccomp
10	para	para	ADP	-	-	-	-	-	-
11	funções	função	NOUN	-	-	-	-	-	-
12	mentais	mental	ADJ	-	-	-	-	-	-
13	de	de	ADP	-	-	-	-	-	-
14	altíssimo	alto	ADJ	-	-	-	-	-	-
15	nível	nível	NOUN	-	-	-	-	-	-
16	,	,	PUNCT	-	-	-	-	-	-
17	tais	tal	DET	-	-	-	-	-	-
18	como	como	ADP	-	-	-	-	-	-
19	o	o	DET	-	-	-	-	-	-
20	raciocínio	raciocínio	NOUN	-	-	-	-	-	-
21	,	,	PUNCT	-	-	-	-	-	-
22	a	o	DET	-	-	-	-	-	-
23	memória	memória	NOUN	-	-	-	-	-	-
24	,	,	PUNCT	-	-	-	-	-	-
25	o	o	DET	-	-	-	-	-	-
26	reconhecimento	reconhecimento	NOUN	-	-	-	-	-	-
27-28	dos	de	ADP	-	-	-	-	-	-
27	de	de	ADP	-	-	-	-	-	-
28	os	o	DET	-	-	-	-	-	-
29	rostos	rosto	NOUN	-	-	-	-	-	-
30	,	,	PUNCT	-	-	-	-	-	-
31-32	das	de	ADP	-	-	-	-	-	-
31	de	de	ADP	-	-	-	-	-	-
32	as	o	DET	-	-	-	-	-	-
33	melodias	melodia	NOUN	-	-	-	-	-	-
34	musicais	musical	ADJ	-	-	-	-	-	-
35	,	,	PUNCT	-	-	-	-	-	-
36	etc.	etc.	ADV	-	-	-	-	-	-

Figura 17: Eu coloquei *tanta farinha quanto* a receita pedia

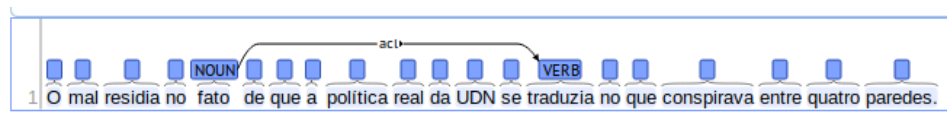


Figura 18: Martin é o cara *mais inteligente de todos*

1.2 Frases do Bosque-UD

2 Elementos discursivos

2.1 É que

A construção “é que” pode aparecer de duas maneiras no corpus: 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:

É: 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 19: Só depois *é que* levanto a cabeça para fazer um lançamento

Figura 20: A melhor equipe *é que* deve jogar

Figura 21: *É que* são uns antipáticos que nunca oferecem copos

OBS.: Consideramos “Não é que” como MWE.

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

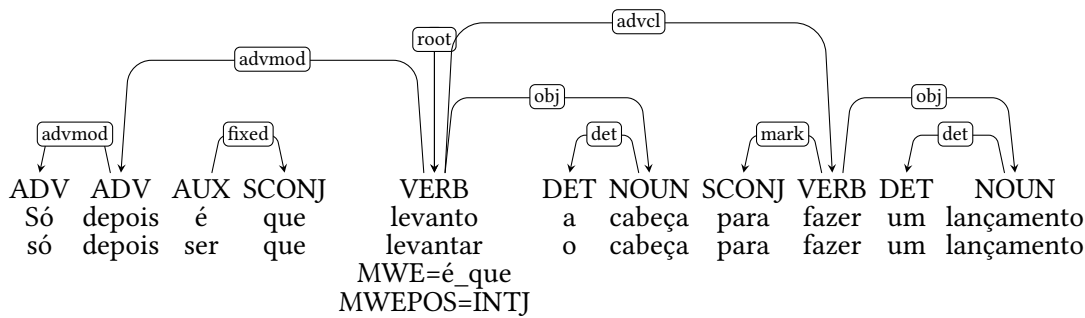


Figura 19: Só depois *é que* levanto a cabeça para fazer um lançamento

Documentação UD em português (e para língua portuguesa)

```
# text = A melhor equipe é que deve jogar.
1  A      o      DET      --      Definite=Daf|Gender=Fem|Number=Sing|FronType=Art      3      det      --      --
2  melhor bom    ADJ      --      Gender=Fem|Number=Sing      3      amod      --      --
3  equipe equipe NOUN      --      Gender=Fem|Number=Sing      4      nsubj      --      --
4  { é }      é      AUX      --      Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin      7      discourse      --      MWZ=é_que
5  { que }    que    SCONJ      --      --      4      fixed      --      --
6  deve      dever  AUX      --      Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin      7      aux      --      --
7  jogar     jogar  VERB      --      VerbForm=Inf      0      root      --      SpaceAfter=No
8  .          .      PUNCT      --      --      4      punct      --      SpaceAfter=No
```

Figura 20: A melhor equipe é que deve jogar

```
# text = É que são uns antipáticos que nunca oferecem copos.
1  { É }      é      AUX      --      5      discourse      --      MWZ=É_que
2  { que }    que    SCONJ      --      Gender=Masc|Number=Sing      1      fixed      --      --
3  são      ser      AUX      --      Mood=Ind|Number=Plur|Person=3|Tense=Pres|VerbForm=Fin      5      cop      --      --
4  uns      um      DET      --      Gender=Masc|Number=Plur|FronType=Art      5      det      --      --
5  antipáticos antipático ADJ      --      Gender=Masc|Number=Plur      0      root      --      --
6  que      que    SCONJ      --      8      mark      --      --
7  nunca    nunca ADV      --      8      advmod      --      --
8  oferecem oferecer VERB      --      Mood=Ind|Number=Plur|Person=3|Tense=Pres|VerbForm=Fin      5      acl:relcl
9  copos    copo   NOUN      --      Gender=Masc|Number=Plur      8      obj      --      SpaceAfter=No
10 .        .      PUNCT      --      5      punct      --      SpaceAfter=No
```

Figura 21: É que são uns antipáticos que nunca oferecem copos

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

Primeiro verbo (ordem alfabética)	#	Primeiro verbo (ordem de frequência)	#
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	depender	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](#)