

A oposição nas relações de coerência *Contrast* e *Concession*

Marquize Silva dos Santos¹, Roana Rodrigues¹, Jackson Wilke da Cruz Souza²

¹Programa de Pós-Graduação em Letras (PPGL)
Universidade Federal de Sergipe (UFS) – São Cristóvão – SE

²Programa de Pós-Graduação em Língua e Cultura (PPGLinC)
Universidade Federal da Bahia (UFBA) – Salvador – BA

marquizesilva45@gmail.com, roana@academico.ufs.br,
jackcruzsouza@gmail.com

Abstract. *The coherence relations Contrast and Concession share the same oppositional effect. Nevertheless, according to studies based on Rhetorical Structure Theory (RST), there are features that distinguish them individually, such as structure and the writer's/speaker's communicative intent. In this study, we investigated more closely the behavior of these relations in a news corpus that had been pre-annotated both with coherence relations and with the discourse signals (DSs) that signal them. The results corroborate and specify the structural differences between these relations; furthermore, their points of convergence are presented, among which the sharing of DMs stands out, especially regarding the use of connectives.*

Resumo. *As relações de coerência Contrast e Concession compartilham o mesmo efeito de oposição. Apesar disso, segundo os estudos com base na Rhetorical Structure Theory (RST), há características que as definem individualmente, como a estrutura e a intencionalidade do escritor/falante. Neste trabalho, investigou-se mais detidamente o comportamento dessas relações em um corpus jornalístico pré-anotado tanto com as relações de coerência quanto com os sinalizadores discursivos (SDs) que as identificam. Os resultados corroboram e especificam a diferença na estrutura dessas relações, além disso, são apresentados seus pontos de encontro, dentre os quais se destaca o compartilhamento de SDs, sobretudo no que se refere ao uso de conectivos.*

1. Introdução

Em gramáticas tradicionais como as de Bechara (2009 [1928]), Cunha e Cintra (2017 [1985]) e Lima (2011), as conjunções *mas*, *porém*, *todavia*, *contudo*, *no entanto*, *entretanto* são próprias das orações adversativas; já as conjunções *embora*, *conquanto*, *ainda que*, *mesmo que*, entre outras, são próprias das orações concessivas. Por outro lado, autoras como Neves (1999) e Margarido (2010), em seus estudos de base funcionalista, defendem que essas conjunções não se aplicam apenas e exclusivamente a suas respectivas orações, mas que elas podem ser compartilhadas entre si, principalmente no uso corriqueiro da língua.

Reconhecendo-se as similaridades e particularidades do fenômeno da oposição, buscou-se, nesta investigação, analisar como ele é sinalizado e identificado tanto em construções que estabelecem a relação de coerência com efeito *adversativo* (de

contraste), quanto nas construções com efeito de *concessão*. Para tanto, utilizou-se como base teórico-metodológica a Rhetorical Structure Theory (RST) [Mann e Thompson, 1987]; uma teoria descritiva também de base funcionalista que estuda o texto em suas partes, analisando as relações de sentido que o formam. Dentre as relações existentes, que também são chamadas de relações de coerência ou retóricas, há a *Contrast* e *Concession*, que indicam justamente a relação de oposição de ideias.

Na RST, as relações são classificadas entre aquelas que possuem dois núcleos, em que cada proposição é independente e igualmente importante, chamadas multinucleares; e aquelas em que há apenas um núcleo (N) e satélites (S), chamadas mononucleares, ou seja, há uma proposição que apresenta a informação central e outra que agrupa informações ao núcleo [Mann e Taboada, 2010]. Dito isso, a primeira diferença entre as relações *Contrast* e *Concession* está em sua estrutura e classificação, a primeira é multinuclear e a segunda é mononuclear.

Uma segunda diferença entre essas relações está na intencionalidade do falante/escritor e no efeito pretendido a ser alcançado por meio da mensagem. Ainda segundo Mann e Taboada (2010), a relação *Concession* tem a intenção de alcançar uma aceitação e inclinação do ouvinte/leitor, ao passo que, na relação *Contrast*, o objetivo é apresentar uma informação, mas não necessariamente buscar uma aceitação por parte do outro. O quadro 1, abaixo, ilustra essas informações.

Quadro 1. Comportamento das relações *Concession* e *Contrast* segundo a RST

Restrições	Relação <i>Concession</i>	Relação <i>Contrast</i>
Sobre N	O leitor julga N válido.	Não se aplica.
Sobre S	O escritor não afirma que S pode ser válido.	Não se aplica.
Sobre N+S	O escritor mostra uma incompatibilidade aparente ou em potencial entre N e S; o reconhecimento da incompatibilidade entre N e S melhora a aceitação de N pelo leitor.	Não se aplica.
Sobre N+N	Não se aplica.	Não mais do que dois Ns; as situações nos Ns são (a) compreendidas como similares em vários aspectos, (b) compreendidas como diferentes em vários aspectos e (c) comparadas em relação a uma ou mais dessas diferenças.
Efeito	O leitor aceita melhor N.	O leitor reconhece as similaridades e diferenças resultantes da comparação sendo feita.

Fonte: Elaboração própria a partir de Pardo (2005).

Sendo assim, nesta pesquisa, espera-se discutir os seguintes aspectos:

- quais os sinalizadores discursivos (doravante SDs), para além dos marcadores discursivos (MDs) protótipicos, sinalizam e caracterizam as relações de coerência com efeito de oposição *Contrast* e *Concession*?
- com base nos sinalizadores anotados, quais os pontos de encontro e distanciamento entre as relações *Contrast* e *Concession*?

É válido salientar que a RST foi a teoria escolhida para o desenvolvimento desta investigação por dois motivos que se retroalimentam, a saber: (i) o caráter sistemático e

formal da teoria em nível discursivo, cujo produto contribui com os estudos descritivos em si e com variadas aplicações na área de Processamento de Língua Natural (PLN), tais como tradução automática [Guzmán *et al.* 2014], sumarização automática [Li; Thadani; Stent, 2016] e analisadores discursivos [Kobayashi *et al.*, 2020]; e (ii) o empenho pela continuidade e avanços dos estudos em RST, com recursos disponíveis para o português brasileiro (PB), em diálogo com os atuais objetivos da área, na chamada eRST (*Enhanced Rhetorical Structure Theory*), com a identificação e descrição não só das relações de coerência, mas também dos elementos que as sinalizam (os SDs), conforme propõem Das e Taboada (2018) e Zeldes *et al.* (2024) para o inglês; e Cardoso *et al.* (2024) e Dantas *et al.* (2024) para o PB.

Este trabalho se organiza da seguinte maneira: na Seção 2, apresentamos os procedimentos metodológicos para a coleta e anotação das proposições com efeitos de oposição a serem estudadas; em 3, descrevemos brevemente os resultados encontrados; seguido das Considerações Finais e propostas de investigações futuras.

2. A respeito das relações *Contrast* e *Concession* no *corpus* CSTNews

Para investigar o comportamento das relações *Contrast* e *Concession*, foi utilizado o CSTNews [Cardoso *et al.*, 2011], um *corpus* de textos jornalísticos do PB que, entre outras, também está anotado com as relações retóricas previstas pela RST. O CSTNews é composto por 50 conjuntos (*clusters*) de textos, sendo que cada um deles possui 2 ou 3 notícias que abordam o mesmo tema. Recentemente, o *corpus* passou por um processo de *anotação sobre anotação*, ou seja, parte do *corpus* (sendo 1 documento de cada *cluster*) foi retomado e anotado agora com os SDs que indicam essas relações. Esse novo processo foi feito, utilizando a ferramenta *rstWeb* [Gessler, Liu, Zeldes, 2019], por um grupo de pesquisadores que tomaram como base para a tarefa o Manual de Anotação de Dantas *et al.* (2024).

O Manual de Anotação de Dantas *et al.* (2024) propõe uma tipologia com 5 categorias para os SDs e seus subtipos, a saber: *marcadores discursivos* (preposição, conjunção, advérbio), *morfológicos* (tempo verbal, sentido verbal), *sintáticos* (oração circunstancial, oração relativa, pronome relativo, numeral), *semânticos* (hiperônimo-hipônimo, sinonímia, antônímia, campo semântico, conhecimento de mundo, fonte, verbo de comunicação, acrônimo) e *gráficos* (parênteses, travessão, pontuação, aspas), evidenciando que a abrangência deles vai além dos MDs prototípicos, dos quais as conjunções fazem parte.

Seguindo a metodologia utilizada por Dantas *et al.* (2024), no presente trabalho, buscou-se no *corpus* as relações com efeito de oposição aqui elencadas (*Contrast* ou *Concession*) que aconteciam numa mesma sentença, ou seja, de forma intrassentencial. A partir desse critério, foram encontradas 16 ocorrências da relação *Contrast* e 40 da *Concession*. Para a análise dos dados, os trechos das ocorrências foram registrados em planilhas, juntamente com os SDs encontrados nelas. É importante mencionar que nem todos os anotadores apontaram os mesmos SDs nos fragmentos selecionados. Esses casos de discrepância foram debatidos e novamente anotados, segundo a compreensão atual do manual proposto [Dantas *et al.*, 2024].

3. Resultados e discussões

Salientamos que os dados aqui abordados descrevem o fenômeno considerando somente o recorte da nossa pesquisa. Ou seja, obviamente, poderá haver outras construções com efeitos de oposição que se distanciam das apresentadas aqui.

Na RST os dados são organizados em forma de árvore, como na Figura 1¹. A relação entre proposições ou EDUs (*Elementary Discourse Unit*), que são as unidades mínimas de sentido do texto, evidenciadas pelas numerações. No caso da Figura 1, tem-se as EDUs 12 e 13, em que a seta sai do satélite (EDU 12) e vai em direção ao que foi pré-anotado como núcleo (EDU 13).

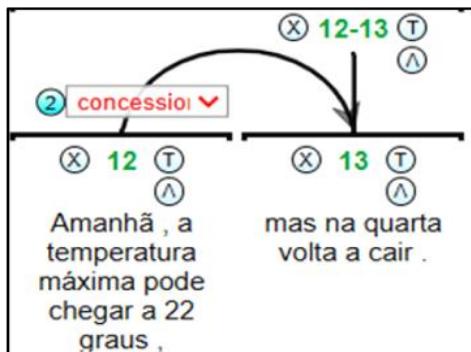


Figura 1. Relação de Concession presente no documento D3_C4_OGlobo²

Na *Concession*, há uma incompatibilidade semântica aparente entre N e S, numa definição da RST. Buscando o conceito gramatical, a oração subordinada concessiva admite um fato contrário à oração principal, mas que não é suficiente para impedi-la [Bechara, 2009 [1928]; Cunha e Cintra, 2017 [1985]; Lima, 2011]. No exemplo ilustrado na Figura 1, o fato de *amanhã a temperatura máxima poder chegar a 22 graus* não impede que *na quarta ela volte a cair*. Há uma incompatibilidade entre as proposições, em que uma informa a temperatura máxima e a outra aponta uma queda dessa temperatura. Com isso, o leitor tende a aceitar melhor a informação de N, pois ela acontecerá independente da informação contrária de S.

Na anotação dos sinalizadores, verificou-se que, em alguns casos, um único sinalizador é capaz de indicar o fenômeno da oposição (*sinalizador único*); e, em outros – que representa a grande maioria dos dados analisados – há mais de um sinalizador que se combina para indicar o efeito de oposição da relação (*sinalizadores combinados*). Na Figura 1, foram anotados os seguintes sinalizadores: *chegar aos 22 graus* e *cair* (semântico – antônima) + *mas* (conjunção - MD) + *vírgula* (pontuação - gráfico).

¹ Os exemplos apresentados nesta seção foram retirados do *corpus* CSTNews.

² A identificação das EDUs analisadas, retiradas do *corpus* CSTNews, se dá por meio de códigos. Conforme se apresenta na Figura 1, tem-se: D3 - terceiro documento do *cluster*; C4 - quarto *cluster* do *corpus*; OGlobo - fonte jornalística.

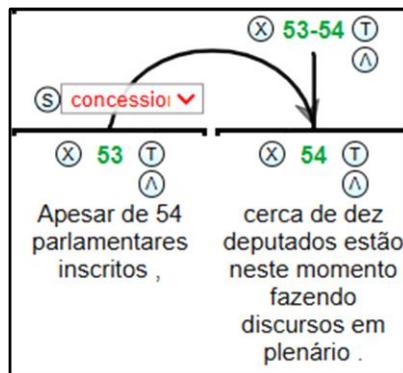


Figura 2. Relação de *Concession* presente no documento D5_C20_GPovo

Na Figura 2, é perceptível a incompatibilidade entre as proposições: o S apresenta uma informação que não impede o acontecimento descrito em N, ele ocorre apesar de S. O MD *apesar de*, que é uma locução prepositiva, deixa esse fato evidente. Além disso, há uma contrariedade entre os números (dados) apresentados; enquanto 54 parlamentares estão inscritos para votação, apenas 10 estão fazendo seus discursos em plenário. A vírgula também tem um papel opositor nessa construção. Ao comparar as informações de S e N, o escritor intenciona que o leitor foque na informação principal, que é a quantidade real de deputados discursando no plenário.

Por meio dos resultados obtidos na análise dos dados desta pesquisa, quanto à relação *Concession*, percebeu-se que os marcadores discursivos (MDs) aparecem em maior quantidade, principalmente as conjunções. Dentre as 37 ocorrências de conjunções, 26 delas são o conectivo *mas*. Além disso, dentre as 11 ocorrências de advérbios, 8 são do advérbio negativo *não*. O segundo SD mais produtivo na relação foi do tipo Semântico, variando entre os seus subtipos.

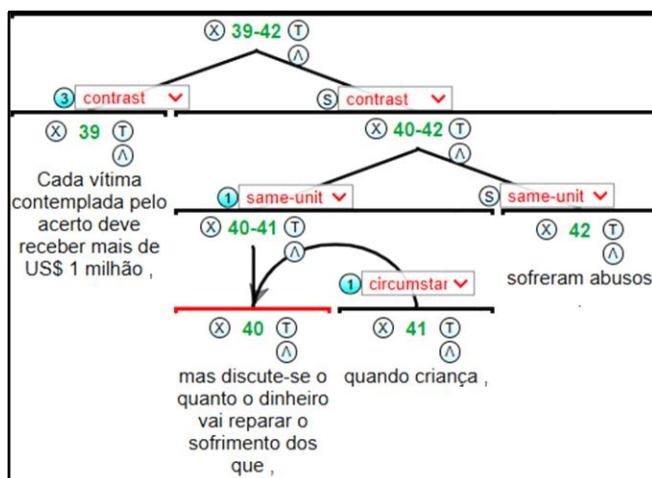


Figura 3. Relação de *Contrast* presente no documento D1_C29_Folha

Na relação *Contrast*, há dois núcleos tidos como igualmente importantes e independentes, que contrastam entre si, ou seja, se trata de uma relação multinuclear. As EDUs 39 e 40-42 estão em oposição. O texto apresenta um acordo feito entre a Igreja Católica e as vítimas de abuso sexual por parte dos padres. Na EDU 39, é dito o valor recebido por cada vítima, enquanto nas EDUs 40-42, se é contraposto como o dinheiro recebido reparará o sofrimento dessas vítimas, ou seja, o escritor não quer que o leitor

aceite melhor uma informação ou outra, mas que as compare e as aceite apenas como informações. De forma prototípica, tem-se o MD *mas* combinado com a vírgula, apresentando uma oposição entre as duas informações. Além dele, há um sinalizador Semântico, do subtipo campo semântico, apresentado em *US\$ 1 milhão /dinheiro*.

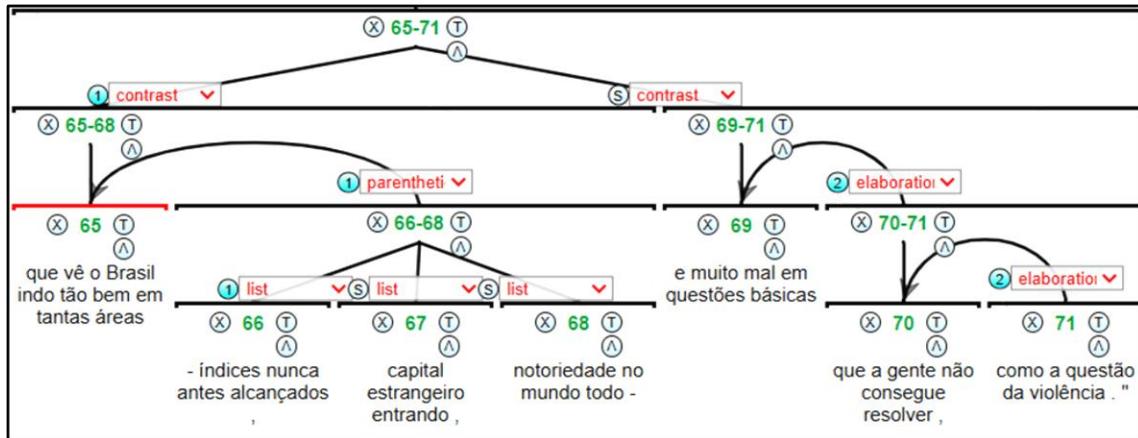


Figura 4. Relação de *Contrast* presente no documento D3_C45_OGlobo

Na Figura 4, a relação *Contrast* está estabelecida entre as EDUs 65-68 e 69-71, identificada, segundo a anotação realizada e a nossa revisão, pelo SD de tipo Semântico, mais especificamente devido à antonímia entre *tão bem* e *muito mal*: a oposição de ideias apresenta, de um lado, o crescimento do país em várias áreas, e do outro os seus problemas em questões básicas. Mais uma vez, não se pretende aceitar mais uma informação ou outra, mas reconhecer que as duas têm a mesma importância para o entendimento geral do que se está sendo dito.

A relação *Contrast* apresenta diferentes tipos de sinalizadores, de acordo com a tipologia de Dantas *et al.* (2024). Os dados mostraram que existe um grande quantitativo de MDs: das 15 ocorrências de MDs, 13 casos foram anotados com a conjunção *mas* e os outros dois foram os advérbios *não* e *nunca*. Depois dos MDs, o tipo Semântico também foi o mais recorrente, apresentando-se, principalmente, por meio da antonímia.

4. Considerações finais

Como visto, as relações de oposição *Contrast* e *Concession* compartilham um mesmo fenômeno, no entanto, possuem características próprias que as definem. A RST apresenta o caráter estrutural (mononuclear ou multinuclear) e o efeito pretendido no leitor/ouvinte como fatores cruciais que as caracterizam. Na *Contrast*, as oposições são identificadas, comparadas e aceitas. Por sua vez, na *Concession*, as oposições são identificadas, mas o efeito pretendido é o de convencer o leitor/ouvinte, fazendo-o aceitar melhor o conteúdo disposto no núcleo.

Considerando-se o número desbalanceado de dados analisados (40 relações *Concession* e 16 relações *Contrast*), é possível afirmar, ainda assim, que o SD de tipo MD é o mais abundante nas relações de oposição, sobretudo a conjunção *mas*. A partir dos dados, é importante ressaltar que os SDs não aparecem sozinhos, mas se relacionam de forma a evidenciar o tipo de relação de sentido existente nas partes de um texto, atuando, portanto, como *sinalizadores combinados*. Dentre os tipos e subtipos que mais apareceram nos dados estão os MDs (conjunções), seguido do SD do tipo gráfico (vírgula) e do semântico (antonímia). Em todos os casos anotados, o *mas* apareceu antecedido do

sinal gráfico *vírgula*, o que caracteriza o comportamento dessas relações de oposição como sendo majoritariamente identificados por, ao menos, dois sinalizadores combinados.

Conforme apontado, nas relações *Concession*, houve uma variedade maior de MDs para além do uso da conjunção *mas*, como casos com *embora*, *mesmo com*, *apesar de*, *não*, *porém*, *dificilmente*, assim como conectivos que apareceram exclusivamente nessas relações (*apesar de*, *embora*, *mesmo com*); já nas relações *Contrast*, houve apenas 2 MDs além do *mas*, sendo os advérbios *não* e *nunca*.

A partir deste trabalho foi possível: (i) estudar e analisar as características das relações de coerência com efeito de oposição do PB, anotando seus pontos de encontro (semelhança estrutural e compartilhamento de SD, principalmente a conjunção *mas*) e de distanciamento (efeitos de sentido e SD exclusivo da relação *Concession*); (ii) estudar, aplicar e propor alterações na taxonomia dos SDs proposta por Dantas *et al.* (2024)³; e (iii) contribuir com a anotação do *corpus* CSTNews com a atuação como adjudicadores das relações *Concession* e *Contrast*.

Para trabalhos futuros, pretende-se ampliar o escopo das anotações e análises dos SDs não apenas em uma mesma sentença, mas em sentenças diferentes, a fim de investigar se os sinalizadores sofrem alterações, como em suas combinações, por exemplo, ou mesmo se sinalizadores diferentes aparecem. Ademais, com relação a possíveis aplicações em PLN, este trabalho subsidia (i) a criação e/ou aprimoramento de sistemas de análise automática de concordância entre anotadores humanos e/ou automatizados; e (ii) a construção e/ou aprimoramento de ferramentas e/ou sistemas que automatizem o processo de análise discursiva a partir de características linguísticas presentes nos enunciados.

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. Além disso, agradecemos à Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) pelo financiamento e suporte.

Referências

- Bechara, E. (2009), *Moderna gramática portuguesa*, Rio de Janeiro: Nova Fronteira.
- Cardoso, P. C. F., Maziero, E. G., Jorge, M. L. R. C., Seno, E. M. R., Felippo, A., Rino, I. H. M., Nunes, M. G. V. e Pardo, T. A. S. (2011) “CSTNews - A Discourse-Annotated

³ Como não é o objetivo deste trabalho, apontam-se aqui apenas alguns exemplos das sugestões propostas de alteração do manual de anotação de Dantas *et al.* (2024): a exclusão ou (re)elaboração dos SDs *sentido verbal* e *pronomé relativo*.

- Corpus for Single and Multi-Document Summarization of News Texts in Brazilian Portuguese". Anais do III Workshop A RST e os Estudos do Texto, p. 88-105.
- Cardoso, P.C.F., Souza, J., Rodrigues, R., Dantas, E., Barbara, L. J. S., Pereira, M. A., Gama, N. S., Almeida, T. J. A. e Cruz, G. (2024), "A Linguagem em Foco: Anotação de Sinalizadores Discursivos? em Textos Jornalísticos", In: XV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, Belém. Anais do XV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana. p. 247.
- Cunha, C., e Cintra, L. (2017). Nova gramática do português contemporâneo. Rio de Janeiro: Lexikon. 800 p., recurso digital.
- Dantas, E., Cruz, G.S.B., Santa Bárbara, L.J., Gama, N.S., Pereira, M.A., Almeida, T.J.A., Souza, J. W. C., Rodrigues, R. e Cardoso, P. C. F. (2024) "Manual de anotação de sinalizadores discursivos", Série de Relatórios Técnicos do NILC.
- Das, D; Taboada, M. (2018). RST Signalling Corpus: a corpus of signals of coherence relations. In Lang Resources & Evaluation. v. 52, 149–184. <https://doi.org/10.1007/s10579-017-9383-x>.
- Gessler, L., Liu, Y. J. e Zeldes, A. (2019) "A discourse signal annotation system for RST trees", In: Proceedings of the Workshop on Discourse Relation Parsing and Treebanking. p. 56-61.
- Guzmán, F.; Joty, S.; Màrquez, L.; Nakov, P. (2014) Using discourse structure improves machine translation evaluation. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore/USA. p. 687-698. DOI: <https://doi.org/10.3115/v1/P14-1065>.
- Kobayashi, N.; Hirao, T.; Kamigaito, H.; Okumura, M.; Nagata, M. (2020) Top-down RST parsing utilizing granularity levels in documents. In: Proceedings of the AAAI Conference on Artificial Intelligence. [s.l.]. p. 8099-8106. DOI: <https://doi.org/10.1609/aaai.v34i05.6321>
- Li., J.J.; Thadani, K.; Stent, A. (2016) The role of discourse units in near-extractive summarization. In: Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue. Association for Computational Linguistics: Los Angeles/USA. p.137-147. DOI: <https://doi.org/10.18653/v1/W16-3617>
- Lima, R. (2011), Gramática normativa da língua portuguesa. Rio de janeiro: José Olympio.
- Mann, W. C. e; Thompson, S. A. (1987) "Rhetorical structure theory: A theory of text organization", Los Angeles: University of Southern California, Information Sciences Institute.
- Mann, W. C. e Taboada, M. (2010) "RSTWeb", <https://www.sfu.ca/rst/03research/index.html>. Outubro.
- Margarido, R. (2010) "Construções (coordenadas) adversativas e construções (subordinadas) adverbiais concessivas em português: pontos de contato e contraste na língua em função", <https://dspace.mackenzie.br/items/ed4ec197-cc80-4f56-8898-04b7cfe95b19>. Janeiro.

- Neves, M. H. M. (1999), Gramática do Português Falado São Paulo: Humanitas/FFLCH/USP; Campinas: Editora da Unicamp. V. 7.
- Pardo, T. A. S. (2005) "Métodos para análise discursiva automática", <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-29082005-172336/pt-br.php/>. Maio.
- Zeldes, A., Aoyama, T., Liu, Y. J., Peng, S., Das, D. e Gessler, L. (2024) "eRST: A Signaled Graph Theory of Discourse Relations and Organization". Computational Linguistics, 1–50.