

ET: uma Estação de Trabalho para busca, revisão e avaliação de corpora anotados morfofossintaticamente

Elvis de Souza
Orientação: Cláudia Freitas

Departamento de Letras
PUC-Rio



Construção de **datasets** para o **PLN** de língua portuguesa

- O processamento de linguagem natural por máquinas tem obtido muito sucesso recentemente por meio do **aprendizado de máquina**
- A abordagem por aprendizado de máquina ocorre por meio de exemplos que fornecemos à máquina, os **datasets**, ou corpora anotados, conjuntos de textos com informação linguística que queremos que as máquinas aprendam
- **A máquina precisa aprender com bons exemplos, linguisticamente consistentes** (Manning, 2011)

Como garantir consistência na anotação linguística?

O problema das locuções verbais

- começou a fazer
- voltou a falar
- acabou de processar

todas as ocorrências de locução verbal estão anotadas da mesma forma?

Vamos olhar no corpus...



dependente do primeiro ou do segundo verbo?

é uma preposição?

qual a função sintática?

Sentença anotada no formato CoNLLU

```
# text = Pela segunda vez desde quando começou a coordenar as ações no Rio, há duas semanas, o Exército mudou o nome das operações.
# sent_id = CF28-1
# source = CETENFolha n=28 cad=Brasil sec=pol sem=94a
# id = 109
1-2  Pela
1   Por    por    ADP    <sam->|PRP|@ADVL>      -         4         case
2   a      o      DET    <-sam>|<artd>|ART|F|S|@>N  Definite=Def|Gender=Fem|Number=Sing|PronType=Art  4         det
3   segunda segundo ADJ    <NUM-ord>|ADJ|F|S|@>N  Gender=Fem|Number=Sing|NumType=Ord  4         amod      ChangedBy=Issue
4   vez    vez    NOUN   <np-def>|N|F|S|@P<    Gender=Fem|Number=Sing  22        obl       ChangedBy=Issue137
5   desde  desde  ADP    PRP|@>A      6         case
6   quando quando  ADV    <rel>|ADV|@ADVL>      9         advmod
7   começou começar  AUX    <aux>|V|PS|3S|IND|@FS-N<  Mood=Ind|Number=Sing|Person=3|Tense=Past|VerbForm=Fin  9         aux
8   a      a      ADP    PRP|@PRT-AUX<  7         compound
9   coordenar  coordenar  VERB   <mv>|V|INF|@ICL-AUX<  VerbForm=Inf  4         acl:relcl
10  as     o      DET    <artd>|ART|F|P|@>N  Definite=Def|Gender=Fem|Number=Plur|PronType=Art  14        det
11  ações  ação   NOUN   <np-def>|N|F|P|@<ACC  Gender=Fem|Number=Plur  9         obj
12-13 no
12  em     em     ADP    <sam->|PRP|@<ACC      -         14        case
13  o      o      DET    <-sam>|<artd>|ART|M|S|@>N  Definite=Def|Gender=Masc|Number=Sing|PronType=Art  14        det
14  Rio    Rio    PROPN  PROP|M|S|@P<    Gender=Masc|Number=Sing  9         obl       ChangedBy=Issue165|SpaceAfter=No
15  ,      ,      PUNCT  PU|@PU      18        punct
16  há     há     ADP    PRP|@ADVL>      18        case
17  duas  dois  NUM    <num>|NUM|F|P|@>N  NumType=Card  18        nummod   ChangedBy=Issue168
18  semanas semana  NOUN   <np-def>|N|F|P|@P<    Gender=Fem|Number=Plur  22        obl       ChangedBy=Issue137|ChangedBy=Is
19  ,      ,      PUNCT  PU|@PU      18        punct
20  o      o      DET    <artd>|ART|M|S|@>N  Definite=Def|Gender=Masc|Number=Sing|PronType=Art  21        det
21  Exército  exército  NOUN   <prop>|<np-def>|N|M|S|@SUBJ>  Gender=Masc|Number=Sing  22        nsubj
22  mudou  mudar  VERB   <mv>|V|PS|3S|IND|@FS-STA  Mood=Ind|Number=Sing|Person=3|Tense=Past|VerbForm=Fin  0         root
23  o      o      DET    <artd>|ART|M|S|@>N  Definite=Def|Gender=Masc|Number=Sing|PronType=Art  24        det
24  nome   nome   NOUN   <np-def>|N|M|S|@<ACC  Gender=Masc|Number=Sing  22        obj
25-26 das
25  de     de     ADP    <sam->|PRP|@N<ARG      -         27        case
```

Precisamos de uma
ferramenta que nos ajude
a buscar, revisar e editar
as sentenças!

Interrogatório

interrogações recentes - nova interrogação

Pesquisa: (?)

Corpus:

bosque_organico.conllu ▼

Teste Salvar busca

Pesquisar

Critério 1: Regex

Se a expressão regular for encontrada em alguma sentença, retorná-la.

Dica: Cada grupo da expressão regular será mostrado com uma cor diferente.

Dica 2: Em um arquivo .conllu, "\t" separa as colunas, e "\n", as palavras.

Exemplos:

1) Retornar sentenças com o lema "dizer" seguido de um "PROPN".

- 1 (.*dizer.*)\n(.*PROPN)

2) Retornar sentenças com substantivos seguidos ou não de adjetivos, além de um nome próprio que seja apostro.

- 1 (.*NOUN.*)\n(.*ADJ.*)*\n(.*PROPN.*appos)

3) Retornar sentenças com tokens cujo POS seja "NUM" e o DEPREL, diferente de "NUMMOD".

- 1 .*t.*\t.*\tNUM\t(?!.*nummod)

2/75

CF62-2

«Um diplomata da Bélgica comprou apartamento para passar as férias de julho e janeiro, mas fora desse período ele aluga o imóvel por US\$ 1.000», **diz Bezerra.**

Esconder contexto

Mostrar anotação

Mostrar opções

Abrir inquérito

2/75

CF62-2

«Um diplomata da Bélgica comprou apartamento para passar as férias de julho e janeiro, mas fora desse período ele aluga o imóvel por US\$ 1.000», diz Bezerra.

Esconder contexto

Mostrar anotação

Mostrar opções

Abrir inquérito

Segundo ele, italianos e belgas já descobriram que comprar imóveis em Natal é um ótimo negócio. «Um diplomata da Bélgica comprou apartamento para passar as férias de julho e janeiro, mas fora desse período ele aluga o imóvel por US\$ 1.000», diz Bezerra.

2/75

CF62-2

«Um diplomata da Bélgica comprou apartamento para passar as férias de julho e janeiro, mas fora desse período ele aluga o imóvel por US\$ 1.000», diz Bezerra.

Esconder contexto

Esconder anotação

Mostrar opções

Abrir inquérito

Segundo ele, italianos e belgas já descobriram que comprar imóveis em Natal é um ótimo negócio. «Um diplomata da Bélgica comprou apartamento para passar as férias de julho e janeiro, mas fora desse período ele aluga o imóvel por US\$ 1.000», diz Bezerra.

```
# text = «Um diplomata da Bélgica comprou apartamento para passar as férias de julho e janeiro, mas fora desse período ele aluga o imóvel por US$ 1.000», diz Bezerra.
# sent_id = CF62-2
# source = CETENFolha n=62 cad=Imóveis sec=eco sem=94b
# id = 236
1 « « PUNCT PU|@PU 7 punct ChangedBy=Issue165|SpaceAfter=No
2 Um um DET <arti>|ART|M|S|@>N Definite=Ind|Gender=Masc|Number=Sing|PronType=Art 3 det _ _
3 diplomata diplomata NOUN <np-idf>|N|M|S|@SUBJ> Gender=Masc|Number=Sing 7 nsubj _ _
4-5 da
4 de de ADP <sam->|PRP|@N< 6 case _ _
5 a o DET <-sam>|<artd>|ART|F|S|@>N Definite=Def|Gender=Fem|Number=Sing|PronType=Art 6 det _ _
6 Bélgica Bélgica PROPN PROP|F|S|@P< Gender=Fem|Number=Sing 3 nmod
7 comprou comprar VERB <first-cjt>|<mv>|V|PS|3S|IND|@FS-ACC> Mood=Ind|Number=Sing|Person=3|Tense=Past|VerbForm=Fin 32 ccomp:parataxis _ _
8 apartamento apartamento NOUN <np-idf>|N|M|S|@<ACC Gender=Masc|Number=Sing 7 obj _ _
9 para para ADP PRP|@<ADVL 10 mark _ _
10 passar passar VERB <mv>|V|INF|@ICL-P< VerbForm=Inf 7 advcl
11 as o DET <artd>|ART|F|P|@>N Definite=Def|Gender=Fem|Number=Plur|PronType=Art 12 det _ _
12 férias férias NOUN <np-def>|N|F|P|@<ACC Gender=Fem|Number=Plur 10 obj _ _
13 de de ADP PRP|@N< 14 case _ _
14 julho julho NOUN <first-cjt>|<np-idf>|N|M|S|@P< Gender=Masc|Number=Sing 12 nmod _ _
15 e e CCONJ <co-prparg>|KC|@CO 16 cc d2d:#106
16 janeiro janeiro NOUN <cjt>|<np-idf>|N|M|S|@P< Gender=Masc|Number=Sing 14 conj _ ChangedBy=Issue165|SpaceAfter=No
17 , , PUNCT PU|@PU 24 punct _ d2d:#106
18 mas mas CCONJ <co-fcl>|<co-vfin>|KC|@CO 24 cc _ d2d:#106
19 fora fora ADV ADV|@ADVL> 24 advmod _ _
20-21 desse
```


Abrir inquérito

1	«	«	PUNCT	PU @PU	-	7	punct
2	Um	um	DET	<arti> ART M S @>N	Definite=Ind Gender=Masc Number=Sing PronType=Art	3	det
3	diplomata	diplomata	NOUN	<np-idf> N M S @SUBJ>	Gender=Masc Number=Sing	7	nsubj
4-5	da	-	-	-	-	-	-
4	de	de	ADP	<sam-> PRP @N<	-	6	case
5	a	o	DET	<-sam> <artd> ART F S @>N	Definite=Def Gender=Fem Number=Sing PronType=Art	6	det
6	Bélgica	Bélgica	PROPN	PROP F S @P<	Gender=Fem Number=Sing	3	nmod
7	comprou	comprar	VERB	<first-cjt> <mv> V PS 3S IND @FS-ACC>	Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin	32	ccomp:parataxis
8	apartamento	apartamento	NOUN	<np-idf> N M S @<ACC	Gender=Masc Number=Sing	7	obj
9	para	para	ADP	PRP @<ADVL	-	10	mark
10	passar	passar	VERB	<mv> V INF @ICL-P<	VerbForm=Inf	7	advcl
11	as	o	DET	<artd> ART F P @>N	Definite=Def Gender=Fem Number=Plur PronType=Art	12	det
12	férias	féria	NOUN	<np-def> N F P @<ACC	Gender=Fem Number=Plur	10	obj
13	de	de	ADP	PRP @N<	-	14	case
14	julho	julho	NOUN	<first-cjt> <np-idf> N M S @P<	Gender=Masc Number=Sing	12	nmod
15	e	e	CCONJ	<co-prparg> KC @CO	-	16	cc
16	janeiro	janeiro	NOUN	<cjt> <np-idf> N M S @P<	Gender=Masc Number=Sing	14	conj
17	,	,	PUNCT	PU @PU	-	24	punct
18	mas	mas	CCONJ	<co-fcl> <co-vfin> KC @CO	-	24	cc
19	fora	fora	ADV	ADV @ADVL>	-	24	advmod
20-21	desse	-	-	-	-	-	-
20	de	de	ADP	<sam-> PRP @A<	-	22	case
21	esse	esse	DET	<-sam> <dem> DET M S @>N	Gender=Masc Number=Sing PronType=Dem	22	det

Uma vez que as
alterações tenham
sido feitas, é hora
de **avaliar** o corpus

Avaliação intrínseca

- Julgamento da **qualidade do *modelo*** treinado a partir do meu corpus.
Quantos % de acerto o modelo teve?

Metric	Precision	Recall	F1 Score	AligndAcc
Tokens	100.00	100.00	100.00	
Sentences	100.00	100.00	100.00	
Words	100.00	100.00	100.00	
UPOS	96.97	96.97	96.97	96.97
XPOS	100.00	100.00	100.00	100.00
UFeats	93.18	93.18	93.18	93.18
AllTags	92.03	92.03	92.03	92.03
Lemmas	97.14	97.14	97.14	97.14
UAS	85.25	85.25	85.25	85.25
LAS	81.15	81.15	81.15	81.15
CLAS	73.96	73.51	73.74	73.51
MLAS	65.69	65.29	65.49	65.29
BLEX	70.61	70.19	70.40	70.19

Analisando os números: o aprendizado melhorou?

Apenas um “sim ou não” não basta: queremos aproximar o olhar para cada categoria morfossintática e entender quais classes gramaticais ou relações sintáticas tiveram melhora ou perda de acertos no aprendizado

Avaliação por categoria morfossintática

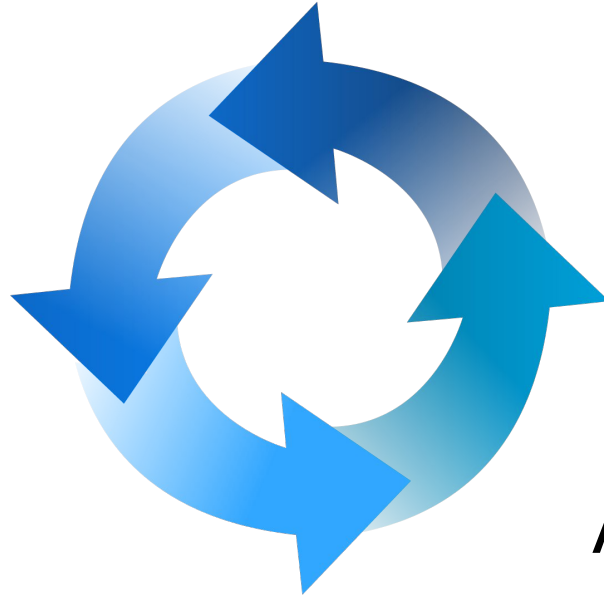
Acurácia por categoria gramatical			
UPOS	GOLDEN	ACERTOS	PORCENTAGEM
ADJ	444	386	86.93693693693693%
ADP	1623	1616	99.5686999383857%
ADV	364	350	96.15384615384616%
AUX	283	272	96.113074204947%
CCONJ	210	203	96.66666666666667%
DET	1561	1543	98.8468930172966%
NOUN	1925	1837	95.42857142857143%
NUM	248	230	92.74193548387096%
PART	17	14	82.35294117647058%
PRON	318	295	92.76729559748428%
PROPN	851	812	95.41715628672151%

Avaliação por
matriz de
confusão,
recurso muito
utilizado em
ciência dos
dados

UD[2]	ADJ	ADP	ADV	AUX	CCONJ	DET	NOUN	NUM	PART	PRON	PROPN	PUNCT	SCONJ	SYM	VERB	X	_	All
UD[1]																		
ADJ	<u>386</u>	<u>1</u>	<u>2</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>27</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>5</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>23</u>	<u>0</u>	<u>0</u>	444
ADP	<u>0</u>	<u>1616</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>4</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>2</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	1623
ADV	<u>2</u>	<u>3</u>	<u>350</u>	<u>0</u>	<u>0</u>	<u>2</u>	<u>3</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	364
AUX	<u>0</u>	<u>0</u>	<u>0</u>	<u>272</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>10</u>	<u>0</u>	<u>0</u>	283
CCONJ	<u>0</u>	<u>1</u>	<u>3</u>	<u>0</u>	<u>203</u>	<u>0</u>	<u>2</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	210
DET	<u>1</u>	<u>1</u>	<u>3</u>	<u>0</u>	<u>0</u>	<u>1543</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>6</u>	<u>6</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	1561
NOUN	<u>29</u>	<u>2</u>	<u>3</u>	<u>2</u>	<u>0</u>	<u>3</u>	<u>1837</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>41</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>8</u>	<u>0</u>	<u>0</u>	1925
NUM	<u>2</u>	<u>2</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>3</u>	<u>230</u>	<u>0</u>	<u>0</u>	<u>10</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	248
PART	<u>0</u>	<u>3</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>14</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	17
PRON	<u>0</u>	<u>0</u>	<u>6</u>	<u>0</u>	<u>0</u>	<u>7</u>	<u>2</u>	<u>1</u>	<u>0</u>	<u>295</u>	<u>2</u>	<u>0</u>	<u>5</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	318
PROPN	<u>4</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>29</u>	<u>2</u>	<u>0</u>	<u>1</u>	<u>812</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>2</u>	<u>0</u>	<u>0</u>	851
PUNCT	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>1343</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	1343
SCONJ	<u>0</u>	<u>1</u>	<u>3</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>12</u>	<u>0</u>	<u>0</u>	<u>84</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	100
SYM	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>29</u>	<u>0</u>	<u>0</u>	<u>0</u>	30
VERB	<u>13</u>	<u>1</u>	<u>1</u>	<u>7</u>	<u>0</u>	<u>0</u>	<u>5</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>5</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>821</u>	<u>0</u>	<u>0</u>	853
X	<u>1</u>	<u>2</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>7</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>17</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>3</u>	<u>0</u>	30
_	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>742</u>	742
All	438	1633	373	281	203	1560	1916	234	14	315	902	1343	91	29	865	3	742	10942

Busca

Revisão



Avaliação

ET - Estação de Trabalho

De código aberto, a Estação de Trabalho funciona com corpora no formato de anotação CoNLLU

Acesso:

<http://comcorhd.lettras.puc-rio.br/interrogatorio>

Código-fonte:

<https://github.com/alvelvis/Interrogat-rio>

Lançamento da versão 2.4 do Bosque-UD em maio/2019 (Nivre et al., 2019)

Revision type	Nature	Amount	Relation
	conversion	615	appos
		497	dep
		2068	obl
		1093	NUM/flat
Rule based		794	npmod/tmod removal
		1544	mark
	analysis	149	aux
	conversion + analysis	2516	MWE
		249	ccomp
Rule + manual	analysis	453	“o que” (DET/PRON)
		391	obl
Manual	analysis	164	xcomp
		70	POS corrections based on MC convergences
		395	POS corrections based on MC divergences
	conversion	151	“o que” (appos)
	UD validation script	1329	miscellaneous
Total		12478	

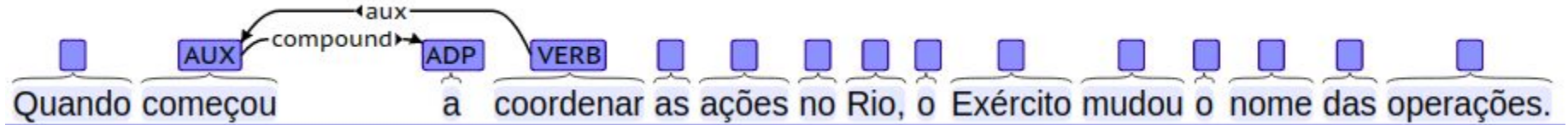
Table 1: Distribution of changes from 2.3 to 2.4.

≈ 5% de um corpus
que contém aprox.
200 mil palavras

Resultados 3/3

Um estudo sobre as locuções verbais

Nossa contribuição da Linguística Computacional para a descrição do português (contribuição teórica) e para o uso no PLN (contribuição empírica)



Referências 1/3

C. D. Manning, “Part-of-speech tagging from 97% to 100%: is it time for some linguistics?” in International conference on intelligent text processing and computational linguistics. Springer, 2011, pp. 171–189.[2]

A. Rademaker, F. Chalub, L. Real, C. Freitas, E. Bick, and V. de Paiva, “Universal dependencies for portuguese,” in Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017), 2017, pp. 197–206.2 Disponível em <https://github.com/alvelvis/Interrogat-rio/wiki>

R. McDonald, J. Nivre, Y. Quirnbach-Brundage, Y. Goldberg, D. Das, K. Ganchev, K. Hall, S. Petrov, H. Zhang, T. Ošcar et al., “Universal Dependency annotation for multilingual parsing,” in Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2013, pp. 92–97.

M. Straka, J. Hajic, and J. Straková, “Udpipe: trainable pipeline for processing conll-u files performing tokenization, morphological analysis, pos tagging and parsing,” in Proceedings of the tenth international conference on language resources and evaluation (LREC 2016), 2016, pp. 4290–4297.

D. Santos, “Podemos contar com as contas? In Sandra Maria Aluísio; Stella E O Tagnin (ed) New Language Technologies and Linguistic Research: A Two-Way Road Cambridge Scholars Publ 2014; 2014, 2014.[6]

F. M. Tyers, M. Sheyanova, and J. N. Washington, “Ud annotatrix: an annotation tool for universal dependencies,” 2017.

M. Janssen, “Dependency graphs and teitok: Exploiting dependency parsing,” in International Conference on Computational Processing of the Portuguese Language. Springer, 2018, pp. 470–478.

Referências 2/3

C. Freitas, E. de Souza, and L. Rocha, “Quantificando (e qualificando) osujeito oculto em português,” in VI Jornada de Descrição do Português, STIL 2019, 2019.

J. Nivre et al., “Universal dependencies 2.4,” 2019, LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (UFAL), Faculty of Mathematics and Physics, Charles University. [Online]. Available: <http://hdl.handle.net/11234/1-2988>

L. Rocha, I. Soares-Bastos, C. Freitas, and A. Rademaker, “Scavengerhunt: what do we find when look for confusions,” in International Conference on the Computational Processing of Portuguese, PROPOR2018, 2018.

D. Zeman, J. Hajič, M. Popel, M. Potthast, M. Straka, F. Ginter, J. Nivre, and S. Petrov, “Conll 2018 shared task: multilingual parsing from raw text to universal dependencies,” in Proceedings of the CoNLL2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, 2018, pp. 1–21.

C. Freitas, L. F. Trugo, F. Chalub, G. Paulino-Passos, and A. Rademaker, “Tagsets and datasets: Some experiments based on portuguese language,” in International Conference on Computational Processing of the Portuguese Language. Springer, 2018, pp. 459–469.

Afonso, S., Bick, E., Haber, R., and Santos, D. (2002). Floresta sintá(c)tica: a treebank for portuguese. In Manuel Gonzalez Rodrigues; Carmen Paz Suarez Araujo (ed) Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)(Las Palmas de Gran Canaria Espanha 29-31 de Maio de 2002) Paris: ELRA. ELRA.

Referências 3/3

Bagno, M. (2012). Gramática pedagógica do português brasileiro. Parábola Ed.

Camara Júnior, J. M. (1992). Dicionário de linguística e gramática: referente à língua portuguesa.

Nivre, J., De Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., et al. (2016). Universal dependencies v1: A multilingual treebank collection. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), pages 1659–1666.

Rademaker, A., Chalub, F., Real, L., Freitas, C., Bick, E., and de Paiva, V. (2017). Universal dependencies for portuguese. In Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017), pages 197–206.

Vilela, M. and Koch, I. (2001). Gramática da língua portuguesa: Gramática da palavra, gramática da frase, gramática do texto/discurso. Coimbra: Almedina.

Obrigado!

Elvis de Souza

elvis.desouza99@gmail.com

Cláudia Freitas

claudiafreitas@puc-rio.br

Departamento de Letras - PUC-Rio