

Documentação da anotação morfossintática do PetroGold

Elvis de Souza, Aline Silveira, Tatiana Cavalcanti,
Maria Clara Castro, Cláudia Freitas
[ComCorHd](#) - [ICA](#) - PUC-Rio

25/11/2021

Este documento relata decisões de anotação do corpus PetroGold, e contempla sobretudo decisões associadas a especificidades do gênero acadêmico e domínio do petróleo, além de decisões pouco especificadas durante a anotação do corpus Bosque, em sua versão UD. Para questões genéricas associadas à anotação UD em língua portuguesa, veja-se este documento: <https://comcorhd.lettras.puc-rio.br/Documenta-o-UD-PT>.

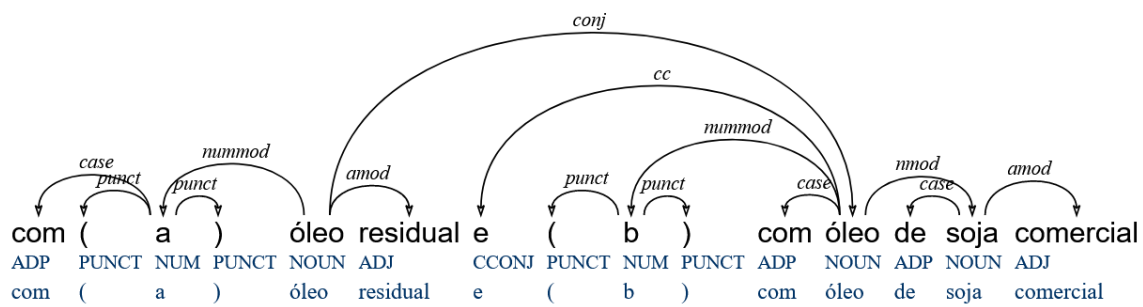
O corpus encontra-se disponível no endereço <https://petroles.puc-rio.ai>.

Algumas das informações aqui presentes são consideradas “workbench”, isto é, são anotações que utilizamos internamente mas não lançamos publicamente, por fugirem um pouco das diretrizes UD. Essas anotações estão indicadas, assim como as respectivas conversões para a anotação conforme a abordagem UD.

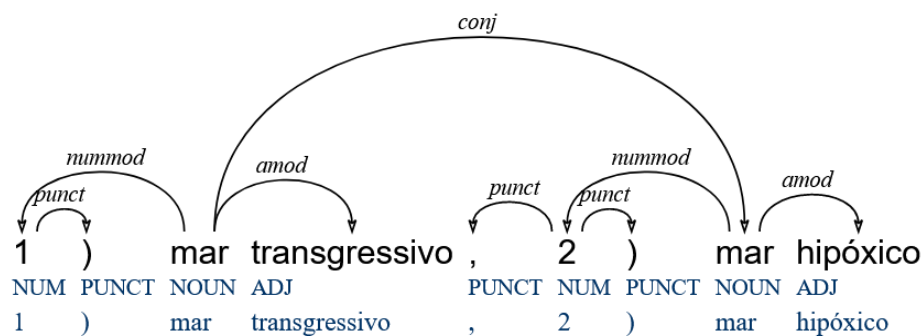
Itens enumerados e títulos de seção que começam com número

Em listas numeradas, independentemente de os itens serem algarismos - de qualquer tipo, como arábico ou romano - ou letras, o item enumerado será o head e o item enumerador receberá anotação NUM/nummod. Essa decisão se baseia nas discussões em UD disponíveis em <https://github.com/UniversalDependencies/docs/issues/156> e <https://universaldependencies.org/u/dep/list.html>. A pontuação que vier junto ao número(ou letra) será dependente dele pela relação *punct*. Já nos casos de lista itemizada com travessão, bola ou outro tipo de marcação que não seja alfanumérica, a decisão padrão é a de utilizar a anotação PUNCT/punct, segundo as guidelines do UD (<https://universaldependencies.org/u/dep/list.html>).

(271-20150810-MONOGRAFIA_0-405)

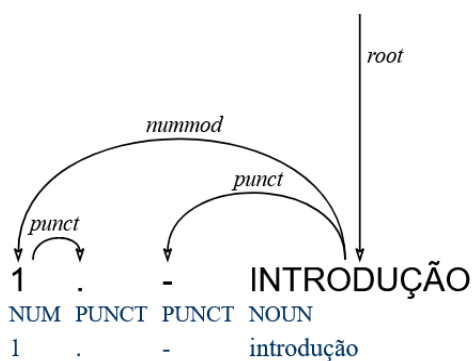


(284-20150917-MONOGRAFIA_0-396)



De modo semelhante anotamos os títulos de seção que começam com algarismos.

(247-20140910-MONOGRAFIA_0-1)

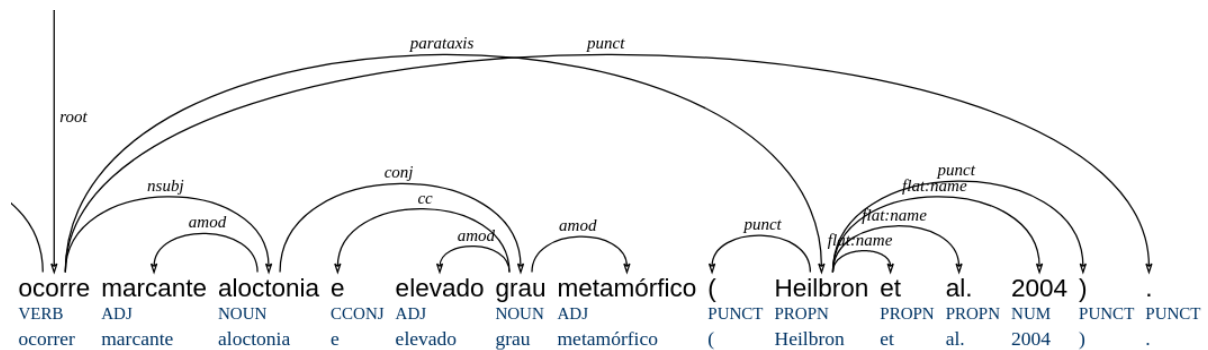


Referências bibliográficas

Quando estão entre parênteses, são *parataxis*, apontam para o root da sentença, independentemente da posição na frase e compõem um *flat:name*.

(247-20140910-MONOGRAFIA_0-77)

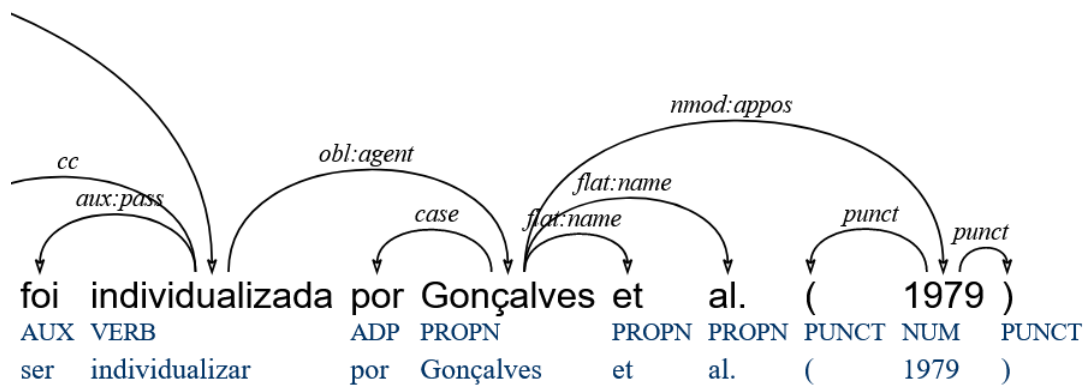
#text = Para o interior da faixa, *ocorre* marcante aloctonia e elevado grau metamórfico (Heilbron et al. 2004).



Quando estão fora de parênteses, o ano da referência é *nmod:appos* do sobrenome e o head do termo ocupa a função sintática de acordo com a posição na frase. A etiqueta *nmod:appos* será melhor explorada na seção “*nmod:appos*”.

(247-20140910-MONOGRAFIA_0-134)

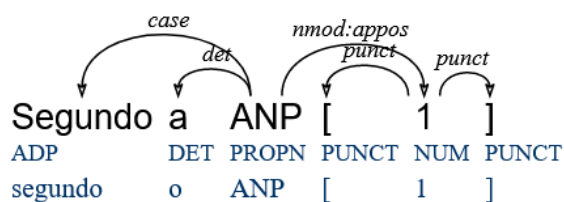
#text = A Plataforma de Florianópolis localiza-se na porção norte da bacia e foi individualizada por Gonçalves et al. (1979).



Em casos excepcionais (ou não convencionais) como este abaixo, o número é *nmod:appos* do nome.

#text = Segundo a **ANP [1]** (Agência Nacional de o Petróleo, Gás Natural e Biocombustíveis) , o biodiesel é um combustível composto de mono-ésteres alquílicos de ácidos graxos de cadeia longa derivados de óleos vegetais ou gorduras animais .

(279-20140530-MONOGRAFIA_0-7)



Tabelas e figuras

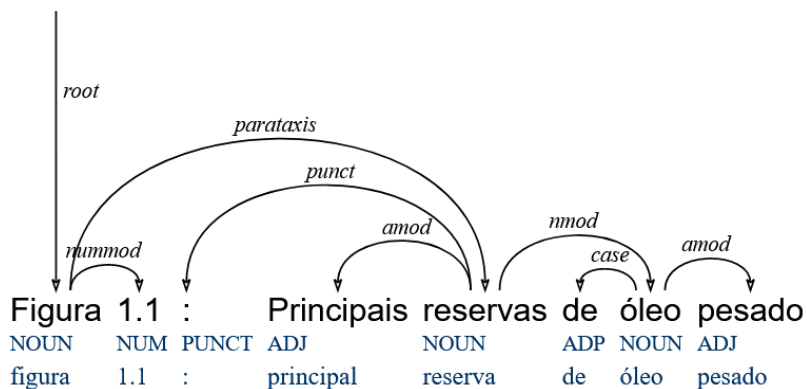
As figuras e tabelas se comportam de maneira similar às referências bibliográficas quando estão inseridas em uma frase.

Quando estão entre parênteses, são *parataxis* e apontam para o root da sentença, independentemente da posição na frase.

Quando estão fora de parênteses, o head do termo ocupa a função sintática de acordo com a posição na frase.

Nos casos em que formam uma frase completa, tem-se a seguinte anotação:

(77-20141209-TESEMSC_0-7)



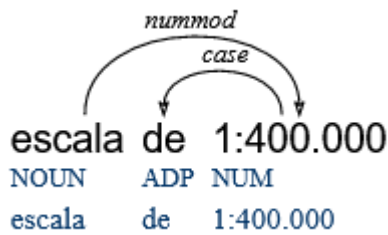
Fonte:

O que vier depois de “Fonte:” será apostro de “Fonte”.

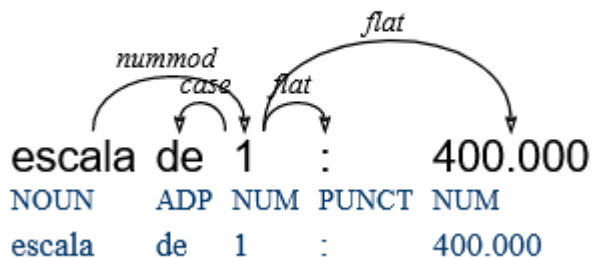
Escalas/proporções

Proporções como 1:400.000 são um único token, cuja POS é NUM:

text = A partir deste grid de dados, um mapa do campo magnético total anômalo foi gerado na escala de **1:400.000** (figura 1.2 e anexo 2).



Quando a escala estiver tokenizada **incorretamente** (figura abaixo), tudo deverá ser tratado como *flat*. O caso é análogo a números e datas, que são anotados como *flat*.

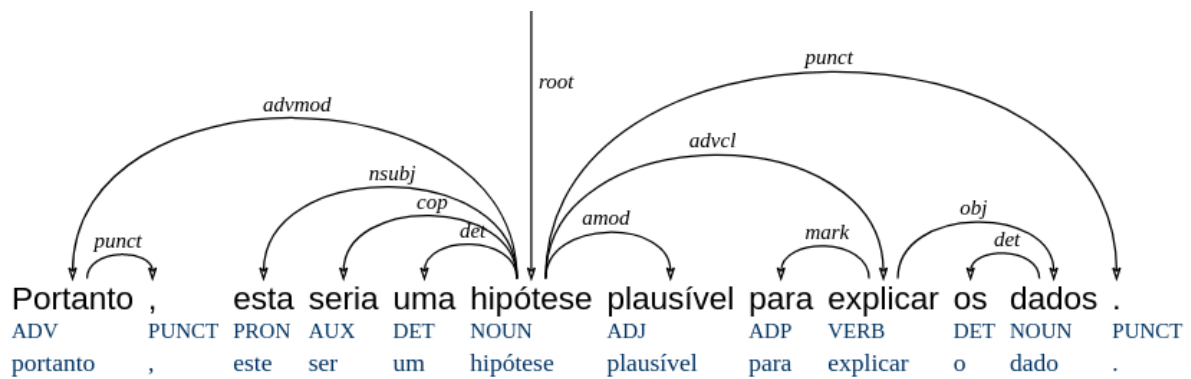


Locuções conjuntivas ou conjunções que iniciam frases

Elementos que iniciam orações com um único verbo são considerados ADV/advmod. Isso acontece porque, caso fossem anotadas como CCONJ/cc, seria necessário resgatar um elemento *conj*, que não está nesta frase, e sim na frase anterior. É o caso de “Portanto”, “Todavia”, “Sendo assim”, “Além disso”.

Preposições que funcionam como conjunções, conectando orações

Nos casos em que temos palavras tipicamente classificadas como preposições que fazem a conexão entre orações, anotamos a função sintática *mark*, mas a classe permanece sendo preposição.



MWEs (Expressões multi-palavras)

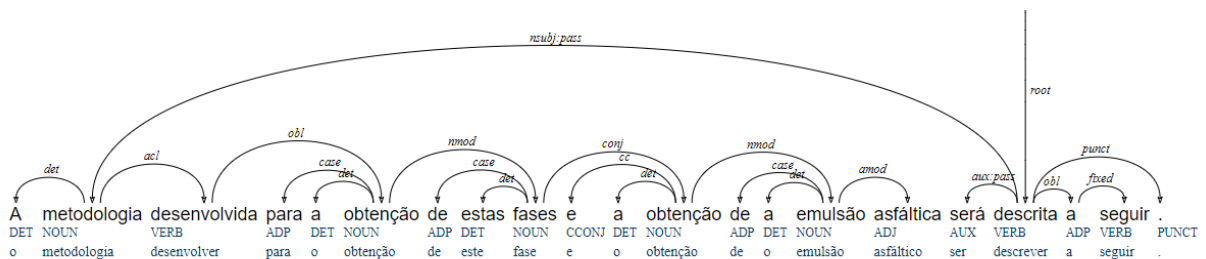
MWEs são expressões gramaticalmente cristalizadas na língua que se comportam como uma unidade semântico-funcional. Uma dada combinação pode ou não ser uma MWE, conforme o contexto. A seguir, detalhamos e exemplificamos a anotação dessas expressões, que podem ser de natureza gramatical (*fixed*) ou nominal (*compound* e *flat*).

MWEs gramaticais (*fixed*)

combinação “A SEGUIR” – três situações diferentes.

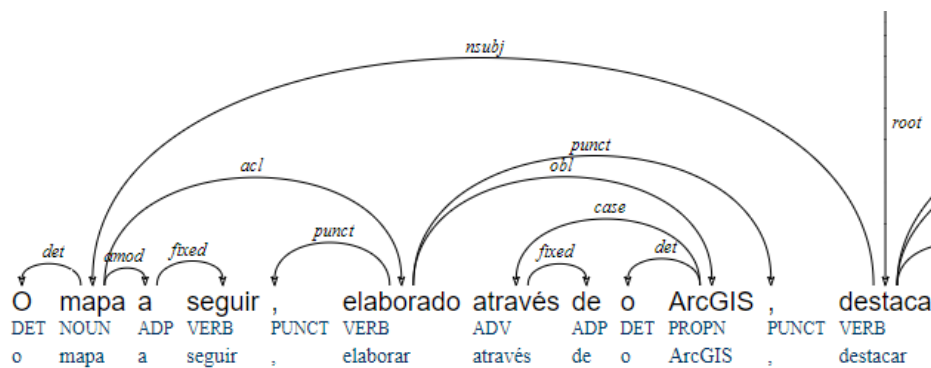
1) MWEPOS=ADV: quando aponta para um verbo.

(Ex.: 241-20140227-MONOGRAFIA_0-148)



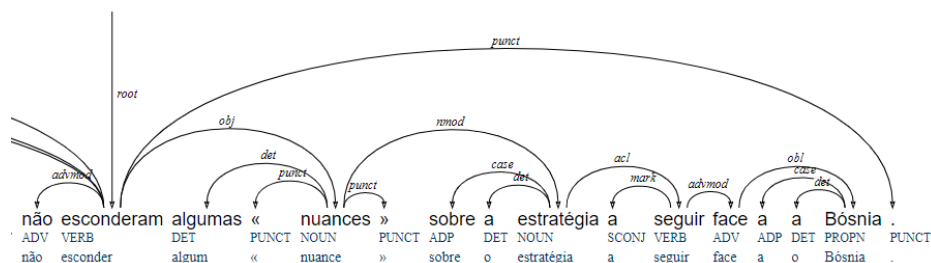
2) MWEPOS=ADJ: quando aponta para um nome. Semelhante a “seguinte”.

(Ex.: 161-20150810-MONOGRAFIA_0-504)



3) “A seguir” que não é MWE:

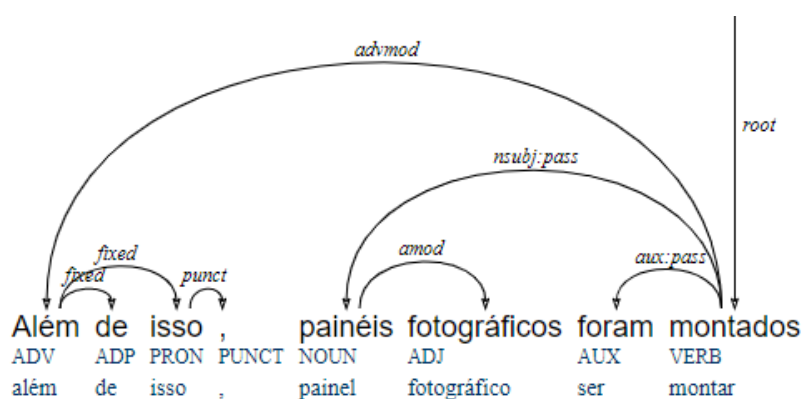
(Ex.: CP783-2)



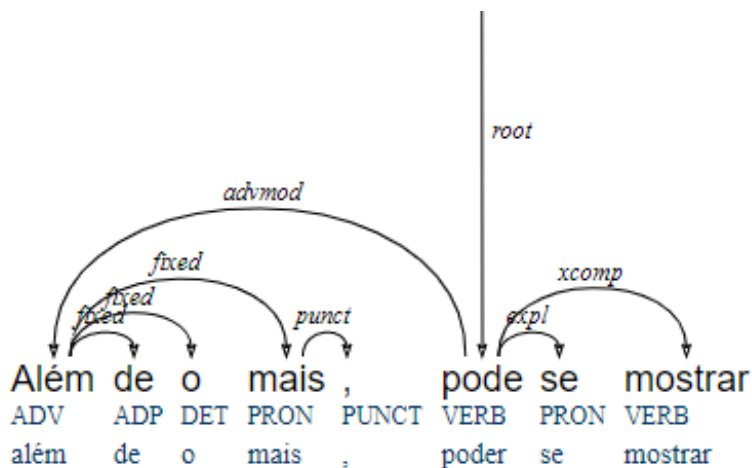
“ALÉM DISSO”, “ALÉM DO MAIS” E “ALÉM DO QUE”

- MWEPOS=ADV e apontam para o root da frase.

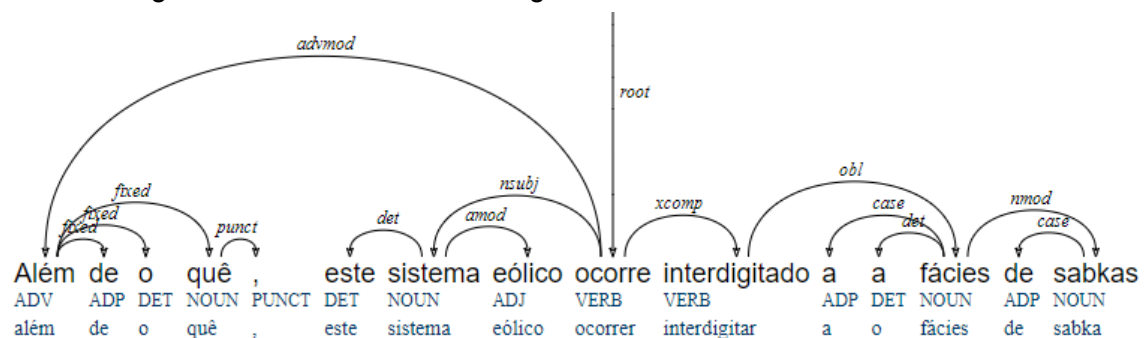
1) **Além disso**: [284-20150917-MONOGRAFIA_0-62] “**Além disso**, painéis fotográficos foram montados de modo a obter imagens mais abrangentes, ilustrando relações laterais e verticais.”



2) **Além do mais**: [20-20140904-TESEDSC_0-442] “**Além do mais**, pode-se mostrar que a estimativa das coordenadas horizontais da fonte magnética através da Deconvolução de Euler, é estável se a janela de dados for centralizada nas coordenadas do ponto de máxima amplitude do sinal analítico.”



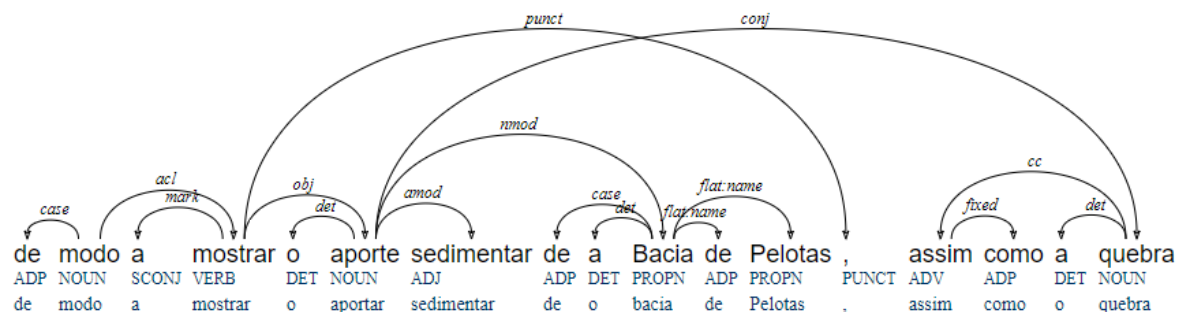
3) **Além do que:** [283-20150917-MONOGRAFIA_0-403] “**Além do quê**, este sistema eólico ocorre interdigtado à fácies de sabkas, sugerindo maior umidade em todo o sistema.”



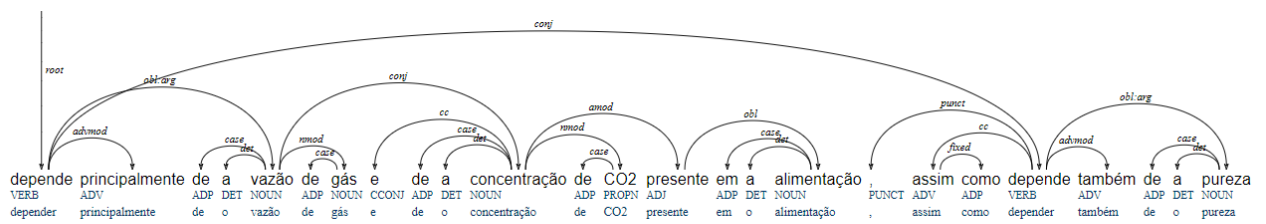
ASSIM COMO – Duas situações diferentes.

1) MWEPOS=CCONJ: locução conjuntiva coordenativa aditiva. Possível de ser substituída por “e”:

(Ex.: 247-20140910-MONOGRAFIA_0-240: As seções sísmicas de direção NW-SE são perpendiculares à linha de costa, de modo a mostrar o aporte sedimentar da Bacia de Pelotas, **assim como** a quebra da plataforma e o mergulho das camadas.)

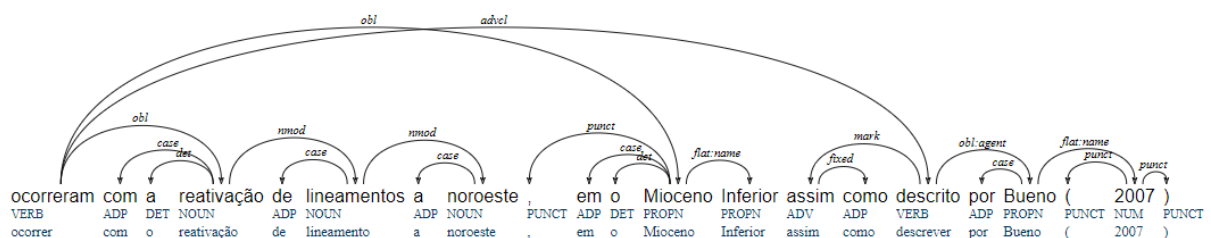


(Ex2.: 334-20150521-MONOGRAFIA_0-527: Conforme mostrado no diagrama da Figura 41, a escolha de qual processo usar para a remoção de CO2 depende principalmente da vazão de gás e da concentração de CO2 presente na alimentação, **assim como** depende também da pureza que se deseja obter.)



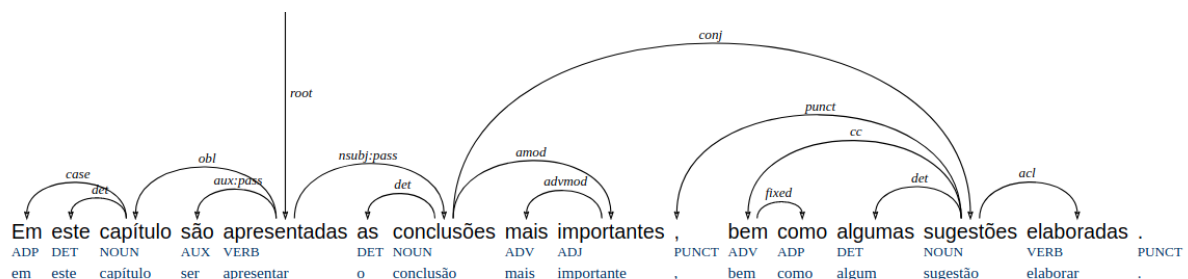
2) MWEPOS=SCONJ: locução conjuntiva subordinativa conformativa. Possível de substituir por “conforme”.

(Ex.: 247-20140910-MONOGRAFIA_0-299: Estas feições podem ser representar a Formação Curumim ou intrusões que ocorreram com a reativação de lineamentos a noroeste, no Mioceno Inferior **assim como** descrito por Bueno (2007).)



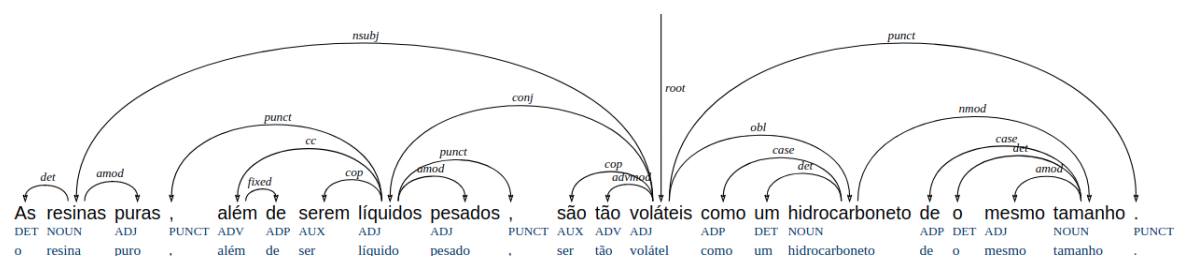
BEM COMO: MWEPOS=CCONJ

(Ex.: 119-20141209-TESEMSC_0-784: Neste capítulo são apresentadas as conclusões mais importantes derivadas deste trabalho, **bem como** algumas sugestões e recomendações elaboradas a partir do conhecimento adquirido.)



ALÉM DE: MWEPOS=CCONJ

(Ex.: 77-20141209-TESEMSC_0-53: As resinas puras, **além de** serem líquidos pesados ou sólidos pastosos, são tão voláteis como um hidrocarboneto do mesmo tamanho.)



MWEs nominais: compound e flat:name

Algumas combinações nominais, que poderiam ser analisadas como palavras compostas, estão anotadas como compound, e isso se aplica principalmente a termos técnicos em que uma das palavras é um nome próprio.

efeito estufa
meio ambiente
desvio padrão
raio X
óleo diesel

Wolburgiopsis plastica
Wolburgiopsis chinamuertensis
Dastilbe moraenses nov. esp.

Nomes próprios e flat:name

flat:name

- Nomes próprios que modificam outros nomes próprios (PROPN) são anotados como flat:name

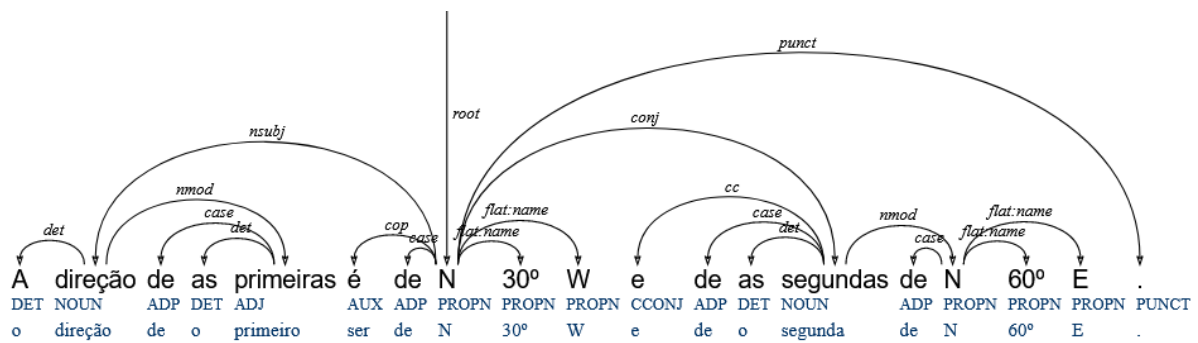
Bacia **Pelotas**(flat:name/Bacia)
Serra do **Mar**(flat:name/Serra)
Alto de **Cabo**(flat:name/Alto) **Frio**(flat:name/Alto)
São **Paulo**(flat:name/São)

Coordenadas geográficas e flat

Coordenadas geográficas são formadas por tokens distintos e recebem a etiqueta *flat:name*. Esta decisão é motivada principalmente por serem, os elementos das coordenadas geográficas, anotados como nomes próprios (PROPN).

text = A direção das primeiras é de **N 30° W** e das segundas de **N 60° E**.

(20-20140904-TESEDSC_0-297)

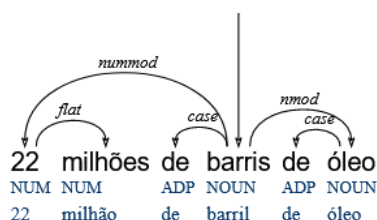
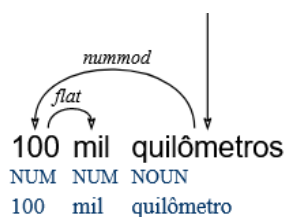


Numerais (e determinantes)

As palavras milhares|centenas|dezenas|milhões|bilhões|dúzias e mil são analisadas como numerais (upos = NUM e deprel = flat quando associadas a um outro numeral, como "5 mil"). Os números acompanhados por esse *flat* tem relação *nummod* com o restante do sintagma.

Consequentemente, damos um tratamento unificado às construções elencadas em (a) abaixo, independentemente da presença da preposição.

- (a) 2 milhões de barris
- 2 mil barris
- centenas de barris
- duas centenas de barris



De modo análogo, consideramos que, em português, as construções

- (b) Muitas das pessoas se arrependeram

Muitas pessoas se arrependeram

têm uma leitura (um sentido) muito próxima, sendo a presença da preposição uma questão mais estilística que gramatical. Em ambos os casos, o *head* da combinação [muitas das pessoas] ou [muitas pessoas] é *pessoas*, e *muitas* é DET/det.

Modificadores nominais: entre o adjunto adnominal e o aposto

nmod:appos

Esta é uma etiqueta criada por nós e que **não está presente na versão release, onde todos os casos de nmod:appos são transformados em nmod.**

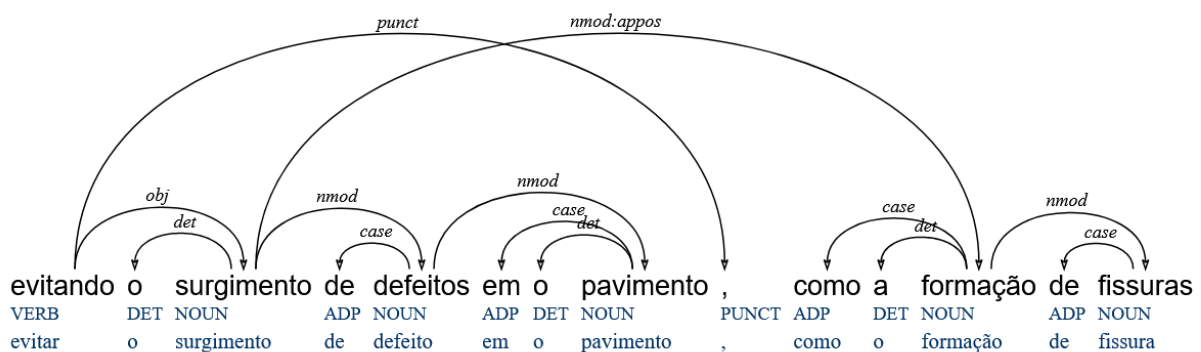
É a deprel utilizada no caso de estruturas em que não há equivalência (ou correferência) entre os dois termos (diferente do appos), mas há uma relação que, embora não explícita, é clara para o leitor (diferente da parataxis). **Ver também seção “Como/ tal como”:**

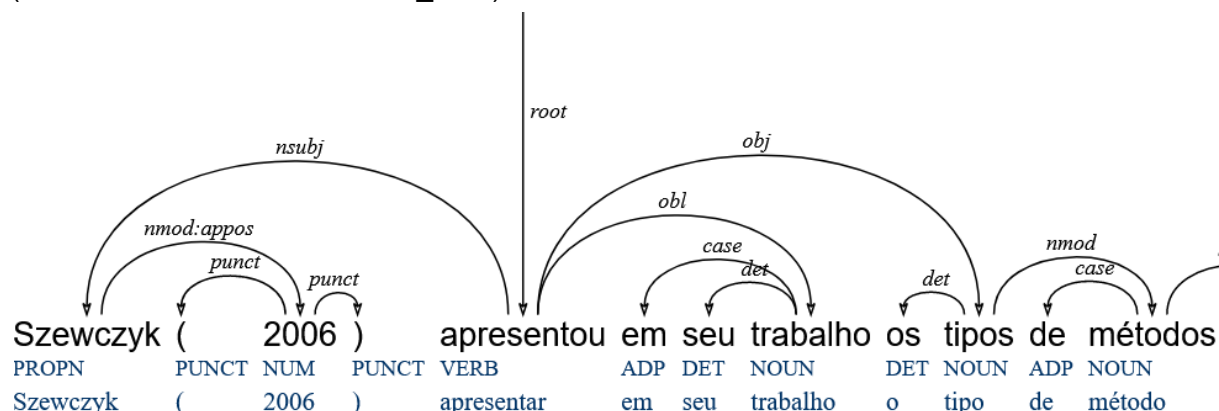
(0-20150121-TESEMSC_0_resumo-9)

Os resultados obtidos no estudo da recuperação mostraram que o método UOP-326 apresenta uma pior **recuperação** (78%) quando comparado aos **métodos** voltamétrico (96%) e **CG/EM** (95%)

- 78% → nmod:appos de "recuperação"
- 96% → nmod:appos - a porcentagem não corresponde a “métodos” diretamente, mas a "recuperação" que está implícita
- 95% → nmod:appos de “CG/EM”

(241-20140227-MONOGRAFIA_0-93)





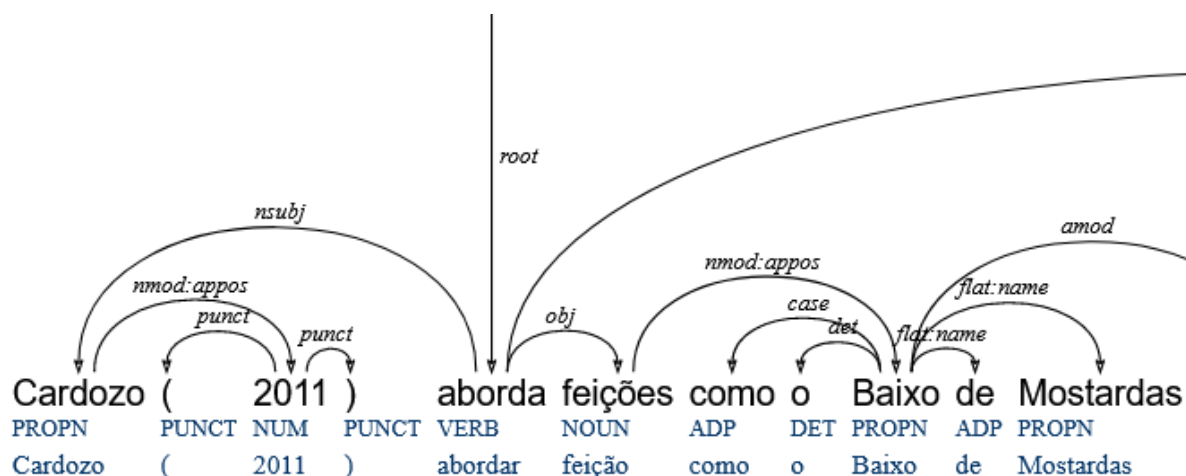
Também é anotado como *nmod:appos* (e, excepcionalmente neste caso, posteriormente convertido para *nummod*, e não *nmod*) o ano de referências bibliográficas, quando se encontra entre parênteses. Abaixo, "1979" é *nmod:appos* de "Gonçalves". (ver também a seção *Referências bibliográficas*)

A Plataforma de Florianópolis localiza-se na porção norte da bacia e foi individualizada por **Gonçalves et al. (1979)**

SN como SN; SN tais como SN

Os casos de "como.../ tal como..." são classificados como *nmod:appos* (e posteriormente convertidos em *nmod*)

#text = Cardozo (2011) aborda feições como o **Baixo** de Mostardas presente na porção sul da Bacia de Pelotas.



Quando não estiver claro a que elemento o núcleo da parataxis - que está entre parênteses - se refere, a relação será com o root.

text = No perfil magnético (figura 6.2), o corpo sedimentar é *representado* pelo contraste negativo de susceptibilidades (**regiões claras na imagem**).



The diagram illustrates the syntactic structure of the sentence "Pode-se notar que a SE ocorre o maior espessamento sedimentar (depocentro controlado por a falha de a borda sul)". The tree shows the following components and relationships:

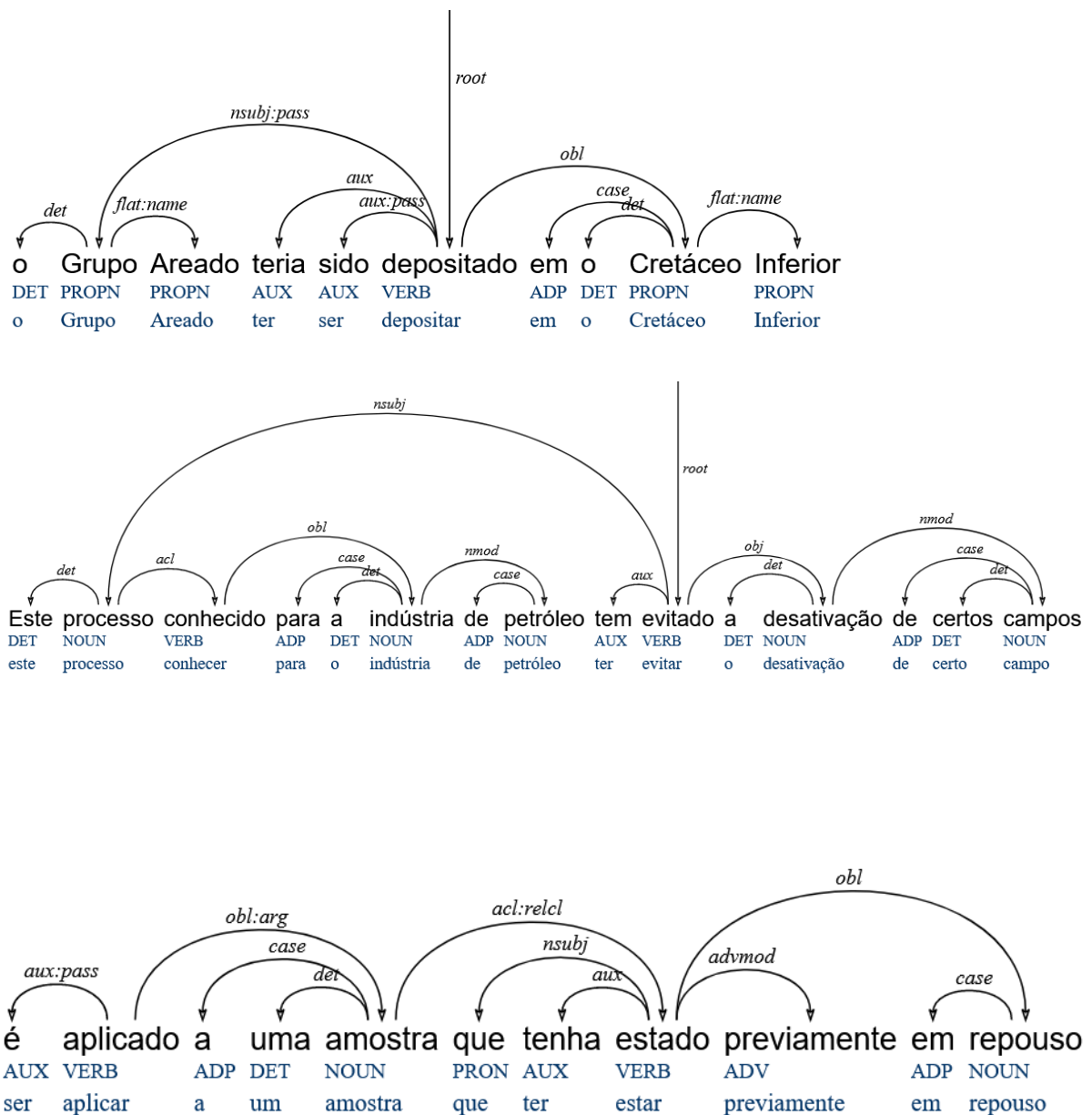
- Root:** The main clause structure.
- parataxis:** Connects the main clause to the subordinate clause.
- punct:** Marks the beginning of the subordinate clause.
- xcomp:** Connects the main clause verb "Pode" to the subordinate clause verb "ocorre".
- ccomp:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- mark:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- case:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- obl:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- obj:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- amod:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- amod:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- punct:** Marks the beginning of the subordinate clause.
- acl:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- obl agent:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- case:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- case:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- amod:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".
- amod:** Connects the main clause verb "notar" to the subordinate clause verb "ocorre".

The words are color-coded: blue for auxiliary verbs (Pode, SE, ocorre), green for main verbs (notar, espessamento, sedimentar, depocentro, controlado), red for conjunctions (que), orange for prepositions (a, de, por), purple for pronouns (se), and grey for other words (o, maior, falha, borda, sul).

Seguindo a ideia de “economia” dos auxiliares em UD, no Petrolês só são considerados auxiliares os verbos envolvidos na formação de tempo composto (ter, haver, ir, ser), de voz passiva e verbos de ligação (ser e estar).

Verbo TER como auxiliar

O verbo TER na posição de auxiliar tem POS AUX e deprel *aux* como informam os exemplos abaixo.



Locuções verbais

Apenas consideramos locuções verbais (combinação AUX + VERB) as combinações associadas a **tempo composto** e **voz passiva**. O que a GT chama de locuções aspectuais, modais etc é anotado como VERB + VERB, e a relação entre os verbos é *xcomp*.

Formas Participiais

Voz passiva com “estar”: em casos de “ESTAR ” + particípio, “ESTAR” é anotado como AUX e o particípio como VERB, a não ser em casos em que a forma participial é sinônima de um adjetivo “clássico” (um adjetivo que não está na forma participial).

Ela *estava*(AUX cop) *dividida* = estava *indecisa* (ADJ)

O mapa *estava*(AUX aux:pass) *dividido* em setores = Alguém *dividiu* o mapa (VERB)

→ Termos específicos do domínio

estratificação *cruzada*; carvão *ativado*; clastos não *fraturados/cimentados*; caulinita bem *cristalizada* → ADJ/amod

água *produzida*; argila *adicionada/incorporada*; amostra *saturada*; planos *estriados*; campo magnético *induzido*; arenitos bem *selecionados*; *referidos* esqueletos; sistema *integrado*; duas barras *superpostas*; circuito *fechado*; → VERB/acl

→ Outros casos:

• *Determinado*

1. Verbo “determinar”: Essas estimativas são *determinadas* a partir das relações → VERB/root
2. “determinado” seguido de substantivo: Geralmente são os seguintes os fatores que controlam as propriedades que uma *determinada* argila possui → DET/det

• *Dado*

1. “dado” seguido de adjetivo: *dado* geográfico → NOUN
2. Verbo “dar”: a deformação teria se *dado* ao longo de as falhas de transferência → VERB
3. “dado” seguido de substantivo: um *dado* material pode pertencer a um ou mais de estes grupos → DET

• *Detalhado*

1. Há também uma descrição *detalhada* dos métodos usados → ADJ

2. Os procedimentos usados são *detalhados* a seguir → VERB

- *Devido*

1. Este equipamento permite calcular o volume de sólidos a partir do volume de uma câmara (96 cm³), sendo feitas as *devidas* correções para alterações de temperatura.
→ ADJ

Comparativas de igualdade

Na anotação das "comparativas de igualdade", construções híbridas que têm características de subordinação/adverbial e de coordenação/adição, **privilegiamos a dimensão de adição** porque informativamente é o que nos parece mais interessante (e ainda nos ajuda na ajuda na anotação de entidades). E, porque são construções reconhecidamente "híbridas", considerá-las aditivas não chega a ser uma heresia.

Nos dois exemplos abaixo, tratar como aditiva nos permite mais diretamente inferir coisas sobre "demais bacias da margem continental" (exemplo 1) e sobre a "argila A2" (exemplo 2)

(1) A Bacia de Pelotas, **assim como** as demais bacias da margem continental brasileira, tem sua origem atrelada ao....

(2) **Assim como** a argila A2 , a argila A3 possui alta concentração de ferro

Estão anotadas como CCONJ/conj:

X, assim como Y :

--> quando o "assim como" equivale a "conforme", fazemos a leitura adverbial.

X, bem como Y,

X, além de Y,

Tanto X quanto Y

Comparativas sem verbo

Nos casos de frases em que a comparação é feita sem um verbo, consideramos o substantivo como complemento do adjetivo (obl).

398-20160721-MONOGRAFIA_0-280: A proposta é que este sistema possua horários predefinidos de operação, para estimular a recuperação de petróleo, atuando em conjunto com outros métodos de menor custo **que o eletromagnético**.

27	menor	ADJ	28	amod
28	custo	NOUN	25	nmod
29	que	ADP	31	case
30	o	DET	31	det
31	eletromagnético	NOUN	27	obl

10-20150122-MONOGRAFIA_0_intro-91

Nesses trabalhos Rocco e colaboradores, filmes de blendas poliméricas contendo PBE, dopados com 10% em massa de LiClO₄, mostraram em câmara seca condutividade de 10-5 S/cm e janela de estabilidade eletroquímica **maior que 5 V**.

32	janela	NOUN	27	conj
33	de	ADP	34	case
34	estabilidade	NOUN	32	nmod
35	eletroquímica	ADJ	34	amod
36	maior	ADJ	34	amod
37	que	ADP	39	case
38	5	NUM	39	nummod
39	V	SYM	36	obl
40	.	PUNCT	23	punct

Comparativas com verbo

Quando a comparativa acontece com um verbo, o substantivo é *advcl* dele.

26	apresenta	VERB	0	não interessa
27	menor	ADJ	28	amod
28	custo	NOUN	26	obj
29	que	SCONJ	31	mark
30	o	DET	31	det
31	eletromagnético	NOUN	26	advcl

Frases adverbiais consecutivas

Nos casos das frases adverbiais consecutivas, o verbo da oração consecutiva aponta para o head da oração principal com a *deprel advcl*. Exemplo:

Diagram illustrating the syntactic structure of the sentence "a taxa de deformação é tão baixa que não se percebe grandes mudanças" (The deformation rate is so low that one does not perceive large changes). The diagram shows the hierarchical structure of the sentence, with labels indicating the grammatical function of each word or phrase.

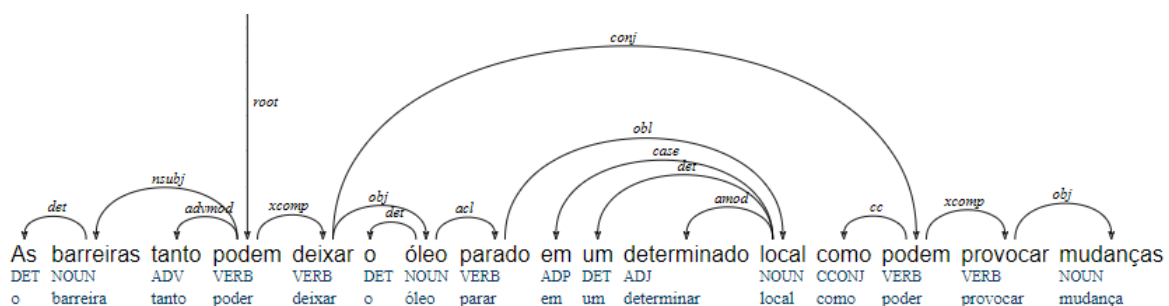
Labels and corresponding words/phrases:

- det*: a
- nmod*: taxa
- case*: de
- nmod*: deformação
- cop*: é
- advmod*: tão
- advmod*: baixa
- mark*: que
- advcl*: não
- advcl*: se
- advmod*: percebe
- expl*: grandes
- expl*: mudanças
- obj*: grandes
- amod*: mudanças

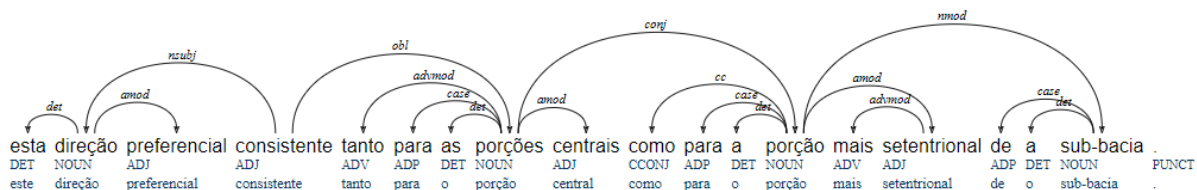
The diagram shows the hierarchical structure of the sentence, with labels indicating the grammatical function of each word or phrase. The words are arranged in two rows, with the first row showing the words and the second row showing their grammatical categories (DET, NOUN, ADP, NOUN, AUX, ADV, ADJ, SCONJ, ADV, PRON, VERB, ADJ, NOUN). The labels are placed above the words, and arrows indicate the grammatical relationships between them.

Tanto: ADV/advmod; aponta para o que vem depois, seja nominal ou verbal
 Como/Quanto: CONJ/cc; aponta para o que vem depois, seja nominal ou verbal
 (“quanto” só vale como conjunção aditiva quando acompanhado de “tanto”)

(Ex.: 411-20150702-MONOGRAFIA_0-275: As barreiras **tanto** *podem* deixar o óleo parado em um determinado local **como** *podem* provocar mudanças no posicionamento do mesmo em direção de áreas menos sensíveis, ou servem para fazer a deslocação desse óleo para uma área longe de para prováveis incidentes (ITOPF, 2011).)



(Ex.: 283-20150917-MONOGRAFIA_0-204: Depreende-se assim que o paleofluxo fluvial para a Sub-Bacia Abaeté era para NW/N, sendo esta direção preferencial consistente **tanto** para as porções centrais **como** para a porção mais setentrional da sub-bacia.)



Tanto o óleo cru **quanto** seus *derivados* estão entre os contaminantes...

Este material (microssílica e fibra mineral) serve **tanto** como(ADP/case) *fonte* de sílica adicional (no caso da microssílica) **quanto** *reforço* mecânico, principalmente no caso da fibra mineral, a qual incrementou a resistência à tração.

Este tipo de grão favorece a atenuação natural de hidrocarbonetos que se apresentaram em todas as amostras, mas devido à quantidade de matéria orgânica cerca de 10%, rica em biopolímeros (carboidratos 398-1760 µg/g, proteínas 118-824 µg/g e lipídios 12.4-154 µg/g) esses compostos se depositaram nos sedimentos, **tanto** de *superfície* **como** os *profundos*(conj - superfície).

Frases adverbiais proporcionais

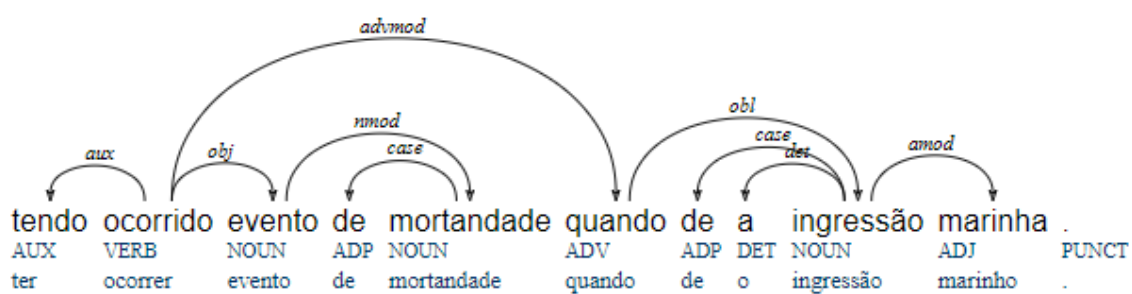
161-20150810-MONOGRAFIA_0-499: Como mencionado anteriormente, a capacidade de produção extra de petróleo, por injeção de CO₂, está diretamente relacionada à capacidade de armazenamento de gás carbônico do reservatório, logo **quanto maior for a capacidade de aporte de CO₂ maior será a produção de óleo..**

33	logo	CCONJ	35	cc
34	quanto	CONJ	35	mark
35	maior	ADJ	43	advcl
36	for	AUX	35	cop
37	a	DET	38	det
38	capacidade	NOUN	35	nsubj
39	de	ADP	40	case
40	aporte	NOUN	38	nmod
41	de	ADP	42	case
42	CO ₂	PROPN	40	nmod
43	maior	ADJ	20	conj
44	será	AUX	43	cop
45	a	DET	46	det
46	produ	NOUN	43	nsubj
47	ão	PUNCT	46	goeswith
48	de	ADP	49	case
49	óleo	NOUN	46	nmod

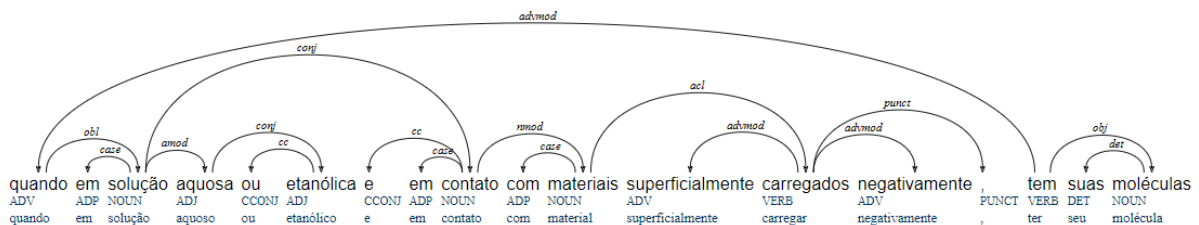
“QUANDO EM...”/ “QUANDO DE...” – Oração adverbial com elipse do verbo.

O advérbio quando é *advmod*, filho do núcleo da oração que o precede ou sucede. E o termo que seria filho do verbo omitido passa a ser *obl* do advérbio quando.

Ex1.: 284-20150917-MONOGRAFIA_0-393 - As condições ambientais, às quais os ostracodes foram submetidos, são interpretadas por Arai & Coimbra (1990) como de água doce, tendo ocorrido evento de mortandade **quando da ingressão marinha**.



(Ex2.: 24-20150122-MONOGRAFIA_0-303 - Este método baseia-se no fato que o azul de metileno [C₁₆H₁₈N₃SCl(3H₂O)], **quando em solução aquosa** ou etanólica e em contato com materiais superficialmente carregados negativamente, tem suas moléculas rapidamente absorvidas mediante um mecanismo de troca iônica irreversível.)

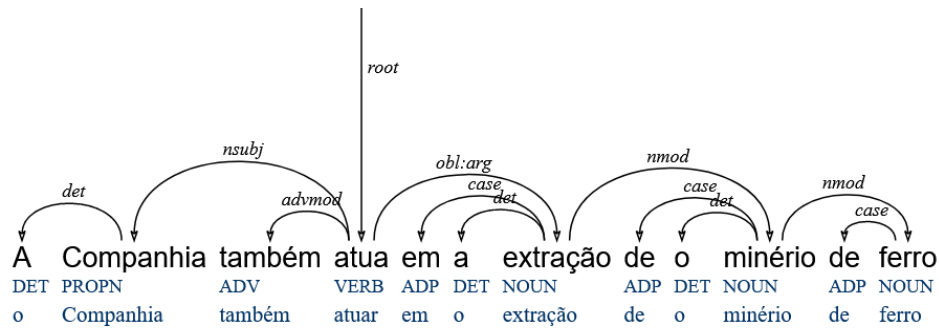


Obj e obl e obl:arg

Objetos não preposicionados são sempre *obj*.

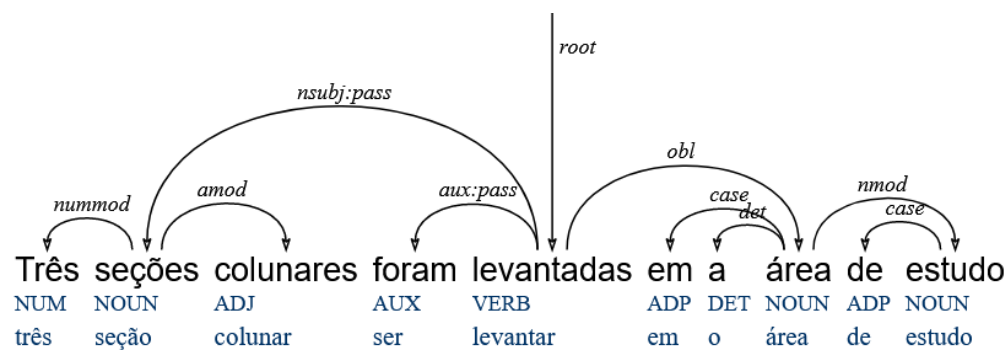
Objetos preposicionados e SEM valor de advérbio (tempo, lugar, etc) são *obl:arg*.

(161-20150810-MONOGRAFIA_0-330)



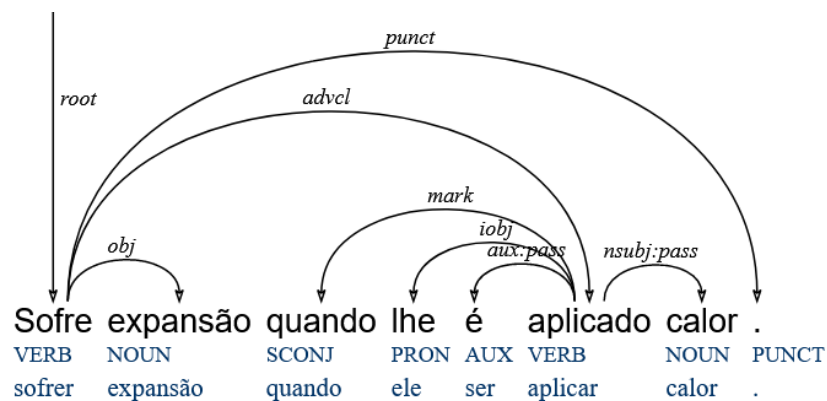
Casos diferentes dos dois acima são *obl*.

(284-20150917-MONOGRAFIA_0-17)



APENAS pronomes oblíquos não preposicionados são *iobj*.

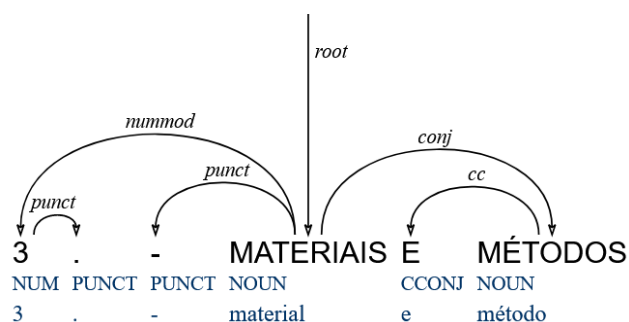
(241-20140227-MONOGRAFIA_0-109)



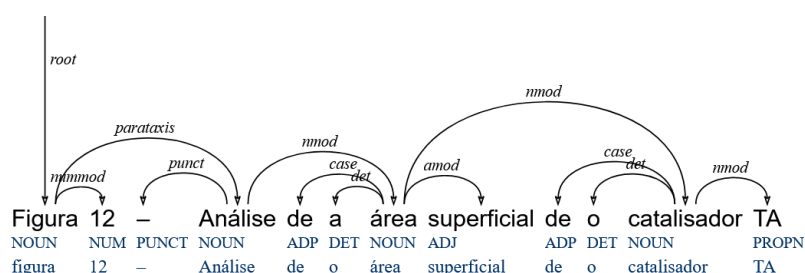
Frases sem verbo

Nas frases sem verbo, o root é o núcleo do sintagma nominal - ou o primeiro nome (NOUN ou PROPN). Exemplos:

(25-20150123-MONOGRAFIA_0-11)



(271-20150810-MONOGRAFIA_0-307)



Frases mal escritas

Nas frases abandonadas e sem verbo, o root vai ser o primeiro elemento nominal mais acima. No exemplo abaixo, o root da frase será “ambiente”.

(411-20150702-MONOGRAFIA_0-327)

Uma vez que dentro desse **ambiente** estuarino, local onde é considerado um berçário das mais variadas espécies marítimas, bem como aves que dependem das copas do mangue e de todo local ao redor do mesmo para reprodução e alimentação).

Sentenciação

Quando a frase não termina no ponto final

334-20150521-MONOGRAFIA_0-168: O Golfo do México e o Mar do Norte são duas regiões que podem ser utilizadas como referência ao tratar da exploração em campos marginais, ambas as regiões possuem áreas de produção offshore bem desenvolvidas (Tabela 6) (BOTHAMLEY et al., 2004). (Fonte: Adaptado de BOTHAMLEY et al., 2004)

Tipos de frases eliminadas

161-20150810-MONOGRAFIA_0-555: Utilizando o valor da pressão original do campo de Marlim que é igual a 4082 psi ou 28,13 MPa (Oliveira, 2008), o valor de produção total acumulada de petróleo em Marlim, que é de 2131 milhões de barris e o valor de volume original de óleo in place, que é de 8100 milhões de barris, foi possível calcular o valor atual de pressão do reservatório: Com o auxílio do gráfico, o valor atual de pressão do reservatório foi de: Utilizando o valor da pressão original do campo de Barracuda que foi estimado, a partir do valor de pressão original de Marlim (devido as características e profundidades semelhantes), em 29 MPa, o valor de produção total acumulada de Barracuda, que é 386 milhões de barris e o valor de volume original de óleo in place, que é de 2250 milhões de barris, foi possível calcular o valor atual de pressão do reservatório: Com o auxílio do gráfico, o valor atual de pressão do reservatório foi de: Pressão atual do (MPa) = $82\% \times 29\text{MPa} = 23,8 \text{ MPa}$

Outras questões

Lema de “vários”

Lema = vário; PronType = Ind

Matriz-sustentados

Quando tem hífen, é um token só, ADJ, Gender=Masc, Number=Plur e amod

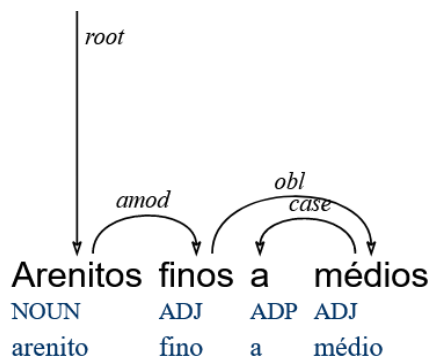
Quando não tem, “sustentados” é compound

“Finos a médios”

- “arenitos finos a médios”:

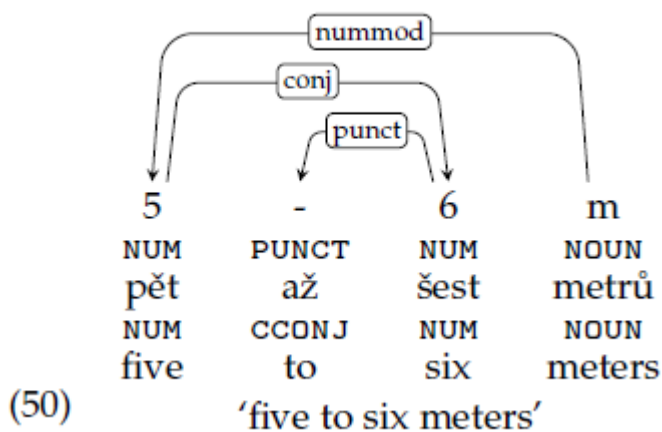
Nesses casos, o segundo adjetivo estabelece relação de *obl* com o primeiro adjetivo.

(Ex.: 284-20150917-MONOGRAFIA_0-224)

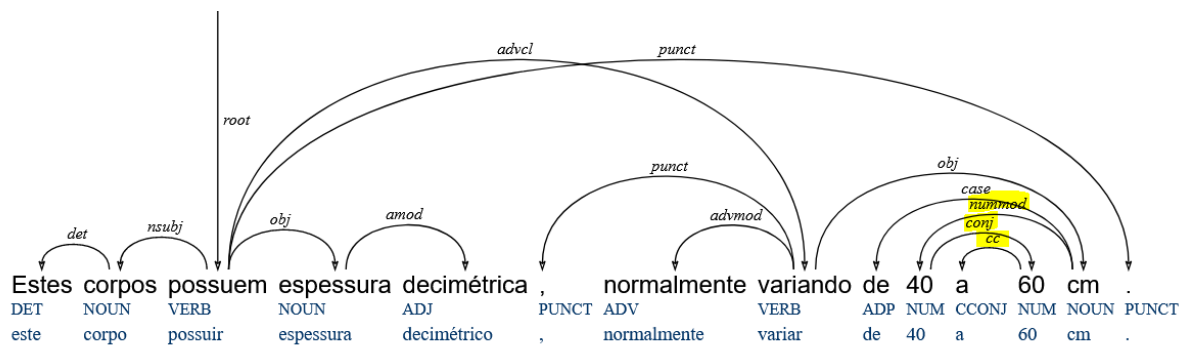


- "... variando de 40 a 60 cm"

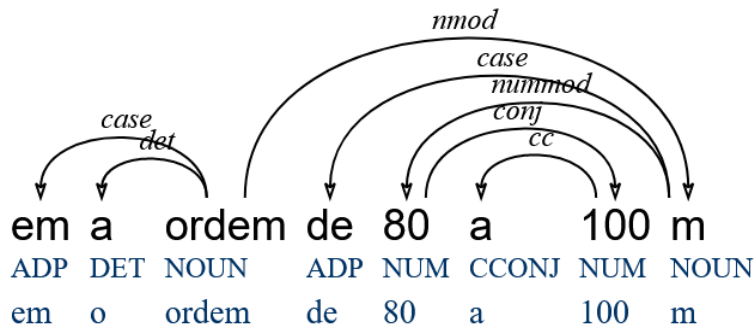
Nesses casos, a unidade de medida é o núcleo do sintagma, o primeiro número modifica a unidade de medida(*nummod*) e o segundo número estabelece relação de coordenação com o primeiro(*conj*). A preposição é encarada como o elemento que coordena os números, portanto, passa a ser classificada como CCONJ/cc. Essa decisão se baseia do artigo de Marneffe et al. (2021) a partir do seguinte exemplo:



(Ex.: 284-20150917-MONOGRAFIA_0-195)



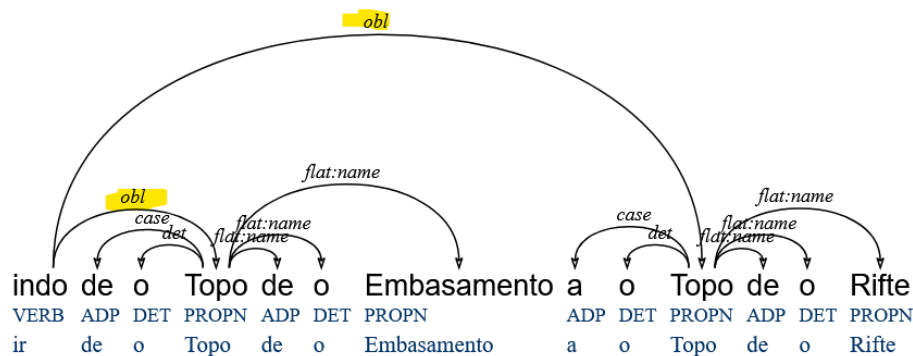
(Ex.: 283-20150917-MONOGRAFIA_0-421)



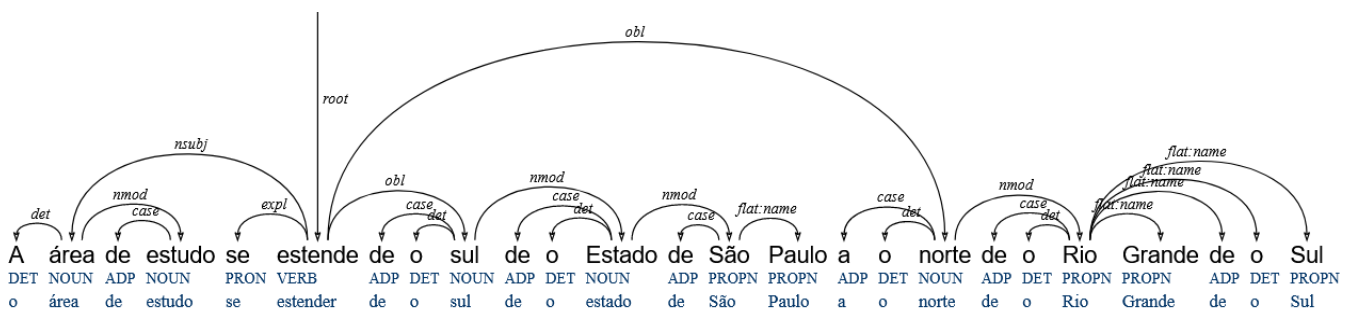
- “Foi de... a...”/ Estende-se de... a...

Nesses casos, os dois complementos de valor adverbial são classificados como *obl* do verbo a que pertencem.

(Ex.: 247-20140910-MONOGRAFIA_0-259)



(Ex.: 247-20140910-MONOGRAFIA_0-13)



Verbos leves

Neste momentos, os complementos dos verbos leves estão sendo anotados como *obl:arg* (quando preposicionados).

FAZER

uso_obj

parte_obj

... de algo_obl:arg

DAR

origem_obj

continuidade_obj

início_obj

lugar_obj

... a algo_obl:arg

TER

influência_obj

origem_obj

lugar ("a operação só terá lugar quando for disponibilizado o equipamento..." - acontecerá;
"...uma conferência internacional que terá lugar em Roma em Dezembro")

... em algo_obl:arg

A PRIORI

MWE fixed

MWEPOS=ADJ