

Aposições anafóricas e catafóricas no português e sua anotação no esquema *Universal Dependencies*

Magali Sanches Duran¹, Maria das Graças Volpe Nunes^{1,2}

¹Núcleo Interinstitucional de Linguística Computacional (NILC)

²Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo (USP)

magali.duran@uol.com.br, gracac@icmc.usp.br

Abstract. *This paper discusses the syntactic analysis of anaphoric and cataphoric appositions in Portuguese, which fall under the referentiality strategy called encapsulation. We present a review of studies that address the phenomenon and propose guidelines for annotating it from the perspective of dependency syntax, using labels from the Universal Dependencies approach.*

Resumo. *Este artigo discute a análise sintática de aposições anafóricas e catafóricas na língua portuguesa, as quais se enquadram na estratégia de referencialidade chamada de encapsulamento. Apresenta-se uma revisão de trabalhos que abordam o fenômeno e propõem-se diretrizes para anotá-lo sob a ótica da sintaxe de dependências, usando etiquetas da abordagem Universal Dependencies.*

1. Introdução

O exercício de anotação sintática de corpus apresenta ao anotador o desafio de reconhecer, na prática, fenômenos estudados pela linguística e descritos nas gramáticas. Porém, há alguns fenômenos que não são imediatamente reconhecidos, seja pelo fato de o anotador desconhecer sua descrição, seja pelo fato de tais fenômenos ainda não constarem dos manuais de anotação, ou, ainda, por não terem sido largamente descritos e consensualmente reconhecidos por gramáticos e linguistas.

Neste artigo discutem-se dois fenômenos que impuseram desafios ao projeto POeTiSA¹: as aposições catafóricas (exemplos 1, 2 e 3) e as aposições anafóricas (exemplos 4, 5 e 6), cujos correferentes estão indicados em negrito:

1. Com ares mediterrâneos, o Kez nasce de uma boa **ideia**: popularizar o bagel, aquele pão judeu redondo com um anel vazado no centro, denso, de crosta ligeiramente úmida.
2. Nas últimas semanas, ele aprendeu uma nova **técnica**: dedura o comentário para a empresa onde seu autor trabalha.
3. Há um **agravante**: como a JBS tem fábricas nos EUA, as propinas pagas no Brasil são uma violação da lei americana que proíbe empresas de lá de pagar suborno no exterior.
4. Até hoje não se sabe quem foi o autor - **o que** significa que ninguém foi punido.

¹<https://sites.google.com/icmc.usp.br/poetisa>

5. Depois, mesmo sofrendo o gol quando ainda faltavam mais de cinco minutos, soube segurar o resultado, **algo** que não vinha acontecendo em sua vida mais recente.
6. Passaram de R\$78,5 bilhões para R\$111,8 bilhões nos últimos dez anos, um **crescimento** real de 42% no período.

Esses fenômenos são denominados, em mais de uma gramática, “oração apositiva” (1, 2 e 3) e “aposto de oração” (4, 5 e 6). São casos, portanto, que orbitam a esfera do aposto, função na sintaxe que engloba muitos fenômenos e nem sempre é objeto de consenso entre os gramáticos e linguistas.

Todos os exemplos apresentados fazem parte de uma estratégia de referenciação chamada “encapsulamento”. No encapsulamento há uma palavra encapsuladora (pronome ou sintagma nominal) que é usada no lugar de toda uma oração por ela encapsulada. Segundo Koch (2008, p. 106) o encapsulamento é uma estratégia para construir novos objetos de discurso, que passam a constituir um referente para novas predicções. Quando ocorrem dentro de uma mesma sentença, tornam-se um problema de sintaxe. E é precisamente a análise sintática dessas construções de encapsulamento que inspirou as reflexões e propostas que se apresentam neste artigo. A análise utiliza a sintaxe de dependências (Tesnière, 2015) e o esquema de anotação da Universal Dependencies (UD) (De Marneffe *et al.* 2021; Nivre *et al.* 2020), exigindo basicamente duas definições em relação aos fenômenos em foco: qual etiqueta adotar para nomear a relação de dependência e quais são o *head* e o dependente da relação em cada caso.

Na Seção 2, apresentam-se brevemente a sintaxe de dependências e o esquema de anotação da abordagem UD, adotados como método de análise. Na Seção 3, revisa-se criticamente o tema das posições em gramáticas e estudos linguísticos. Na Seção 4 são discutidos exemplos de *cópus* e apresentadas propostas para sua anotação usando a sintaxe de dependências. Por fim, a Seção 5 traz conclusões e possibilidades de trabalhos futuros.

2. A sintaxe de dependências e a abordagem Universal Dependencies

A sintaxe de dependências faz uso de relações que ligam as palavras de uma sentença, duas a duas, sempre determinando qual é o *head* (ou governante) e qual é o dependente da relação. A UD é inspirada na sintaxe de dependências de Tesnière (1959 e 2015) e possui um conjunto de etiquetas para anotar a categoria morfossintática das palavras (17 *part-of-speech tags* ou *PoS tags*) e um conjunto de etiquetas para anotar as relações sintáticas entre palavras (37 *dependency relations* ou *deprel*).

A UD possui diretrizes² sobre como realizar a anotação e um fórum virtual onde são discutidas dúvidas e dificuldades de anotadores que adotaram a abordagem nas mais diversas línguas³. As diretrizes da UD já foram descritas e exemplificadas em manuais de anotação para língua portuguesa (Duran, 2021 e Duran, 2022) e há dois *cópus* de português brasileiro já disponíveis no site da UD: o Bosque-UD (Rademaker *et al.*

² Disponível em <https://universaldependencies.org/>

³ Até o momento da escrita deste artigo, 141 diferentes línguas apresentavam pelo menos um *cópus* anotado seguindo essa abordagem.

2017) e o PetroGold (Souza *et al.* 2021). Além disso, outras iniciativas de anotação de corpus de português nos moldes da UD encontram-se em andamento, como os relatados por Pardo *et al.* (2021), Oliveira *et al.* (2022) e Coneglian *et al.* (2022).

As etiquetas da UD guardam alguma equivalência com a nomenclatura gramatical brasileira, mas são muitas as diferenças entre os fenômenos cobertos e, por falta de espaço, não serão explicadas aqui. Para a discussão aqui instada, é essencial dizer que o arco das relações de dependência é direcional, partindo do *head* (ou governante) da relação e apontando sua flecha para o dependente da relação. Além disso, quando o head ou dependente de uma relação é uma oração, a ponta do arco correspondente à oração é colocada no núcleo do predicado; já quando o *head* ou dependente de uma relação é um sintagma, a ponta do arco correspondente é colocada no núcleo do sintagma. A Figura 1 ilustra uma árvore de dependências UD anotada com a ferramenta Arborator-Nilc (Miranda & Pardo, 2022).

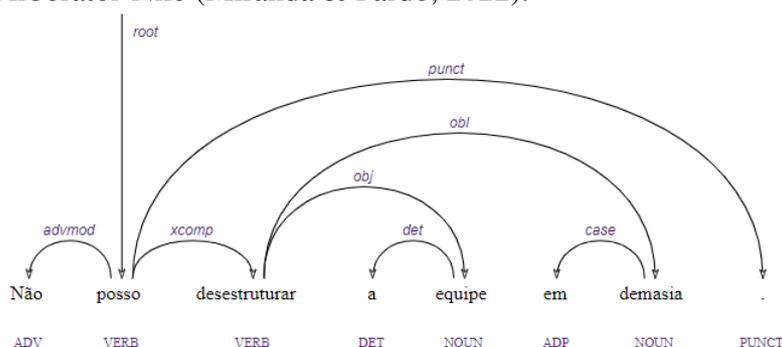


Fig. 1. Árvore de dependências de sentença no esquema UD

3. O estatuto sintático das posições catafóricas e anafóricas

Nesta seção discute-se o tratamento das posições catafóricas e anafóricas em gramáticas e estudos linguísticos do português, a fim de reunir subsídios para a anotação dessas construções dentro do esquema de anotação de corpus da UD.

3.1. Posições catafóricas

Uma posição catafórica é aquela em que um pronome ou substantivo genérico contido em uma oração é explicado ou comentado por uma outra oração. Esse fenômeno é comumente reconhecido nas gramáticas como uma oração que tem função de aposto (Cegalla, 2020, p. 385; Azeredo, 2013, p.139; Rocha Lima, 2011, p. 328; Faraco e Moura, 1994, p. 331). Essa oração é classificada como “subordinada substantiva apositiva”, antecedida por dois-pontos, e é a única das orações substantivas que não é selecionada pela semântica de verbos e nomes. Pode-se afirmar, portanto, que se trata de uma oração adjuntiva, ou seja, que não preenche uma lacuna da oração principal.

Teoricamente, as orações apositivas são orações dependentes e subordinam-se a um pronome ou substantivo contido na oração principal. Em sua forma finita, essas orações são introduzidas por uma conjunção integrante (“que” ou “se”) e, em sua forma nominal, são orações reduzidas (quase sempre de infinitivo). Azeredo (2013 p. 139) traz exemplos que se enquadram perfeitamente nessa definição, como o (7):

7. “Ele só pediu um favor: que o tirassem daquele hospital.”

Porém, alguns gramáticos chamam também de apositivas orações independentes, precedidas de dois-pontos, com verbo finito, ou seja, nem reduzidas e nem introduzidas por uma conjunção subordinativa. Isso pode ser observado nos exemplos 8 e 9, fornecidos por Cegalla (2020, p. 385), e nos exemplos 10, fornecido por Rocha Lima (2011, p. 332) e 11, fornecido por Faraco e Moura (1996, p.331).

8. "E confesso uma verdade: eu era um homem puro." (Povina Cavalcanti)
9. "A notícia veio de supetão: iam meter-me na escola." (Graciliano Ramos)
10. "Dei-lhe tudo: ofereci-lhe o meu nome; tornei-a dona de todo o meu dinheiro, e levei-a à minha posição social."
11. "Então aconteceu o pior: veio vento sul."

Nogueira e Leitão (2004) estudam as orações substantivas apositivas, "as quais estão associadas às estratégias de referenciação catafórica" e "cujo conteúdo é encapsulado e antecipado pela expressão referencial" (*op. cit.*, p.138).

A sentença 12 é fornecida pelas autoras como exemplo de oração apositiva, o que mostra que, para elas, assim como para Cegalla, Rocha Lima e Faraco e Moura (anteriormente citados), não importa se há ou não marcas explícitas de subordinação ligando a aposição à oração principal: em qualquer hipótese a aposição catafórica em forma oracional será classificada como oração apositiva.

12. "Só vejo uma situação: o Governo do PT elegeu-se pregando uma coisa e atualmente faz totalmente o inverso." (Nogueira e Leitão, 2004, p. 139)

Na sentença 12, "situação" é a palavra encapsuladora e "o Governo do PT ... inverso" é a oração coordenada encapsulada, conforme ilustrado na Figura 2.

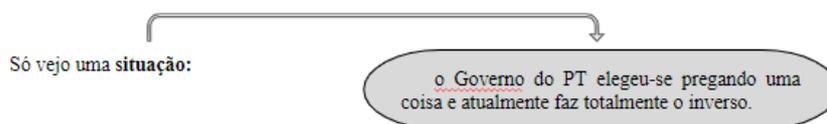


Fig. 2. Encapsulamento catafórico (da esquerda para a direita)

3.2. Aposições anafóricas

A relação inversa também pode ocorrer: uma aposição que contém um pronome ou substantivo que comenta o conteúdo da oração principal. Alguns gramáticos, como Bechara (1999), denominam esse tipo de aposição de "aposto de oração", mas há gramáticos que usam outra denominação, como Nougé (2015), que o chama de "aposto resumitivo", e ainda gramáticos, como Rocha Lima (2011), que reconhecem a possibilidade de um aposto que se refira a uma oração, porém não o nomeiam.

A ideia de que a "oração apositiva" é o inverso do "aposto de oração" parece encontrar respaldo nos estudos de linguística textual e discursiva acerca das estratégias de encapsulamento usadas para construir coesão textual. Francis (1994, p. 98) faz a distinção entre o "*advance label*", que funciona cataforicamente, e "*retrospective label*", que funciona anafóricamente. Observa-se que o "*advance label*" produz o encapsulamento catafórico (para frente), o que engloba o conceito estendido de "oração apositiva", ao passo que o "*retrospective label*" produz o "encapsulamento anafórico" (para trás), o que engloba o conceito de "aposto de oração".

As construções chamadas de “aposto de oração” são tema dos estudos de Sousa (2016). Ao descrever o aposto de oração, Sousa afirma que a função pode ser desempenhada por um substantivo, seguido de oração adjetiva, ou pelo pronome “o”, seguido de uma oração adjetiva, ou simplesmente por um substantivo. As sentenças 13, 14 e 15, fornecidas no trabalho da autora, exemplificam essas três realizações:

13. “Paradoxalmente, a partir da década de 90, a prática de extorsão mediante sequestro foi sensivelmente maior do que antes, fato que se repetiu, aliás, com outros tantos delitos mais ou menos hediondos.”
14. “(...) atenuam responsabilidades por meio do argumento de que também os outros partidos cometeram pecados e que destes não se falou o quanto era preciso, argumento insuficiente (mesmo se a imputação é legítima), já que, além do tamanho da operação, há, no caso do PT, uma circunstância agravante pelo fato de se tratar de um partido que se apresenta como modelo de virtude cívica.”
15. “Mais do que planilhas e números, estamos falando de pessoas, na ativa ou não, que terão maior poder de compra, o que gera, de imediato, melhoria na qualidade de vida do cidadão.”

No exemplo 13 há claramente uma oração encapsulada e um substantivo encapsulador (“fato”), como mostra a Figura 3.



Fig. 3. Encapsulamento anafórico (da direita para a esquerda)

Já o exemplo 14, ao que parece, não constitui um aposto de oração, mas um aposto comum, que liga “argumento de que...” e “argumento insuficiente”.

E quanto ao exemplo 15, não parece ter comportamento similar ao do exemplo 13. Enquanto um substantivo ou pronome indefinido pode ser adjetivado (exemplo 16), o pronome “o” não o pode (exemplo 17).

16. (...) a prática de extorsão mediante sequestro foi sensivelmente maior do que antes, **fato/algo interessante** que se repetiu (...)
17. *(...) terão maior poder de compra, **o interessante que** gera (...)

Além disso, o pronome relativo “que” da oração adjetiva restritiva cujo antecedente é um substantivo ou pronome indefinido pode ser preposicionado (exemplo 18), enquanto o pronome “que” que acompanha o pronome “o” não pode (19).

18. (...) a prática de extorsão mediante sequestro foi sensivelmente maior do que antes, **fato/algo de que** nem todos se lembram.
19. *(...) terão maior poder de compra, **o de que** todos se orgulham.

Na verdade, se houvesse uma preposição, ela antecederia “o que”, como mostra o exemplo 20.

20. (...) terão maior poder de compra, do que todos se orgulham.

Ao que parece, os pronomes “o” e “que” nesse uso são indissociáveis, formando uma locução pronominal, análoga ao pronome relativo “o qual” e suas flexões. Mas, se

“o que” for um pronome relativo, como se chamaria uma oração relativa que acrescenta informação a uma outra oração? Pelo fato de modificar um predicado e de ser adjuntiva, poderia se chamar “oração subordinada adverbial relativa”? Independentemente da denominação, o que distingue o aposto de oração da oração relativa é o fato de que o dependente da relação de dependência passa a ser o predicado da oração relativa, como mostrado na Figura 4.



Fig. 4. Encapsulador e encapsulado em forma de oração

É interessante observar ainda que, se a anotação não fosse de relações sintáticas, mas de relações de correferência, a relação sempre partiria do encapsulador em direção ao encapsulado, independentemente de o encapsulador estar à direita ou à esquerda do encapsulado.

4. Relações de dependência entre posições catafóricas e anafóricas

Inicialmente justifica-se a utilização do termo “aposição” e não “aposto” para designar as construções em foco pelo fato de que, no esquema de anotação UD, a relação de dependência **appos** só pode ser atribuída a relações entre substantivos (próprios e/ou comuns) e entre substantivos e pronomes, e apenas na direção da esquerda para a direita. Como o esquema UD tem como objetivo servir a várias línguas, não existe flexibilidade para utilizar a relação **appos** para anotar o que na literatura é descrito como “oração apositiva” e “aposto de oração”. De fato, isso geraria um problema de inconsistência entre as línguas anotadas no esquema, prejudicando a comparação de suas estruturas sintáticas.

Pelos mesmos motivos, também não é permitido criar novas etiquetas de relações de dependência, a não ser de forma consensual entre todos os cientistas que participam da iniciativa. Quando ocorrem acordos de mudanças no conjunto de etiquetas ou na forma de empregá-las, as diretrizes da UD são alteradas e todos os inscritos recebem notificação sobre as alterações. Como é alto o custo de reanotar os 245⁴ corpúscos já disponíveis na UD (em 141 línguas), mudanças no esquema são raras. Isso não impede que haja muita discussão entre aqueles que adotam o esquema, e a questão do aposto em que um dos termos é uma oração é um tema recorrente⁵.

Diante das restrições do esquema de anotação, acredita-se que a relação de dependência mais adequada para anotar as posições catafóricas e anafóricas na UD seja

⁴ Dado de 15/05/2023

⁵ As discussões acerca do aposto oracional podem ser encontradas nos seguintes links:

<https://github.com/UniversalDependencies/docs/issues/762>

<https://github.com/UniversalDependencies/docs/issues/751>

<https://github.com/UniversalDependencies/docs/issues/523>

a **parataxis**. Essa opção parece ser bem adequada quando não há marcas de subordinação entre o termo encapsulador e o termo encapsulado nas aposições, porém é uma solução pouco defensável sob o ponto de vista das gramáticas, já que a parataxe não deveria se confundir com a hipotaxe (subordinação). Mas é uma concessão que se faz enquanto não há melhor opção.

Assim, a aposição catafórica anotada com a relação **parataxis** teria como *head* o nominal que encapsula uma oração, como ilustrado na Figura 5, que mostra parte da árvore sintática do exemplo (1).

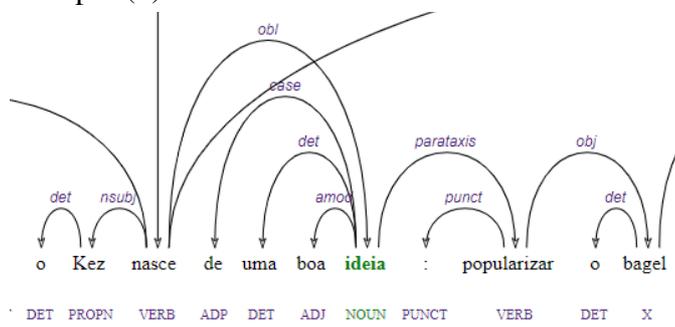


Fig. 5. Anotação no esquema UD de aposição catafórica (oração apositiva)

Já a aposição anafórica anotada com a relação **parataxis** tem como *head* o predicado da oração encapsulada e, como dependente, a palavra encapsuladora, como mostra a Figura 6 (parte da árvore sintática do exemplo 5).

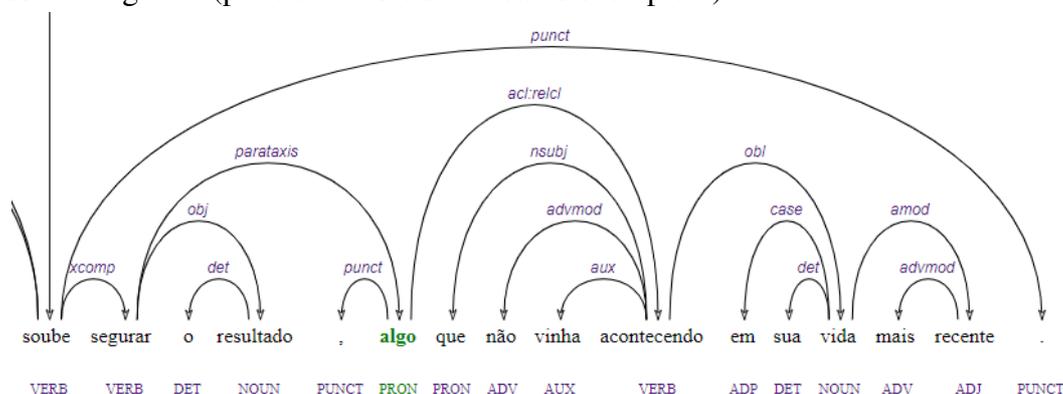


Fig. 6. Anotação no esquema UD de aposição anafórica (aposto de oração)

Contudo, no caso das aposições anafóricas iniciadas por “o que”, decidiu-se usar uma outra anotação, reconhecendo os dois pronomes como uma locução que funciona como um pronome relativo. Para isso, os dois pronomes são unidos pela relação de dependência **fixed**, utilizada para palavras funcionais constituídas de mais de um token. Nesse caso, a relação de dependência não se dá entre a oração encapsulada e o pronome encapsulador, mas entre a oração encapsulada e a oração que contém o encapsulador, à semelhança do que ocorre nas orações relativas cujo antecedente é um nominal.

Inicialmente, optou-se por utilizar a relação **parataxis** inclusive nesse caso de relativa, pois a relativa **acl:relcl**, prevista na UD, só se aplica a antecedentes nominais. Porém, tem se mostrado ser plausível a possibilidade de reconhecer, como oração adverbial relativa, as aposições iniciadas por “o que” cujo antecedente é oracional. Isso cumpriria dois requisitos: reconhecer seu estatuto de oração subordinada e seu estatuto

de adjunção. A Figura 7 ilustra a anotação da aposição anafórica encabeçada por “o que” na condição de pronome relativo.

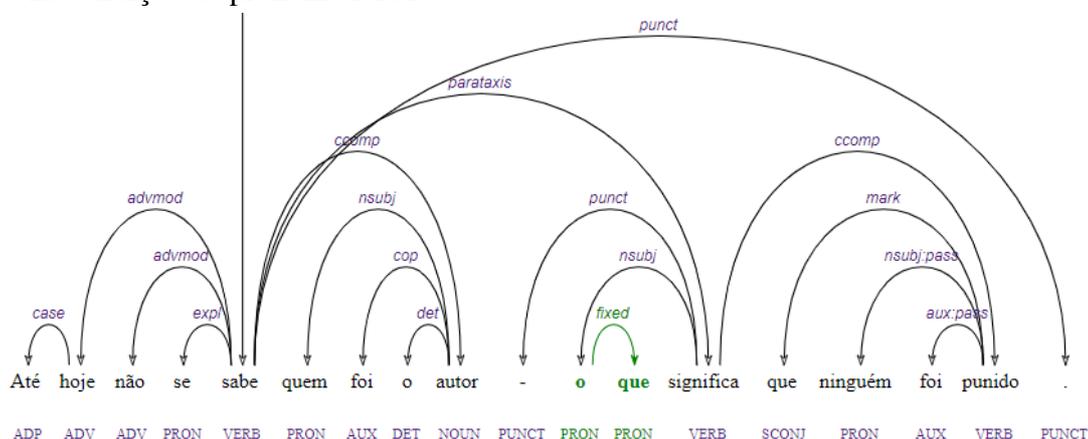


Fig. 7. Anotação no esquema UD de oração relativa com antecedente oracional

Cabe esclarecer que a UD permite que sejam acrescentadas sub-relações a fim de discriminar tipos de uma mesma relação de dependência. Assim, caso seja interessante para o projeto de anotação, a relação **parataxis** utilizada para anotar as aposições catafóricas e anafóricas poderia receber uma sub-relação, como **parataxis:appos**, por exemplo. O mesmo poderia ser feito no caso das relativas encabeçadas por “o que”, que poderiam ser anotadas como **parataxis:relcl** (**relcl** de *relative clause*) ou como **advcl:relcl** (adverbial relativa).

5. Conclusões e trabalhos futuros

As construções apositivas catafóricas e anafóricas, denominadas, respectivamente, orações apositivas e apostos de oração em gramáticas e estudos linguísticos, motivaram uma série de reflexões neste artigo. Questionou-se o fato de orações sem marcas de subordinação estarem sendo reconhecidas como orações subordinadas apositivas, o que sugere que se trata de um tema que merece estudos mais profundos sob o ponto de vista sintático. Da mesma forma, questionou-se a anotação de orações encabeçadas por “o que” que encapsulam o conteúdo de uma oração precedente, pois testes demonstraram que “o que” se comporta como uma locução e não como um encapsulador nominal comum (“o”) seguido de oração relativa introduzida por “que”. Propõe-se reconhecer “o que” como uma expressão fixa com função de pronome relativo, à semelhança de “o qual” e suas flexões, quando encabeçam orações relativas.

Um trabalho futuro é discutir com outros grupos que anotam córpus seguindo o esquema da UD a possibilidade de reconhecer as orações relativas encabeçadas por “o que”, e cujo antecedente é uma oração, como orações adverbiais relativas.

Uma vez que se tenha chegado a um consenso, com outros grupos que empregam o esquema UD em português, a respeito das relações a serem adotadas na anotação das aposições catafóricas e anafóricas, as decisões deverão ser incluídas nos manuais que contêm as diretrizes de anotação, a fim de promover a disseminação das respectivas análises entre os anotadores.

Agradecimentos

Este trabalho foi realizado no âmbito do Centro de Inteligência Artificial da Universidade de São Paulo (C4AI - <http://c4ai.inova.usp.br/>), com o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (processo FAPESP #2019/07665-4) e da IBM. Este projeto também foi apoiado pelo Ministério da Ciência, Tecnologia e Inovações, com recursos da Lei N. 8.248, de 23 de outubro de 1991, no âmbito do PPI-Softex, coordenado pela Softex e publicado como Residência em TIC 13, DOU 01245.010222/2022-44.

Referências

- Azeredo, J. C. S. Fundamentos de Gramática do Português. E-book. Jorge Zahar Editores, Rio de Janeiro, 2013.
- Bechara, E. Moderna Gramática Portuguesa. 39ª edição. Editora Nova Fronteira: Rio de Janeiro, 2019.
- Castilho, A. T. Gramática do Português Brasileiro. Editora Contexto, São Paulo, 2010.
- Cegalla, D. P. Novíssima Gramática da Língua Portuguesa. 49ª edição. Companhia Editora Nacional, 2020.
- Coneglian, A. V. L.; Guimarães, A. L. A. R.; Ferreira, T. C.; Pagano, A. S. Anotação de textos não canônicos: um estudo exploratório de Grande sertão: veredas pelas dependências universais. *In: Proceedings of the Universal Dependencies Brazilian Festival*, pp. 1-11. Association for Computational Linguistics, Fortaleza, Brazil, 2022.
- Cunha, C. F.; Lindley Cintra, L. F. Nova gramática do Português contemporâneo. 7ª edição. Lexikon Editora Digital, Rio de Janeiro, 2017.
- Duran, M. S. Manual de Anotação de PoS tags: Orientações para anotação de etiquetas morfossintáticas em Língua Portuguesa, seguindo as diretrizes da abordagem Universal Dependencies (UD). Relatório Técnico do ICMC 434. ICMC, Universidade de São Paulo, 55p., Set. 2021. Disponível em: https://drive.google.com/file/d/1BddPswN-_Ioo-A5GslDA1cO1kqbcCahb/view?usp=sharing
- Duran, M. S. Manual de Anotação de Relações de Dependência – Versão Revisada e Estendida. Relatório Técnico do ICMC 440. ICMC-USP. São Carlos-SP, Out 2022, 166p. Disponível em: <https://drive.google.com/file/d/1ile8Wfxu1qdrZOMLGqkvVuQ4fXvHgVMo/view?usp=sharing>
- Faraco, C. E.; Moura, F. M.. Gramática. Editora Ática, São Paulo, 1994.
- Francis, G. Labelling discourse: an aspect of nominal-group lexical cohesion. In: Coulthard, M. (Ed.), *Advances in Written Text Analysis*. Routledge, London, p. 83-101, 1994.

- Marneffe, M.; Manning, C.; Nivre, J.; Zeman, D. Universal Dependencies. *Computational Linguistics* 47 (2) p. 255-308. MIT PRESS, 2021.
- Miranda, L. G. M.; Pardo, T. A. S. An Improved and Extended Annotation Tool for Universal Dependencies-based Treebank Construction. *In: Proceedings of the PROPOR Demonstrations Workshop*, p. 1-3, 2022.
- Neves, M. H. de M. Gramática de Usos do Português. Ed. Unesp, São Paulo, 2000.
- Nivre, J.; Marneffe, M.; Ginter, F.; Hajič, J.; Manning, C.; Pyysalo, S.; Schuster, S.; Tyers, F.; Zeman, D. Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection. *In: Proceedings of the 12nd International Conference on Language Resources and Evaluation (LREC 2020)*, p. 4034-4043, 2020.
- Nogueira, M. T.; Leitão, R. J. A oração substantiva apositiva: aspectos textual-discursivos. *Veredas*, v.8, n.1 e n.2, p. 137-151. Juiz de Fora, 2004.
- Nougué, C. Suma Gramatical da Língua Portuguesa: Gramática Geral e Avançada. *É Realizações*, São Paulo, 2015.
- Oliveira, L. F. A.; Pagano, A.; Oliveira, L. E. S.; Moro, C. Challenges in Annotating a Treebank of Clinical Narratives in Brazilian Portuguese. *In: Computational Processing of the Portuguese Language, Volume 13208*, 2022.
- Pardo, T. A. S.; Duran, M. S.; Lopes, L.; Di Felippo, A.; Roman, N. T.; Nunes, M. G. V. Portinari - A large multi-genre treebank for Brazilian Portuguese. *In: Proceedings of the XIV Symposium in Information and Human Language (STIL 2021)*, p. 1-10, 2021.
- Perini, M.. Gramática Descritiva do Português. 2ª ed., 380 p. Ática, São Paulo, 1996.
- Rademaker, A.; Chalub, F.; Real, L.; Freitas, C.; Bick, E.; de Paiva, V. (2017). Universal dependencies for Portuguese. *In: Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017)*, p. 197–206, 2017.
- Rocha-Lima, C. H. Gramática normativa da língua portuguesa. Editora José Olympio, São Paulo, 2010.
- Sousa, R. S. N A aposição encapsuladora em artigos de opinião no português. Tese (Doutorado) – Universidade Federal do Ceará, Programa de Pós-Graduação em Linguística, Fortaleza (CE), 2016.
- Souza, E; Silveira, A.; Cavalcanti, T.; Castro, M. C.; Freitas, C.. PetroGold – Corpus padrão ouro para o domínio do petróleo. *In: Anais do Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*, 13. p. 29-38. Sociedade Brasileira de Computação, Porto Alegre, 2021.
- Tesnière, L. (2015). Elements of Structural Syntax. Tradução de Osborne, Timothy; Kahane, Sylvain. John Benjamins, Amsterdam, 2015.