



**UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS**

MARQUIZE SILVA DOS SANTOS

**ESTUDO DAS RELAÇÕES DE COERÊNCIA COM EFEITOS
DE OPOSIÇÃO EM *CORPUS* DE TEXTOS JORNALÍSTICOS
DO PORTUGUÊS BRASILEIRO**

São Cristóvão – SE
2025

**UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS**

MARQUIZE SILVA DOS SANTOS

**ESTUDO DAS RELAÇÕES DE COERÊNCIA COM EFEITOS
DE OPOSIÇÃO EM *CORPUS* DE TEXTOS JORNALÍSTICOS
DO PORTUGUÊS BRASILEIRO**

Dissertação apresentada ao Programa de Pós-Graduação em Letras (PPGL) da Universidade Federal de Sergipe (UFS), como requisito para a aprovação da defesa do mestrado em Letras.

Área de concentração: Estudos Linguísticos

Linha de Pesquisa: Linguagem, Usos e Tecnologias

Orientadora: Profa. Dra. Roana Rodrigues

Coorientador: Prof. Dr. Jackson Wilke da Cruz Souza

São Cristóvão – SE
2025

FICHA CATALOGRÁFICA

MARQUIZE SILVA DOS SANTOS

Dissertação apresentada ao Programa de Pós-Graduação em Letras (PPGL) da Universidade Federal de Sergipe (UFS), como requisito para a aprovação da defesa do mestrado em Letras sob a orientação da Profa. Dra. Roana Rodrigues (UFS) e coorientação do Prof. Dr. Jackson Wilke da Cruz Souza (UFBA).
Área de concentração: Estudos Linguísticos.

Aprovado em: _____ de _____ de 2025.

BANCA EXAMINADORA

Prof^ª. Dr^ª. ROANA RODRIGUES (presidente)
Universidade Federal de Sergipe (UFS)

Prof. Dr. CÉSAR ALEXANDRE NERI SANTOS (examinador interno)
Universidade Federal de Sergipe (UFS)

Prof^ª. Dr^ª. PAULA CHRISTINA FIGUEIRA CARDOSO (examinador externo)
Universidade Federal do Pará (UFPA)

Ao meu SENHOR, que até aqui tem me ajudado.
1 Samuel 7.12

AGRADECIMENTOS

Em mais uma etapa vencida, sou grata ao Senhor que me sustentou em todo o processo, desde o início, quando não tinha esperança de passar na prova, até o final, em meio a tantos percalços vividos durante os anos do mestrado. Em todos os momentos fui sustentada pela boa mão do Senhor.

Sou grata à minha família por todo apoio, e especial à minha mãe Ivanilde, minha tia Silvanete e o meu irmão Thesare, que são meus maiores incentivadores e admiradores.

Agradeço à minha querida orientadora Roana, por confiar e me incentivar tanto a seguir como pesquisadora, que é algo que tenho tanto prazer. Agradeço pela parceria, pela paciência, por compartilhar tanto e me ajudar a crescer academicamente. Sou grata por sua compreensão e acolhimento em momentos difíceis para mim, os quais tive dificuldade em seguir com a pesquisa. Foi extremamente importante para chegar até o final.

Agradeço ao meu coorientador Jackson, que chegou tão gentil, solícito e competente e somou tanto em minha aprendizagem, não só em minha pesquisa, mas em áreas afins também. Obrigada pela paciência em mediar e compartilhar seu conhecimento e por estimular meu raciocínio. Obrigada por ter tornado leve o processo de coorientação, por ter me acolhido e visto potencial onde não percebi.

Agradeço à banca, Paula e César, por ter aceitado o convite tão sollicitamente, por ter feito uma leitura atenta e cuidadosa do texto e pelos apontamentos enriquecedores e construtivos para melhoria do meu trabalho.

A todos que até aqui foram instrumentos em minha vida e em minha pesquisa de alguma forma, muito obrigada!

RESUMO

Dentre as relações de coerência que formam e dão sentido a um texto, as relações *Antithesis*, *Contrast* e *Concession* apresentam um fenômeno comum à luz da *Rhetorical Structure Theory*, a oposição (Egg e Das, 2022). Desse modo, o objeto de estudo deste trabalho são as relações de coerência com efeito de oposição em textos jornalísticos do português brasileiro. Esta pesquisa tem por objetivo principal investigar e descrever as relações de oposição no *corpus* CSTNews (Cardoso *et al.* 2011), buscando pontos que as diferenciam e as conectam. Para isso, buscou-se, além de marcadores discursivos, outros elementos linguísticos que identificassem e justificassem a anotação dessas relações no referido *corpus*. A teoria base para esses estudos é a *Rhetorical Structure Theory* (Mann e Thompson, 1987), ou RST, que se dedica ao estudo das relações de coerência existentes no texto, por meio de suas proposições. Inicialmente, o principal objetivo da RST era auxiliar na geração de textos automáticos que fossem coerentes (Taboada e Mann, 2006), no entanto, ela tem sido usada para outros fins computacionais e para diferentes aplicações linguísticas. O foco nas relações de oposição permite que a teoria seja investigada em partes de forma mais aprofundada, expandindo os estudos e as aplicações da RST no processamento de linguagem natural. Para alcançar os objetivos desta pesquisa, foi utilizado o CSTNews (Cardoso *et al.* 2011), que é um *corpus* multidocumental que contém diversas anotações (manuais e semiautomáticas) linguísticas, incluindo a RST. Além da anotação das relações retóricas, o *corpus* passou por uma reanotação, agora dos sinalizadores discursivos que indicam as relações retóricas, baseada no Manual de anotação de Dantas *et al.* (2024). No *corpus*, buscou-se pelas relações de oposição (*Antithesis*, *Contrast* e *Concession*) e delimitou-se o *subcorpus* de análise somente de relações que aconteciam dentro de uma mesma sentença. A análise dos dados mostrou a necessidade de revisão na anotação dos sinalizadores feita por Dantas *et al.* (2024). Além disso, mostram que as relações de oposição podem ser identificadas pelos mesmos sinalizadores, por meio de combinações que se relacionam para designar o sentido da relação, sendo a mais recorrente Marcadores discursivos, principalmente o *mas*, Gráfico, a *vírgula*, e Semântico. Os resultados desta pesquisa contribuem para a expansão dos estudos da RST, bem como a sua aplicação em sistemas computacionais no Processamento de Linguagem Natural.

Palavras-chave: Relações de coerência; Oposição; Contraste; *Rhetorical Structure Theory*; Anotação de *corpus*.

RESUMEN

Entre las relaciones de coherencia que forman y dan sentido a un texto, las relaciones de *Antithesis*, *Contrast* y *Concession* presentan un fenómeno común a la luz de la *Rhetorical Structure Theory*: la oposición (Egg y Das, 2022). El objeto de este estudio es, por lo tanto, las relaciones de coherencia con efecto de oposición en textos periodísticos del portugués brasileño. El objetivo principal de esta investigación es investigar y describir las relaciones de oposición en el *corpus* CSTNews (Cardoso et al. 2011), buscando puntos que las diferencien y las conecten. Para ello, además de los marcadores del discurso, buscamos otros elementos lingüísticos que pudieran identificar y justificar la anotación de estas relaciones en el *corpus*. La teoría básica para estos estudios es la *Rhetorical Structure Theory* (Mann y Thompson, 1987), o RST, que se dedica al estudio de las relaciones de coherencia en el texto, a través de sus proposiciones. Inicialmente, el objetivo principal de la RST era ayudar a generar textos automáticos que fueran coherentes (Taboada y Mann, 2006), sin embargo, se ha utilizado con otros fines computacionales y para diferentes aplicaciones lingüísticas. El enfoque en las relaciones de oposición permite profundizar la teoría por sus partes, ampliando los estudios y aplicaciones de RST en el procesamiento del lenguaje natural. Para alcanzar los objetivos de esta investigación, utilizamos CSTNews (Cardoso et al. 2011), que es un *corpus* multidocumento que contiene diversas anotaciones lingüísticas (manuales y semiautomáticas), incluyendo RST. Además de la anotación de las relaciones retóricas, el *corpus* se sometió a una reanotación, ahora de las señales discursivas que indican relaciones retóricas, basada en el Manual de Anotación de Dantas et al. (2024). Se buscaron en el *corpus* las relaciones de oposición (*Antithesis*, *Contrast* y *Concession*) y se delimitó el *subcorpus* para analizar sólo las relaciones que se producen dentro de una misma oración. El análisis de los datos mostró la necesidad de revisar la anotación de señalamientos realizada por Dantas et al. (2024). Además, muestran que las relaciones de oposición pueden ser identificadas por los mismos señalizadores, a través de combinaciones que se relacionan para designar el significado de la relación, siendo los más recurrentes los marcadores de Discurso, principalmente *pero*, Gráfico, la coma, y Semántico. Los resultados de esta investigación contribuyen la ampliación de los estudios sobre RST, así como a su aplicación en sistemas informáticos de Procesamiento del Lenguaje Natural.

Palabras clave: Relaciones de coherencia; Oposición; Contraste; *Rhetorical Structure Theory*; Anotación de *corpus*.

ABSTRACT

Among the coherence relations that form and give meaning to a text, the Antithesis, Contrast and Concession relations present a common phenomenon in the light of Rhetorical Structure Theory: opposition (Egg and Das, 2022). The object of this study is therefore the relations of coherence with the effect of opposition in Brazilian Portuguese journalistic texts. The main objective of this research is to investigate and describe the relations of opposition in the CSTNews corpus (Cardoso et al. 2011), looking for points that differentiate and connect them. To do this, in addition to discourse markers, we looked for other linguistic elements that could identify and justify the annotation of these relations in the corpus. The basic theory for these studies is Rhetorical Structure Theory (Mann and Thompson, 1987), or RST, which is dedicated to the study of coherence relationships in the text, through its propositions. Initially, the main objective of RST was to help generate automatic texts that were coherent (Taboada and Mann, 2006), however, it has been used for other computational purposes and for different linguistic applications. The focus on opposition relations allows the theory to be investigated in more depth in parts, expanding the studies and applications of RST in natural language processing. To achieve the objectives of this research, we used CSTNews (Cardoso et al. 2011), which is a multi-document corpus that contains various (manual and semi-automatic) linguistic annotations, including RST. In addition to the annotation of rhetorical relations, the corpus was re-annotated, now for the discursive flags that indicate rhetorical relations, based on the Annotation Manual by Dantas et al. (2024). In the corpus, we looked for relations of opposition (Antithesis, Contrast and Concession) and delimited the subcorpus of analysis only to relations that occurred within the same sentence. The analysis of the data showed the need to revise the annotation of the signalers made by Dantas et al. (2024). In addition, they show that oppositional relations can be identified by the same signposts, through combinations that are related to designating the meaning of the relation, the most recurrent being Discourse markers, mainly but, Graphic, the comma, and Semantic. The results of this research contribute to the expansion of RST studies, as well as its application in computer systems in Natural Language Processing.

Keywords: Coherence relations; Opposition; Contrast; Rhetorical Structure Theory; Corpus annotation.

LISTA DE FIGURAS

Figura 1 - Diagrama da análise de uma porção de texto	15
Figura 2 - Exemplo da relação Elaboration	27
Figura 3 - Exemplo relação Antithesis	34
Figura 4 - Exemplo da relação Concession	34
Figura 5 - Exemplo da relação Contrast	35
Figura 6 - Relação Elaboration sinalizando o mas em pergunta retórica	37
Figura 7 - Relação Elaboration introduzindo um subtópico	38
Figura 8 - Relação Elaboration na tomada de turno	39
Figura 9 - Exemplo de relação Concession	41
Figura 10 - Exemplo da relação Concession com núcleo invertido	42
Figura 11 - Relações RST no corpus CSTNews	45
Figura 12 - Tipologia dos sinalizadores em Dantas et al. (2024)	47
Figura 13 - Ilustração dos dados de análise na planilha	52
Figura 14 - D3_C12_GPovo - Antithesis	55
Figura 15 - D3_C50_OGlobo - Antithesis	57
Figura 16 - D3_C4_OGlobo - Concession	59
Figura 17 - D5_C20_GPovo - Concession	60
Figura 18 - D2_C42_Estadão - Concession	61
Figura 19 - D1_C35_Folha - Concession	62
Figura 20 - D1_C29_Folha – Contrast	64
Figura 21 - D1_C44_Folha – Contrast	65
Figura 22 - D3_C45_OGlobo - Contrast	66
Figura 23 - Mas na relação Concession	68
Figura 24 - Mas na relação Contrast.	68

LISTA DE QUADROS

Quadro 1 - Construção Concessiva	23
Quadro 2 - Construção Adversativa	23
Quadro 3 - Relação Elaboration	30
Quadro 4 - Definição das relações de oposição segundo a RST	33
Quadro 5 - Resumo do processo metodológico	43
Quadro 6 - Exemplo de comparação entre análises gold e silver	49
Quadro 7 - Sinalizadores únicos da relação Antithesis	54
Quadro 8 - Sinalizadores combinados na relação Antithesis	54
Quadro 9 - Anotação D3_C12_GPovo - Antithesis	56
Quadro 10 - Sinalizadores únicos da relação Concession	58
Quadro 11 - Sinalizadores combinados na relação Concession	58
Quadro 12 - Anotação D2_C42_Estadão - Concession	61
Quadro 13 - Sinalizadores únicos da relação Contrast	63
Quadro 14 - Sinalizadores combinados da relação Contrast	63
Quadro 15 - Anotação D1_C44_Folha – Contrast	65

LISTA DE TABELAS

Tabela 1 - Relações de Oposição nos documentos do corpus	50
Tabela 2 - Planilha com dados da relação Antithesis	76
Tabela 3 - Planilha com os dados da relação Concession	78
Tabela 4 - Planilha com dados da relação Contrast	91
Tabela 5 - Lista de relações RST	97

SIGLAS E ABREVIACÕES

RST: *Rhetorical Structure Theory*.

PIBIC: Programa Institucional de Bolsas de Iniciação Científica.

TCC: Trabalho de Conclusão de Curso.

ENEM: Exame Nacional do Ensino Médio

EDU: *Elementary Discourse Unit*.

MD: Marcador discursivo.

SD: Sinalizador discursivo.

POeTISA: *POrtuguese processing - Towards Syntactic Analysis and parsing*.

PLN: Processamento de Linguagem Natural.

N: Núcleo.

S: Satélite.

NURC: Norma Urbana Culta.

SA: Sumarização Automática.

SUCINTO: *Summarization for clever information access*.

SUMÁRIO

APRESENTAÇÃO.....	12
1. INTRODUÇÃO	14
2. ESTADO DA ARTE.....	18
2.1 Elementos linguísticos de OPOSIÇÃO	18
2.1.1. Efeitos de oposição: a respeito da relação de antítese	18
2.1.2. Efeitos de oposição: a respeito das relações de contraste e concessão	20
2.2 <i>Rhetorical Structure Theory</i> (RST): princípios e bases da teoria.....	26
2.3 Noções discursivas de Oposição via RST	32
2.4 Lições aprendidas	40
3. PROCEDIMENTOS METODOLÓGICOS	43
3.1 A respeito do <i>corpus</i> CSTNews.....	43
3.2 Proposta taxonômica de sinalizadores discursivos para o português	46
3.3 <i>Subcorpus</i> de análise.....	48
4. ANÁLISES E DISCUSSÕES.....	54
4.1 Análise da relação <i>Antithesis</i>	54
4.2 Análise da relação <i>Concession</i>	57
4.3 Análise da relação <i>Contrast</i>	62
4.4 Reflexões e contribuições à anotação dos sinalizadores.....	67
5. CONSIDERAÇÕES FINAIS.....	70
REFERÊNCIAS	73
APÊNDICES	76
ANEXO A	97

APRESENTAÇÃO

Este trabalho se dedica ao estudo das relações de coerência com efeitos de oposição em textos jornalísticos do português brasileiro na perspectiva da RST (*Rhetorical Structure Theory*). A justificativa para a escolha dessa temática está centrada em 4 pontos: (i) minha experiência em trabalhos anteriores; (ii) meu gosto pessoal pela descrição linguística; (iii) minha experiência com correção de redações do tipo Enem para cursinho; (iv) integrar e contribuir para o grupo de pesquisa sobre a RST.

Durante os anos de 2018 a 2019, fiz parte de um projeto no Programa Institucional de Bolsas de Iniciação Científica (PIBIC) intitulado “Estudo bibliográfico relativo à abordagem cognitiva, semiótica e discursiva no ensino da argumentação”, sob a orientação da Profa. Dra. Isabel Cristina Michelin de Azevedo. Em meu TCC, migrei da argumentação para os estudos linguísticos comparados, estudando o uso das preposições do português brasileiro e do espanhol de forma comparada sob orientação da professora Roana Rodrigues.

A motivação para mudar a área de estudo se deu durante o 5º período da graduação, cursando a disciplina de Língua Espanhola V, ministrada pela professora Roana Rodrigues, na qual tive o primeiro contato com os estudos linguísticos comparados na morfossintaxe. A afinidade pelos estudos linguísticos e gramaticais se uniram aos estudos comparados entre o português e o espanhol, minhas línguas de habilitação na graduação em Letras – Português e Espanhol.

Ao concluir a graduação, trabalhei como *freelancer* em um cursinho preparatório para o Enem como corretora de redações. Nessa época, a RST me foi apresentada superficialmente e a teoria tinha relação com as competências 3 e 4 da redação do Enem, o que uniria meu trabalho e a pesquisa.

Após decidir sobre a RST, o desafio foi conhecer a teoria, até então desconhecida por mim, e delimitar o objeto de estudo. A cada leitura, percebi quão importante é essa teoria e como ela contribui para o avanço não só dos estudos linguísticos, mas também de sua aplicação à computação e geração automática de dados. Com os avanços da Inteligência Artificial, os estudos e aplicações da RST são imprescindíveis.

Com base nas conclusões do trabalho de Rodrigues, Souza e Cardoso (2023), identificou-se que a conjunção *mas*, que é um marcador discursivo característico da relação *Contrast*, foi utilizada em EDUs – unidades mínimas de significado do texto – anotadas com a relação *Concession* e com a *Antithesis*. A partir disso, investigar-se-ia a diferença entre essas três relações com foco não apenas no marcador discursivo *mas*, comum as três relações, mas

em outros tipos de sinalizadores que podem caracterizar o fenômeno da oposição presente nessas relações. Desse modo, o fenômeno da oposição torna-se o objeto de estudo dessa pesquisa, buscando diferenciá-lo nessas três relações.

1. INTRODUÇÃO

Um texto nasce a partir de ideias que são estruturadas gramaticalmente e relacionadas entre si para que haja clareza e significação no que se é transmitido, isto é, para que haja coerência no texto. Para Koch (2005), a coerência é o princípio da interpretabilidade do texto, o que faz com que o texto faça sentido para os seus usuários em uma situação comunicativa. Dessa forma, a autora defende o processo cooperativo existente entre o produtor e o receptor do texto, no qual alguns componentes são acionados (sintáticos, semânticos, pragmáticos e socioculturais). A construção do sentido do texto é formada a partir da junção entre elementos linguísticos, semânticos, pragmáticos e os conhecimentos socioculturais do falante, que o permite compreender o texto em sua totalidade.

Nessa perspectiva, as relações de coerência, também chamadas de relações retóricas ou discursivas (Das; Taboada, 2018) estabelecem o sentido entre as partes de um texto. De acordo com Egg e Das (2022), as relações de coerência *Antithesis*, *Contrast* e *Concession* indicam um mesmo fenômeno, que é o da oposição.

As relações retóricas podem ser identificadas em um texto a partir de sinalizadores discursivos (SD). Dentre os tipos de sinalizadores, há os marcadores discursivos (MD). Para Das e Taboada (2018), os marcadores discursivos são expressões lexicais, que podem fazer parte de diferentes classes gramaticais, como as conjuntivas, adverbiais e frases preposicionais, geralmente chamados de conectivos, pois estão exercendo, de forma explícita, essa função no texto. Já os sinalizadores discursivos, estão para além de marcadores de discurso: eles formam uma variedade de sinais textuais, tais como os recursos léxicos, semânticos, sintáticos, gráficos e de gênero. É importante ressaltar que os SDs sinalizam as relações, mas não restringem o sentido de cada uma delas.

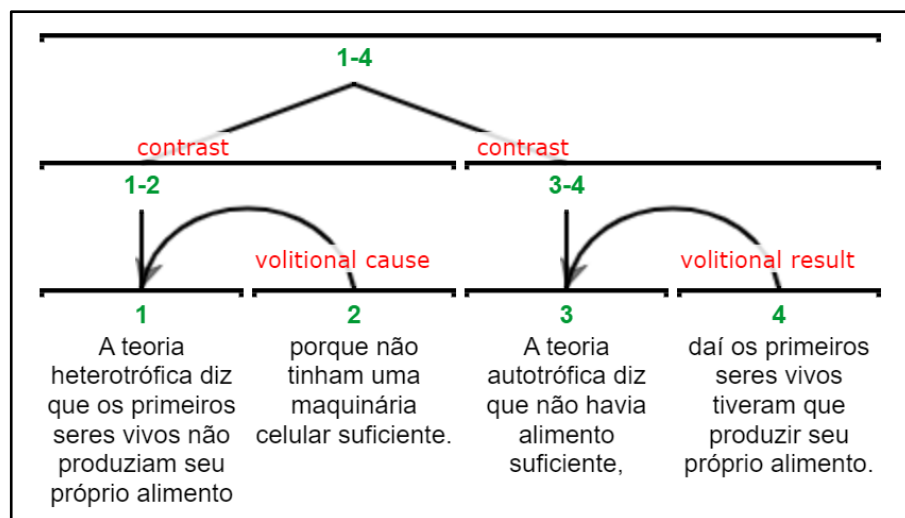
Há marcadores que podem ser usados para indicar mais de uma relação, como é o caso da conjunção *mas*, que pode marcar as relações de oposição, seja *Antithesis*, *Contrast*, ou *Concession*, como se pode ver em (1), (2) e (3), exemplos retirados de Pardo (2005, p. 144, p. 157 e p. 190, respectivamente).

- (1) Os modelos são fechados, eles indicam o que deve ser feito, *mas* não dizem como as atividades que proporcionam a melhoria devem ser realizadas.

- (2) O usuário tem a impressão de que tudo está acontecendo no seu equipamento, *mas* na realidade ele está compartilhando discos, impressoras, processadores e tudo mais que for necessário e o sistema permitir.
- (3) Obter esse resultado corresponde a introduzir estratégias pedagógicas nas aplicações hipermídias, *mas* essa tarefa é difícil porque os sistemas de autoria hipermídia tradicionais (HyperCard e ToolBook, por exemplo) possuem apenas recursos genéricos para auxiliar o autor no desenvolvimento de aplicações.

Nos três exemplos, o *mas* está exercendo uma relação de oposição de ideias, só que em relações retóricas diferentes. Nota-se que as relações não podem ser analisadas considerando apenas esse fator, pois ele é ambíguo e pode desempenhar diferentes funções em diferentes relações. Se essas relações não podem ser caracterizadas por meio, apenas, de MDs, outros sinalizadores devem ser considerados. Além disso, nem sempre as proposições se apresentarão com um MD, mas, ainda assim, é possível identificar a relação de coerência existente por meio de outros SDs, como exemplificado na Figura 1.

Figura 1 - Diagrama da análise de uma porção de texto



Fonte: Baseado em Antonio (2017, p. 88)

Na RST, a análise do texto é feita e representada por meio de um sistema arbóreo, como ilustra a Figura 1, que será mais bem explicado na seção 2.3. A relação *Contrast*, nesse exemplo, está sinalizada por uma oposição entre os prefixos *hetero* e *auto*, das palavras heterotrófica e autotrófica, o que deixa evidente que uma teoria está se opondo à outra. Nesse

caso, então, não foi necessário um MD para que o contraste fosse identificado, mas, por meio do recurso morfológico, o tipo de relação foi identificada.

A fim de estudar o fenômeno da oposição nas relações retóricas, a *Rhetorical Structure Theory*, ou RST, será a teoria base dessa pesquisa, uma vez que ela se dedica ao estudo do texto em suas unidades mínimas de sentido – conhecidas como EDU, *Elementary Discourse Unit* –, analisando como a organização dessas partes foram construídas para que haja coerência em todo o texto. A escolha de trabalhar com esse modelo teórico se dá pelo fato de ele apresentar uma visão diferente de outros modelos linguísticos acerca da organização do texto e um estudo mais completo do discurso (Taboada; Mann, 2006).

A ideia inicial para esta pesquisa se deu por meio do trabalho de Rodrigues, Souza e Cardoso (2023). Os autores analisam os sinalizadores das relações retóricas do *corpus* CSTNews (Cardoso *et al.*, 2011), pré-anotado com o modelo RST. Baseados em autores como Antonio (2017), Das e Taboada (2018) e Liu e Zeldes (2019), os autores justificam seu estudo pela necessidade de criação de tipologias de SDs para além dos marcadores para o português brasileiro (PB), uma vez que as relações retóricas podem ser interpretadas por serem entidades cognitivas, uma linha de pesquisa que não aborda as relações não sinalizadas.

Nessa perspectiva de estudar os SDs e não apenas os marcadores, delimitou-se o fenômeno da oposição, com suas relações correspondentes no modelo RST, como objeto de estudo deste trabalho. Apesar de terem uma característica comum, a oposição, as relações que fazem parte desse fenômeno são diferentes e possuem características próprias. No entanto, como visto nos exemplos (1) a (3), por vezes, essas relações podem ser sinalizadas pelo mesmo marcador.

Diante do que foi exposto, ao longo deste trabalho, buscou-se discutir as seguintes perguntas:

- Fazendo parte do mesmo fenômeno de oposição, o que caracteriza as relações *Antithesis*, *Contrast* e *Concession*?
- Sendo possível indicar mais de uma relação, o que caracteriza a conjunção *mas*?
- Quais outros sinalizadores são usados para marcar essas relações?

Esta pesquisa teve como objetivo principal caracterizar as relações que compreendem o fenômeno da oposição no *corpus* CSTNews (Cardoso *et al.* 2011). Investigou-se para além de marcadores discursivos, sinalizadores linguísticos e estruturais que justifiquem a anotação

dessas relações no referido *corpus*. Como resultado prático, visou-se também apresentar melhorias ao *corpus*, por meio dos estudos feitos nesta pesquisa.

Por fim, por meio do projeto “RST para além dos marcadores discursivos¹”, esta pesquisa tem o objetivo de contribuir com os estudos do POeTISA², que é um projeto de longo prazo que visa aumentar os recursos baseados em análise sintática e desenvolver ferramentas de PLN (Processamento de Linguagem Natural) relacionadas ao português brasileiro. Uma vez que o projeto POeTISA busca produzir um grande e abrangente *corpus* multigênero de classes gramaticais baseados em dependências universais e em textos anotados sintaticamente, principalmente em textos de notícias e conteúdos gerados por usuários, esta pesquisa vai ao seu encontro, trabalhando a anotação, já em nível discursivo, do *corpus* CSTNews.

Este trabalho está dividido em cinco seções a partir da introdução. Na primeira, discorreu-se sobre a teoria que fundamenta esta pesquisa, com subseções que abordam o fenômeno da oposição, a RST e conceito linguísticos e gramáticos. Na segunda, detalhou-se os passos metodológicos utilizados na pesquisa, além de abordar estratégia de análise de concordância. Na terceira, apresentou-se a análise e discussões sobre os dados encontrados. Por fim, na quarta, apresentou-se um resumo sobre o trabalho desenvolvido, suas contribuições e perspectivas de trabalhos futuros.

¹Disponível em: <https://sites.google.com/view/rst-poetisa>. Acesso em 23 de jul. de 2024.

² Poetisa. Disponível em: <https://sites.google.com/icmc.usp.br/poetisa/the-project>. Acesso em 25 de jan. de 2024.

2. ESTADO DA ARTE

Nesta seção, serão apresentados os conceitos-chave acerca do objeto de estudo – a oposição –, da teoria base na qual a pesquisa está ancorada, a RST, e das relações de oposição segundo a RST. A correspondência feita entre as relações RST e a descrição gramatical é a partir de suas proposições, mostrando que, em certa medida, essas teorias (RST e gramática) se encontram e divergem.

2.1 Elementos linguísticos de OPOSIÇÃO

Assim como para Egg e Das (2022) as relações *Antithesis*, *Contrast* e *Concession* abordam o fenômeno da oposição na RST, nas gramáticas tradicionais, tem-se a Antítese, como figura do pensamento, numa perspectiva mais semântica, e as orações Adversativas (contraste) e Concessivas numa perspectiva sintático-funcional.

É tomando como base esses paradigmas que as relações *Antithesis*, *Contrast* e *Concession* serão abordadas, caminhando pelas descrições formais apresentadas pelas gramáticas e funcionalistas do discurso.

2.1.1. Efeitos de oposição: a respeito da relação de antítese

A antítese está entre as figuras do pensamento. Segundo Cegalla (2010), essas figuras são recursos estilísticos que se realizam na esfera do pensamento, no âmbito da frase. A antítese se caracteriza pela aproximação de palavras ou expressões de sentidos opostos.

- (4) "A areia, *alva*, está agora *preta*, de pés que a pisam" (Jorge Amado *apud* Cegalla, 2010, p. 626)
- (5) "Como eram possível *beleza* e *horror*, *vida* e *morte* harmonizarem-se assim no mesmo quadro?" (Érico Veríssimo *apud* Cegalla, 2010, p. 626)

Em (4) e (5), os sentidos contrários estão caracterizados pela antonímia das palavras, *alva* e *preta*, *beleza* e *horror*, *vida* e *morte*, que se opõem totalmente umas às outras. Ao encontro de Cegalla (2010), Lima (2011) diz que a antítese é a contraposição de uma palavra ou frase a outra de significação oposta.

- (6) “*Amigos e inimigos* estão, amiúde, em posições trocadas. Uns nos querem *mal*, e fazem-nos *bem*. Outros nos almejam o bem, e nos trazem o mal” (Rui Barbosa *apud* Lima, 2011, p. 613).

Para Azeredo (2008), antítese é a relação entre duas unidades de significado - palavra, sintagma ou enunciados - que expressam conteúdos opostos.

- (7) “Até agora, estou entre o *elogio*, que me *desvanece*, e a *restrição*, que me *deprime*” (Rodrigues, 1996 *apud* Azeredo, 2008, p. 497).
- (8) “O primeiro efeito dessa lei antifumo, radical e cheia de furos, não foi *apagar* os cigarros, mas *acender* uma grande polêmica” (Ventura, 1999 *apud* Azeredo, 2008, p. 497).

Azeredo (2008) segue a mesma perspectiva dos gramáticos anteriores, mas amplia das palavras e frases de significação oposta para enunciados de conteúdos opostos. Apesar de apresentar exemplos que seguem a antonímia (7) e (8), *enunciados de conteúdos opostos* podem indicar mais que palavras antônimas específicas.

Numa perspectiva semiótica, Fiorin (1988, p. 63) apresenta a Antítese como “uma relação de disjunção sêmica na sucessividade do sintagma”, ou seja, uma separação no sistema de sinais que transmite o pensamento. No entanto, o linguista afirma que só pode se opor elementos semânticos que tiverem uma unidade mínima de significação dentro de um campo semântico comum. Para exemplificar o que diz, Fiorin (1988) apresenta um trecho do livro “Memórias póstuma de Brás Cubas”, de Machado de Assis.

"Gastei trinta dias para ir do Rocio Grande ao coração de Marcela, não já cavalcando o corcel do cego desejo, mas o asno da paciência, a um tempo manhoso e teimoso. Que, em verdade, há dous meios de granjear a vontade das mulheres: o violento, como o touro de Europa, e o insinuativo, como o cisne de Leda e a chuva de ouro de Dânae, três inventos do padre Zeus, que, por estarem fora de moda, aí ficam trocados no cavalo e no asno" (Op.cit, cap. XV, p.534 *apud* Fiorin, 1988, p. 63).

Nesse trecho, há algumas antíteses. Primeiro há uma oposição entre *o corcel do cego desejo* e *o asno da paciência*; em seguida, uma oposição temática entre a conquista violenta e a conquista insinuativa. Por fim, uma oposição comparativa trazida da mitologia grega em *touro de Europa*, *cisne de Leda* e *chuva de ouro de Dânae*. Todas essas oposições fazem parte de um campo semântico comum.

Fiorin (1988) mostra que, por meio de análises semióticas do mundo natural, cria-se uma ilusão de referencialidade, em que o enunciador pretende levar o enunciatário a aceitar o seu discurso.

Emílio (1999) apresenta a antítese aplicada à linguagem jornalística, valorizando a estilística nas contradições do jornalismo existentes nos textos conflitantes. A ambiguidade é o principal ponto encontrado nos textos jornalísticos, segundo a autora. Para ela, esse aspecto é extremamente informativo e proporciona uma gama de possibilidades interpretativas ao leitor, que se sente atraído pela notícia e busca decodificá-la de diversas formas.

(9) Pescadores catarinenses pescaram uma vaca.

Em (9), o título da notícia busca chamar a atenção do leitor por meio da distância lógica entre *peixe* e *vaca* no ato de pescar. No entanto, o objetivo é usar a função paradoxal para fundir dois percursos diferentes em um só.

A antítese é apresentada como figura do pensamento, que usa palavras, expressões e até construções que se opõem no aspecto semântico. Para além disso, Emilio (1999) apresenta a antítese em textos jornalísticos sendo usada como meio de chamar a atenção do leitor ao causar certo estranhamento nele, apontando para a estilística que a antítese proporciona.

2.1.2. Efeitos de oposição: a respeito das relações de contraste e concessão

Embora a RST não esteja pautada na gramática normativa, os conceitos sobre os fragmentos com efeito de oposição podem se relacionar de alguma forma aos estudos gramaticais, especificamente, a relação entre *Contrast* e oração coordenada adversativa e *Concession* e relação subordinada adverbial concessiva, bem como o emprego de seus conectivos, sobretudo as conjunções.

Iniciando por uma vertente mais normativa da língua, na Nova Gramática do Português Contemporâneo, Cunha e Cintra (2017 [1985]) apresentam as orações adversativas como sendo aquelas que são acompanhadas por conjunções adversativas. Os autores defendem que as conjunções adversativas são aquelas que ligam termos ou orações de igual função que apresentam uma ideia de contraste, são elas: *mas, porém, todavia, contudo, no entanto, entretanto*;

(10) Apetece cantar, *mas* ninguém canta.

O mesmo acontece em relação às orações concessivas, os referidos gramáticos as restringem às conjunções concessivas, as quais iniciam uma oração subordinada em que se admite um fato contrário à oração principal, mas que não é suficiente para impedi-la, são algumas delas: *embora, conquanto, ainda que, mesmo que*, entre outras. Os exemplos (10) e (11) foram retirados de Cunha e Cintra (2017 [1985], p. 594 e 600).

(11) Não saberei nunca escrever sobre ele, *embora* tenha tentado mais de uma vez.

Cunha e Cintra (2017 [1985]) defendem que ambas as orações – adversativa e concessiva – apresentam uma ideia de oposição, ainda que tenham diferenças e sejam marcadas por conjunções diferentes.

Na gramática de Bechara (2009 [1928]), não dá definições específicas sobre os referidos tipos de orações, mas o faz com as conjunções. Sobre as conjunções adversativas ele diz que elas enlaçam unidades e marcam a oposição entre elas, sendo as conjunções adversativas por excelência o *mas* e o *porém*. Já as concessivas, iniciam uma oração que expressa que um obstáculo não será impedimento para a realização do que é declarado na oração principal, são as principais conjunções: *ainda que, embora, posto que, se bem que, apesar de que*, etc. O autor não apresenta o fator oposição como sendo comum aos dois tipos de conjunções, como fizeram Cunha e Cintra (2017 [1985]).

Em sua gramática normativa, Lima (2011) cita os tipos de orações coordenadas, dentre elas a adversativa, mas não apresenta a definição delas, vai fazê-lo na parte das conjunções. Já a oração concessiva, para o autor, expressa um fato, real ou suposto, que poderia se opor à realização de outro fato principal, contudo não frustrará o cumprimento deste. As principais conjunções seguem as mesmas indicadas pelos gramáticos anteriores.

Sobre as orações adversativas, na parte das conjunções, Lima (2011) diz que elas relacionam um pensamento contrastante, sendo o *mas* a conjunção contrastiva por excelência. As demais conjunções – *porém, contudo, no entanto, todavia, entretanto* –, segundo o gramático, possuem mais valor de concessiva atenuada (12) e menos valor contrastivo, diferente da conjunção *mas*.

(12) Ele falou bem; *todavia*, não foi como eu esperava.

Em (12), o autor mostra uma aproximação semântica existente entre a adversativa e a concessiva nesse sentido menos contrastivo e mais de concessiva “atenuada”. Como correspondente a essa construção adversativa, ele dá um exemplo concessivo.

(13) *Embora* ele tenha falado bem, não foi como eu esperava.

Nesses exemplos (12) e (13), pode-se perceber que essa aproximação semântica está caracterizada na ordem em que a conjunção aparece na oração. Isso ficará ainda mais evidente mais adiante com Neves (1999), fazendo uma correlação entre essas construções.

Diferente de Bechara (2009 [1928]) e Cunha e Cintra (2017 [1985]), Lima (2011) apresenta mais detalhes sobre essas orações e o uso de suas conjunções, não restringindo sua definição apenas às conjunções prototípicas, mas mostrando em quais momentos essas conjunções podem se misturar entre elas, já que ambas marcam a oposição de ideias. O autor defende que as concessivas podem se apresentar de duas formas, em orações desenvolvidas ou reduzidas. Quando apresentada em forma de oração desenvolvida anteposta à oração principal, a concessiva ganha uma força mais contrastiva, dando evidência à conjunção de oposição. Neste caso, também aparecem conjunções que são tipicamente adversativas, como: *contudo*, *entretanto* e *todavia*, como mostra em (14).

(14) Posto que se tivesse rebelado contra o comandante, / (sempre, todavia) acabou por acatá-lhe as ordens.

A Gramática do Português Falado (Neves, 1999) é uma gramática funcionalista que tem como base de análise a Norma Urbana Culta (NURC), tendo em vista o funcionamento e competência comunicativa dos falantes. A obra é resultado de discussões e seminários de pesquisas de diferentes especialistas da Língua Portuguesa. Em um dos capítulos, tratando das concessivas, Neves (1999) faz a diferenciação entre a coordenada adversativa e a subordinada concessiva.

O primeiro ponto que a autora apresenta é que a mescla entre essas orações é corrente na língua, ao tempo em que lista as conjunções adversativas como possibilidades de marcadores das concessivas. Em seguida, Neves (1999) faz uma correlação entre as orações como mostra nos exemplos abaixo e nos Quadros 1 e 2.

- (15) Isso faz com que ... a pessoa ... *embora* ... queira fugir da regra / ... não consiga ... num é?
- (16) Eles querem sempre ... *por mais que* a gente dê, eles querem sempre mais coisas, né?
- (17) Eles fazem um molho com pimenta muito gostoso ... *se bem que* é muito ... que é muito forte ... né ... a gente sente assim aquele gosto muito picante ...

Quadro 1 - Construção Concessiva

Oração concessiva	Oração nuclear
embora queira fugir da regra	a pessoa não consegue
por mais que a gente dê	eles sempre querem mais
se bem que é muito forte	eles fazem um molho muito gostoso

Fonte: Neves (1999, p. 563)

Quadro 2 - Construção Adversativa

Primeira coordenada	Coordenada adversativa
(a pessoa) quer fugir da regra	mas não consegue
a gente dá muito	mas eles sempre querem mais
(o molho) é muito forte	mas é muito gostoso

Fonte: Neves (1999, p. 563)

A autora explica que a diferença está na base argumentativa da concessiva. Enquanto nesta o falante refuta uma objeção, na adversativa o falante apenas admite um fato. Se o conceito dessas orações for resgatado, tem-se exatamente essa explicação. A oração adversativa simplesmente expressa uma oposição entre proposições, enquanto a concessiva expressa que um obstáculo não será suficiente para impedir um fato apresentado na oração principal; ainda que o falante tente se opor a um fato, essa oposição não será suficiente para impedi-lo. Há, portanto, um esforço argumentativo neste caso. Também é interessante, nos exemplos acima, que as conjunções utilizadas para cada construção são de fato as prototípicas dos tipos de orações abordados, convergindo com os gramáticos que defendem que o *mas* é o marcador adversativo por excelência. Há outros exemplos em que a autora deixa isso bem claro.

- (18) nós temos as reuniões ... muito mais participação, porque, *mesmo que* alguns professores faltem porque tenham outros ... outros afazeres no ambulatório, *mas* sempre tem um bom número de reuniões.
- (19) eu gosto de comer couve-flor ... gosto mui / engraçado que eu gosto muito de chuchu *embora* todo mundo ache chuchu uma coisa assim sem graça ... aguada... *mas* eu gosto ...

Em (18) e (19), há presença tanto da concessiva quanto da adversativa, e fica clara a diferença apresentada na argumentação de uma e na afirmação da outra, confirmando, também, o conceito de subordinação (dependência) e coordenação (independência).

O trabalho de Margarido (2010) apresenta uma perspectiva semântica entre as construções adversativas e concessivas também pela perspectiva descritiva. Segundo a autora, a área da semântica de oposição engloba tanto conjunções adversativas, quanto concessivas. Gouvêa (2001 *apud* Margarido, 2010) defende que os conectivos adversativos podem passar a ideia de concessão, apesar de nem sempre esse valor estar presente nas adversativas.

Vilela e Koch (2001 *apud* Margarido, 2010) inclui conjunções adversativas (*mas, contudo, todavia, porém, não obstante etc.*) em um grupo de oposição, bem como cria um grupo de concessiva opositora para as conjunções concessivas (*embora, apesar de que, ainda que*), dando a ideia de que as concessivas também têm valor de oposição.

Azeredo (2008 *apud* Margarido, 2010), contudo, defende que a conjunção adversativa *mas* pode expressar uma oposição entre dois conteúdos (20) ou uma quebra de expectativa criada pela primeira proposição (21), que é um valor normalmente atribuído à construção concessiva.

- (20) A secretária dele é antipática, *mas* competente.
- (21) O lutador era magrinho, *mas* derrubava todos os seus adversários.

Van Dijk (1989 *apud* Margarido, 2010) coloca construções adversativas e concessivas no grupo junto das contrastivas, podendo indicar: cursos de eventos excepcionais (22) e (23); eventos inesperados ou indesejados (24) e (25); não satisfação de condições possíveis, prováveis ou necessárias (26).

- (22) João é muito prestativo, *mas* fez um trabalho ruim na pintura de sua casa.

(23) *Embora* tenhamos dormido tarde, conseguimos pegar o barco.

(24) Eu fui pescar, *mas* não consegui pegar nada.

(25) *Embora* Peter seja muito esperto, não é muito gentil.

(26) Peter quer comprar um carro, *mas* não tem dinheiro.

Halliday e Hasan (1997 *apud* Margarido, 2010) fazem o mesmo, incluem as construções adversativas (27) e concessivas (28) em um grupo chamado *relação adversativa*; porém, o sentido que é empregado a esse grupo é o de *contrariedade de expectativa*, resultante da comunicação entre falante e ouvinte.

(27) Todo esse tempo, Tweedledee fez o possível para fechar o guarda-chuva com ele mesmo dentro... *mas* não teve êxito, e acabou virando e se enrolando todo no guarda-chuva, ficando somente a cabeça para fora.

(28) *Embora* tenha dado o melhor de si, ela fracassou.

Margarido (2010) pontua que os referidos autores, apesar de postularem o valor de *contrariedade de expectativa* aos dois tipos de construções, os linguistas incluem apenas as conjunções adversativas (*mas* e *no entanto*) ao grupo de relação contrastiva. Nesse ponto também é interessante o fato de que Azeredo (2001 *apud* Margarido, 2010) e Halliday e Hasan (1997 *apud* Margarido, 2010) apresentam o mesmo valor para as referidas construções, a *quebra de expectativa* ou *contrariedade de expectativa*, que, segundo Azeredo (2001 *apud* Margarido, 2010), geralmente é atribuída à concessiva, mas ele adiciona o *mas*, assim como Halliday e Hasan (1997 *apud* Margarido, 2010) incluem apenas conjunções adversativas nesse caso. O valor semântico fica ainda mais mesclado entre essas construções.

A autora ainda cita dois linguistas que apresentam uma visão diferente das dos demais, eles fazem a separação entre as adversativas e as concessivas, ainda que essas tenham uma aproximação semântica. Bello (1988 *apud* Margarido, 2010) aponta dois usos para a concessiva *aunque* do espanhol e mostra que em (29) a conjunção ganha valor adversativo de oposição. Essa transformação fica clara não pelo conectivo ou pela posição que foi colocado, mas pela pontuação.

(29) Tenho de sair *embora* chova.

(30) Escreve bem, *embora* devagar.

O mesmo faz König (1985 *apud* Margarido, 2010) com a concessiva *although* do inglês, passa a ter um valor adversativo e, desse modo, o linguista faz uma correlação com uma adversativa.

(31) *Embora* nós não tenhamos sido derrotados, não vencemos também.

(32) Nós não fomos derrotados, *mas* não vencemos também.

O referido autor ainda considera que nem todas as línguas têm conectivos concessivos, mas todas têm conectivos adversativos. Desse modo, ele conclui que as concessivas são uma variação das adversativas, o mesmo é defendido por García (2000 *apud* Margarido, 2010).

Em suma, considerando as proposições das relações retóricas e fazendo uma correlação com as orações adversativas e concessivas, é claro o fato de que essas orações, ainda que tenham o fator contrastivo em comum, são bem definidas e possuem suas próprias conjunções, segundo a norma. Entretanto, na funcionalidade da língua, em seus estudos, Neves (1999) e Margarido (2010) mostram que os conectivos dessas orações são usados tanto em uma como em outra, pois a oposição característica dessas orações faz com que elas sejam mescladas no uso cotidiano.

2.2 *Rhetorical Structure Theory* (RST): princípios e bases da teoria

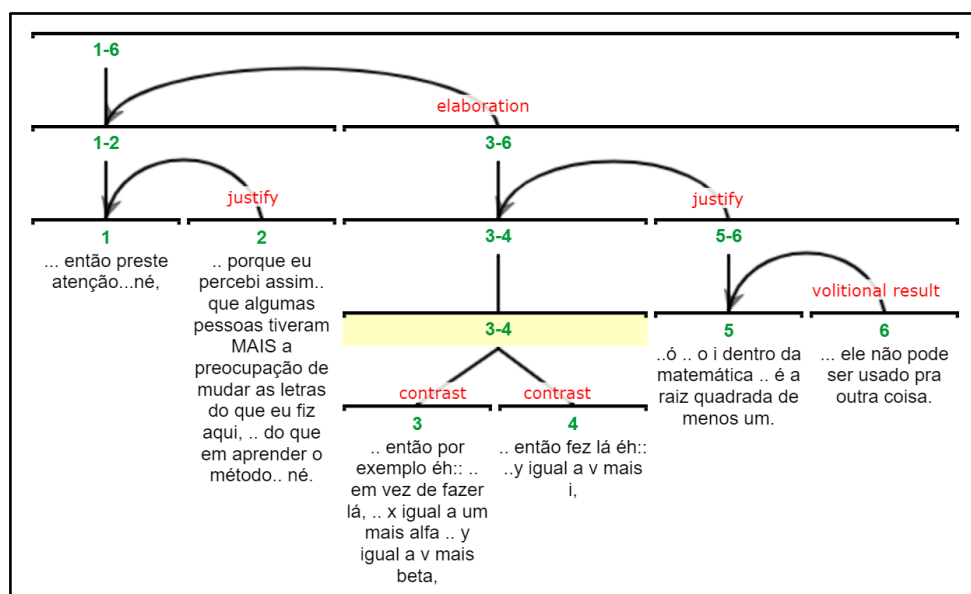
A RST (Mann; Thompson, 1987) é uma teoria descritiva que estuda a organização do texto em suas micro e macroestrutura. Surgiu nos Estados Unidos, na década de 1980, desenvolvida por um grupo de pesquisadores que tinha interesse em PLN. A principal finalidade dela era a geração de textos automáticos; no entanto, tem sido usada para outros fins computacionais e para diferentes aplicações linguísticas, como a tradução automática de texto, corretores ortográficos e gramaticais, ferramentas e sistemas para auxiliar no processo de escrita de textos, entre outros (Taboada; Mann, 2006; Rodrigues; Souza; Cardoso, 2023; Hou; Zhang; Fei, 2020).³

A RST se aplica ao estudo das relações de coerência existentes nas EDUs de um texto, as quais estão conectadas e são representadas em forma de árvore, permitindo uma análise

³ Apesar de sua grande aplicabilidade computacional, a presente dissertação de mestrado foca nos aspectos linguístico-descritivos da teoria e cujo resultado poderá ser utilizado em empreendimentos computacionais em trabalhos acadêmicos futuros.

interpretativa de sua estrutura. Antonio (2017) mostra que as relações podem ser estabelecidas na microestrutura do texto, que se trata de orações e/ou entonação (na língua falada), e na macroestrutura, que se trata de porções maiores de textos, como períodos. O trabalho do autor foi realizado com base nos dados de um *corpus* oral; no entanto, para as análises, ele retextualizou as transcrições, retirando marcas de oralidade. Na análise dessas relações, são descritos aspectos gramaticais e lexicais. Na Figura 2, tem-se um exemplo apresentado pelo autor para ilustrar a análise das relações, por meio do sistema arbóreo, quando há ou não um marcador explícito.

Figura 2 - Exemplo da relação *Elaboration*



Fonte: Baseado em Antonio (2017, p. 84).

No exemplo da Figura 2, na EDU 2, tem-se *porque* como marcador explícito para a justificação, assim como *em vez de* marca o contraste entre informações. Já na relação de *Result*, não há um marcador explícito, a sequência das informações apresentadas já dá indícios para o entendimento de que é o resultado do que está sendo exposto sem que haja um marcador específico.

Ainda na Figura 2, há relações que estão acontecendo dentro de uma mesma sentença (intrassentencial) e outras que ocorrem em sentenças diferentes (interssentenciais). As relações *Justify* – nas duas ocorrências – e *Contrast* estão acontecendo numa mesma sentença. Já as relações *Elaboration* e *Volitional Result* estão acontecendo em sentenças diferentes.

Para a segmentação dos textos, criação das árvores e anotação da RST, como mostra a Figura 2, tem-se a ferramenta RSTTool (O'Donnel, 2000). Para a anotação dos sinalizadores,

como será visto adiante, usou-se a ferramenta rstWeb (Zeldes, 2016; Gessler; Liu; Zeldes, 2019), uma versão mais atualizada.

As relações podem ser classificadas quanto a sua natureza, intencional/argumentativa ou semântica/informativa. Em seu trabalho, Pardo (2005) as diferencia, de modo que as relações intencionais são aquelas em que as intenções do escritor são satisfeitas por meio do reconhecimento e convencimento do leitor. Já as relações semânticas estão desvinculadas das intenções do autor e focam na forma como os conteúdos de um texto se relaciona. No anexo A, há a descrição da natureza de cada relação definida por Pardo (2005).

Segundo Mann e Thompson (1987),

“a Teoria da Estrutura Retórica oferece uma combinação de recursos que se mostrou útil em vários tipos de estudos de discurso. Ela identifica a estrutura hierárquica no texto. Ela descreve as relações entre as partes do texto em termos funcionais, identificando tanto o ponto de transição de uma relação quanto a extensão dos itens relacionados. Fornece análises abrangentes, em vez de comentários seletivos. Não é sensível ao tamanho do texto e foi aplicado a uma ampla variedade de tamanhos de texto” (Mann e Thompson, 1987, p.2. *Tradução nossa*⁴).

Os autores mostram que a teoria não está limitada a um tipo de texto, mas tem seu foco no discurso, por isso o tamanho do texto é irrelevante, uma vez que o formato para a análise da RST é diferente, sendo feita por meio de porções do texto, como se mostra na Figura 2.

Ao falar de níveis linguísticos e o tratamento computacional, Pardo (2005) e, mais recentemente, Caseli, Nunes e Pagano (2023) mostram que os aspectos fonéticos e fonológicos, morfológicos e sintáticos já estão sendo bem desenvolvidos, uma vez que estão mais relacionados à estrutura linguística. No entanto, ao partir para abordagens mais profundas da língua, como a semântica, o discurso e a pragmática, as análises se tornam mais complexas, e os desafios de implementar esse conhecimento para sistemas computacionais se tornam ainda maiores. Isso se deve ao fato da complexidade de resolução e sistematização de fenômenos linguísticos nesses níveis, tanto para humanos quanto para sistemas computacionais.

Para as análises da RST, há um inventário de relações que não é fechado e varia em sua quantidade. Não há uma indicação na teoria sobre o quantitativo máximo das relações; por outro lado, Taboada e Mann (2006) alertam para a dificuldade em anotar e diferenciar as relações

⁴ Do original: “*Rhetorical Structure Theory provides a combination of features that has turned out to be useful in several kinds of discourse studies. It identifies hierarchic structure in text. It describes the relations between text parts in functional terms, identifying both the transition point of a relation and the extent of the items related. It provides comprehensive analyses rather than selective commentary. It is insensitive to text size, and has been applied to a wide variety of sizes of text*” (Mann e Thompson, 1987, p.2).

quando há uma quantidade excessiva delas. Mann e Thompson (1987) apresentaram uma lista com 24 relações retóricas para a língua inglesa; já para o português, Pardo (2005) estabeleceu 32 relações (anexo A), nas quais é possível, a princípio, enquadrar todo tipo de situação comunicativa.

As relações se entrelaçam e formam o texto. Elas podem ser organizadas de duas maneiras, a saber:

- i) *mononuclear* – em que há uma porção do texto que apresentará a informação principal (núcleo – N) e outras que apresentarão informações adicionais ou secundárias (satélite – S) sobre a informação principal;
- ii) *multinuclear* – não há apenas uma informação principal no texto, mas várias informações que são igualmente importantes, sendo cada proposição um núcleo diferente.

É possível observar a diferença entre esses tipos de relações em (33) e (34), exemplos retirados de Pardo (2005, p. 10).

(33) [Embora você não goste,]^S [trabalhar é importante.]^N

(34) [O garoto chegou da escola] ^N [e fez sua lição de casa.]^N [Depois foi brincar com os amigos.]^N

Em (33), tem-se uma relação *Concession*, em que o trecho identificado por N é a informação mais importante, sendo esta proposição nuclear, e o trecho identificado por S, o satélite. Em (34), tem-se uma relação *Sequence*, na qual os dois trechos delimitados por diferentes Ns formam uma sequência de eventos, sendo cada uma das proposições igualmente importantes. Desse modo, a relação *Concession* é mononuclear, e a relação *Sequence* é multinuclear.

Segundo Mann e Taboada (2010), as relações RST, ainda há restrições sobre o N, sobre o S, sobre a combinação núcleo-satélite (N+S) e o efeito. Ao analisar o efeito/intencionalidade do produtor, o analista - responsável por analisar e anotar as relações retóricas - o faz de forma plausível, fazendo um julgamento em que se está considerando o texto, tanto o que está escrito quanto o que dele pode ser interpretado. O conceito de plausibilidade é usado, pois o analista faz seus julgamentos sem conhecer de fato a intenção do escritor/falante ou do leitor/ouvinte.

Quadro 3 - Relação *Elaboration*

Relação RST	Restrições sobre o núcleo ou sobre o satélite individualmente	Restrições sobre núcleo + satélite	Intenção do falante
<i>Elaboration</i>	Nenhuma	O satélite apresenta detalhes adicionais sobre a situação ou sobre algum elemento do assunto que é apresentado no núcleo ou é acessível inferencialmente no núcleo em uma ou mais das maneiras listadas a seguir. Na lista, se o núcleo apresenta o primeiro membro de qualquer par, então o satélite inclui o segundo. <i>grupo: membro</i> <i>abstração: instância</i> <i>todo: parte</i> <i>processo: passo</i> <i>objeto: atributo</i> <i>generalização: específico</i>	O destinatário reconhece o satélite como fornecendo detalhes adicionais para o núcleo. O destinatário identifica o elemento do assunto para o qual o detalhe é fornecido.

Fonte: Traduzido a partir de Mann e Taboada (2010)

Como demonstrada no Quadro 3, *Elaboration* é uma relação que tem a função de acrescentar informações sobre o N por meio do S, objetivando que o destinatário reconheça a sequência que é apresentada, como também é exemplificado em (35), retirado de Antonio (2017, p. 89).

(35) [Quais os motivos das variações entre os indivíduos de uma mesma espécie? São três pontos:]^N [o primeiro ponto importante é a mutação,]^S [o segundo ponto é a recombinação gênica]^S [e o terceiro ponto é a seleção natural.]^S

O fragmento indicado por N está funcionando como núcleo. Em resposta, é apresentada uma lista de informações detalhadas que facilmente é identificada pelo destinatário como acréscimo ao núcleo. Essa listagem está sinalizada por *dois-pontos* que a antecede, e por *enumeração das informações* na sequência.

Para facilitar a observação do analista e a identificação do encadeamento das relações, a análise dos textos é expressa por um sistema arbóreo. Uma vez que as relações tratam desde o nível gramatical ao nível discursivo, a RST recorre aos conceitos da Linguística Funcional para fundamentar suas análises. Antonio (2017) apresenta o modelo de análise *top-down*, o qual parte do texto para as estruturas gramaticais, como apresentado nas Figuras 1 e 2.

No exemplo da Figura 2 (p. 27), relação de *Elaboration*, tem-se nas EDUs 1-2 a informação principal do texto (N), e em 3-6, as informações que acrescentam os detalhes (S). Nota-se que, dentro de uma relação retórica, há outras relações que se unem para garantir a coerência do texto e nem sempre essas relações poderão ser identificadas por marcadores explícitos, como preposições ou conjunções. Nesse caso, serão acionados outros tipos de sinalizadores, como pontuação, léxico, aspectos cognitivos e semânticos, entre outros (Antonio, 2017; Das; Taboada, 2018; Rodrigues; Souza; Cardoso, 2023).

A RST, portanto, tem como objeto de estudo o texto (em seus mais diferentes registros) e as relações de coerência existentes em suas partes. Trata-se de um modelo teórico relevante, com abrangência e escopo em diversas áreas. Segundo Souza, Cardoso e Rodrigues (2023), a RST vem sendo aplicada nos seguintes cenários acadêmico-científicos:

- (i) construção de ontologias: utilizadas para organizar e representar informações, destacando a relação existente entre elas;
- (ii) geração de textos: geração de linguagem natural a partir de um *output* de máquina;
- (iii) sumarização automática: produção automática de uma versão menor coerente e coesa do texto-fonte, a partir de análises discursivas automáticas ou manuais com ênfase nos marcadores discursivos.
- (iv) *parsers* discursivos: desenvolvimento e aprimoramento de *parsers* discursivos baseados na RST;
- (v) análise discursiva: análise automática do discurso de um texto, sendo ele uma estrutura muito elaborada que relaciona todo o seu conteúdo e lhe atribui coerência.
- (vi) tradução automática: uso da RST na tradução automática, construindo árvores RST, fazendo comparações entre estruturas retóricas, entre outros.
- (vii) análise de sentimentos: as relações semânticas da estrutura retórica entre sentenças e parágrafos de um texto podem auxiliar na classificação da polaridade;
- (viii) ferramentas de anotação: para além da criação de *corpus* e analisadores discursivos, o desenvolvimento de interfaces de anotação on-line colaborativos e atualizados.
- (ix) detecção de *fake news*: identificação de *fake news*, com propostas de diferenciação de histórias reais e enganosas, que são objeto de dificuldade de identificação inclusive para humanos.

Uma vez que a RST abrange tantas áreas teóricas e práticas, fica evidente a importância de se ampliar os estudos sobre ela e o desenvolvimento de ferramentas que possibilitem o avanço científico e tecnológico por meio de sua aplicabilidade. Este trabalho, por exemplo,

avança os estudos da RST para além da segmentação das relações retóricas, estudando também os sinalizadores presentes nelas, especialmente os sinalizadores das relações de oposição.

2.3 Noções discursivas de Oposição via RST

Como visto anteriormente, nas relações RST, há restrições sobre o N, sobre o S, sobre a combinação núcleo-satélite (N+S) e o efeito pretendido. Elas também podem ser classificadas em relação de *apresentação* e de *assunto*. Segundo Mann e Taboada (2010), na relação de apresentação, o efeito pretendido pelo escritor/falante é causar uma impressão positiva, uma inclinação e aceitação do núcleo por parte do leitor/ouvinte, como nas relações *Antithesis* e *Concession*; já na relação de assunto, o efeito pretendido pelo escritor é de que o leitor identifique a relação que foi estabelecida, como na relação *Contrast*.

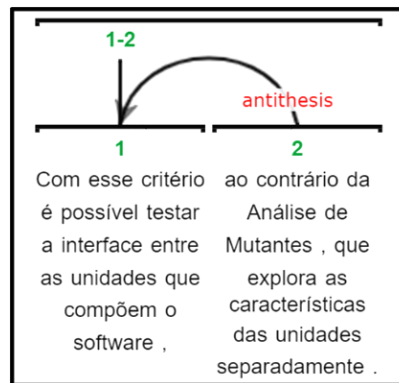
No Quadro 4, estão apresentadas as definições dessas três relações segundo as restrições e efeitos pretendidos.

Quadro 4 - Definição das relações de oposição segundo a RST

RELAÇÃO	RESTRICÇÕES				EFEITO
	sobre N	sobre S	sobre N+S	sobre N + N	
<i>Antithesis</i>	O escritor julga N válido.	Não há.	N e S estão em contraste; por causa da aparente incompatibilidade, não se pode julgar N e S válidos ao mesmo tempo; a compreensão de S e da incompatibilidade entre N e S faz o leitor aceitar melhor N.	Não há.	O leitor aceita melhor N.
<i>Concession</i>	O leitor julga N válido.	O escritor não afirma que S pode ser válido.	O escritor mostra uma incompatibilidade aparente ou em potencial entre N e S; o reconhecimento da compatibilidade entre N e S melhora a aceitação de N pelo leitor.	Não há.	O leitor aceita melhor N.
<i>Contrast</i>	Não há.	Não há.	Não há.	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.	O efeito reconhece as similaridades e diferenças resultantes da comparação sendo feita.

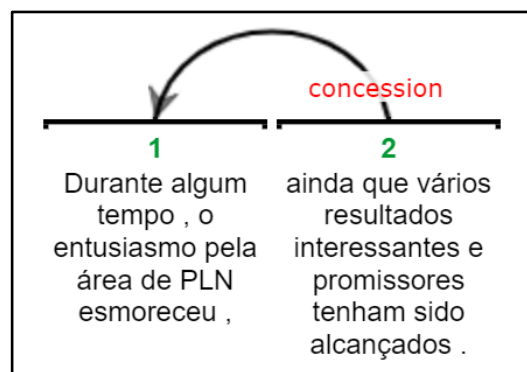
Fonte: Elaboração própria a partir de Pardo (2005)

Para ilustrar o conteúdo do Quadro 4, com base em Pardo (2017), foram feitas árvores RST das referidas relações. A segmentação foi feita indicando apenas as relações de oposição; todavia há outras relações nas sentenças e outras segmentações. Na Figura 3, a EDU 1 representa N, a EDU 2 apresenta uma informação, S, que se opõe a N, deixando clara a incompatibilidade entre as duas proposições. A comparação da EDU 2 em relação a EDU 1 é negativa, de modo que o leitor passa a aceitar melhor a informação apresentada em N.

Figura 3 - Exemplo relação *Antithesis*

Fonte: Baseado em Pardo (2017, p. 143)

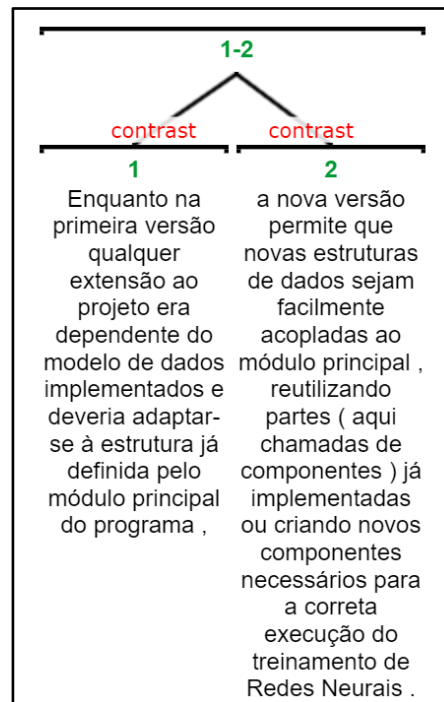
Na relação *Concession*, Figura 4, a EDU 1 é o N que apresenta uma informação válida. A EDU 2 apresenta uma provável incompatibilidade em relação à EDU 1. Essa incompatibilidade não é suficiente para diminuir a aceitação do leitor por N. O núcleo e o satélite são identificados pela direção da seta. A seta sai do satélite em direção ao núcleo.

Figura 4 - Exemplo da relação *Concession*

Fonte: Baseado em Pardo (2017, p.156)

Na relação *Contrast*, há a presença de dois Ns nas EDUs 1 e 2, já que ambos apresentam informações importantes, que são diferentes e estão sendo comparadas. Nesse tipo de relação, não há uma aceitação maior de um ou outro N, o leitor entende que são conteúdos diferentes, mas importantes.

Figura 5 - Exemplo da relação *Contrast*



Fonte: Baseado em Pardo (2017, p. 190)

Na RST, como se tem visto, são observados tanto os aspectos gramaticais explícitos quanto os implícitos no texto. Segundo Pardo (2005, p. 29 - 30), os MDs são a principal marca de indicação das relações retóricas. O autor dá o exemplo de marcadores que indicam oposição, como, *entretanto*, *contudo*, *porém*; quando esses marcadores aparecem em orações, o analista já sabe que está se tratando de uma das relações de oposição: *Antithesis* (36), *Contrast* (37), *Concession* (38).

(36) Utilizando ferramentas de teste, o conjunto de casos de teste pode ser facilmente obtido para a realização do teste de regressão. *Entretanto*, é necessário que esse conjunto seja o mais adequado possível e, principalmente, mínimo para melhorar a relação custo-eficiência desta atividade.

(37) Um concordanciador é um programa que recupera todas as ocorrências de uma determinada cadeia de caracteres em um corpus e as listam, permitindo inclusive que estas listas sejam manipuladas. *Contudo*, certos tipos de análise não podem ser obtidos através de grafia das palavras em questão, como por exemplo, informações de natureza gramatical.

(38) Em cada fase desse processo são gerados modelos (diagramáticos ou textuais) que visam a representar os requisitos de software naquela fase. *Porém*, com essa diversidade de métodos não existe um padrão de modelos para as atividades do desenvolvimento de *software*.

Tanto uma relação pode ser marcada por diferentes marcadores, como também há conectivos que podem sinalizar mais de uma relação. Segundo o autor, a relação de *Concession* pode ser indicada tanto por suas conjunções prototípicas, quanto pelas conjunções de *Contrast*, como o *mas*, (39) e (40) que vai ao encontro do que Mann e Thompson (1987) também defendem.

(39) O usuário tem a impressão de que tudo está acontecendo no seu equipamento, *mas* na realidade ele está compartilhando discos, impressoras, processamentos e tudo mais que for necessário e o sistema permitir.

(40) Obter esse resultado corresponde a introduzir estratégias pedagógicas nas aplicações hipermídias, *mas* essa tarefa é difícil porque os sistemas de autoria hipermídia tradicionais (HyperCard e ToolBook, por exemplo) possuem apenas recursos genéricos para auxiliar o autor no desenvolvimento de aplicações.

Fachada (2019) estuda os valores da conjunção *mas* em artigos de opinião na perspectiva da RST. O *corpus* é formado por 14 artigos de opinião publicados entre novembro de 2016 e maio de 2017 em quatro jornais portugueses: o Jornal de Notícias, o Jornal SOL, o Jornal Económico e o Diário Económico. Apesar de dar mais ênfase ao marcador em si, ao abordar as relações retóricas de oposição, a autora defende que a relação *Contrast* tem como marcador característico a conjunção *mas* e a relação de *Concession* tem como marcador característico a conjunção *apesar*. Em uma análise baseada em *corpus*, a autora defende a multifuncionalidade do *mas* e conclui que essas funções se enquadram muito mais na relação retórica de *Contrast*, podendo apresentar diferentes tipos de contrastes, como nos exemplos (41) a (44).

(41) “Ainda pode perder? Pode. *Mas* se assim for, teremos de aceitar humildemente que tudo aquilo que julgamos saber sobre o comportamento dos eleitores terá de ser revisto”.

(42) “E não foi por ele ser negro, *mas* que sei eu? Perguntem antes a alguém que tenha apoiado a guerra do Iraque, e vibrado com a criatividade do sistema financeiro”.

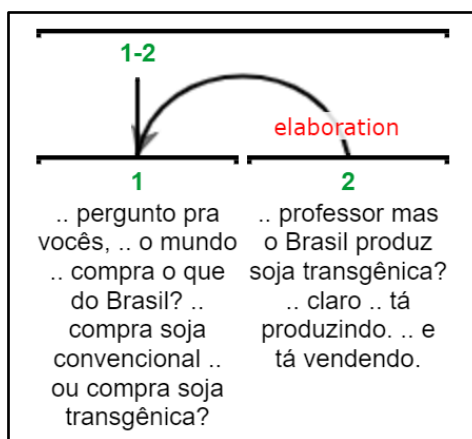
(43) “A crise financeira contribuiu não apenas para piorar essas percepções *mas* também para silenciar todos os outros temas de campanha”.

(44) “É verdade que o Obama não é um utópico e experimentalista, graças a Deus, *mas* sensato e moderado. Sabe falar e ouvir; vamos ver se também sabe fazer”.

Em (41), o *mas* introduz contra-argumentos. O segundo segmento, introduzido pela conjunção, é um argumento que vai contra o primeiro segmento. É apresentado seja por uma concessão ou por uma negação. Em (42), o *mas* tem valor de advérbio conectivo, que vai contribuir para a progressão do texto; nesse caso, há pouca ou nenhuma força contrastiva. Em (43), a conjunção tem valor aditivo e apresenta uma maior força argumentativa. Em (44), o *mas* que inicia o segundo segmento refuta e corrige o primeiro segmento; se apresenta como um advérbio conectivo.

Pecuch (2021) também desenvolveu sua pesquisa com foco no marcador *mas*, só que na relação retórica *Elaboration*, em textos transcritos de aulas e em entrevistas orais. Assim como Fachada (2019), a autora defende que tal conjunção é prototipicamente adversativa, buscando, assim como neste trabalho, embasamento tanto em estudos linguísticos, quanto na RST. Ao concluir que o *mas* é contrastivo, a pesquisadora se debruça em outras funções que a conjunção pode desempenhar, especificamente na relação de *Elaboration*, que são: 1) sinalizar pergunta retórica; 2) introduzir um subtópico; 3) tomada de turno de fala.

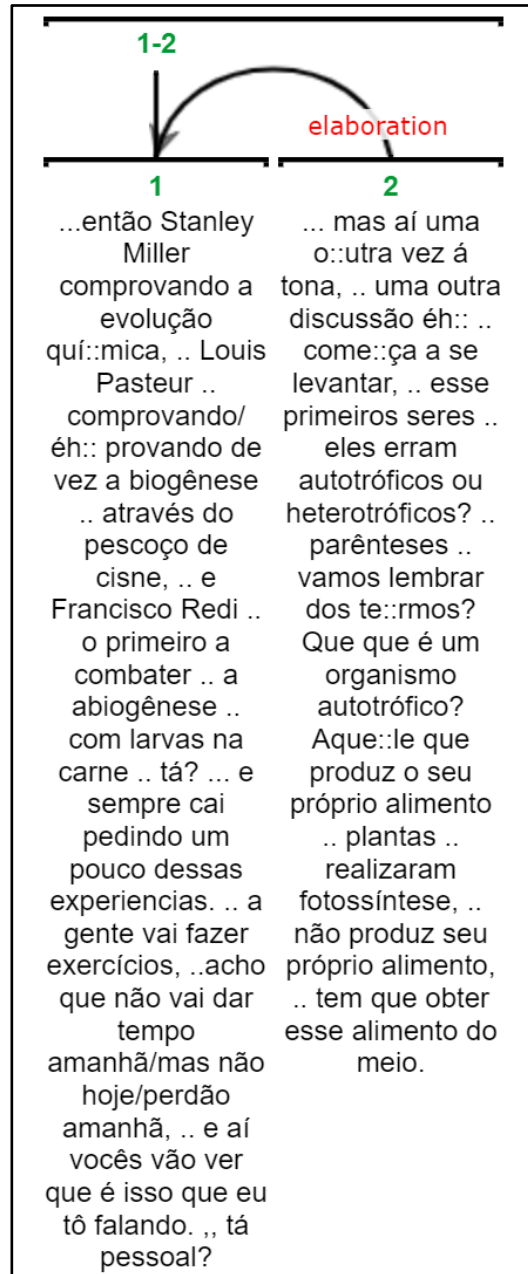
Figura 6 - Relação *Elaboration* sinalizando o *mas* em pergunta retórica



Fonte: Pecuch (2021, p. 52)

Pecuch (2021) mostra que das 15 ocorrências de *Elaboration* com o MD *mas*, 13 foram em aulas e 2 em entrevistas. Comentando o exemplo da Figura 6, a autora explica que esse resultado é muito comum, uma vez que, juntamente com a relação de *Elaboration*, está a relação de *Preparation*, entre a resposta e a pergunta retórica, causando um efeito de interesse no ouvinte sobre o conteúdo que vem em seguida. A EDU 2 elabora as informações da EDU 1.

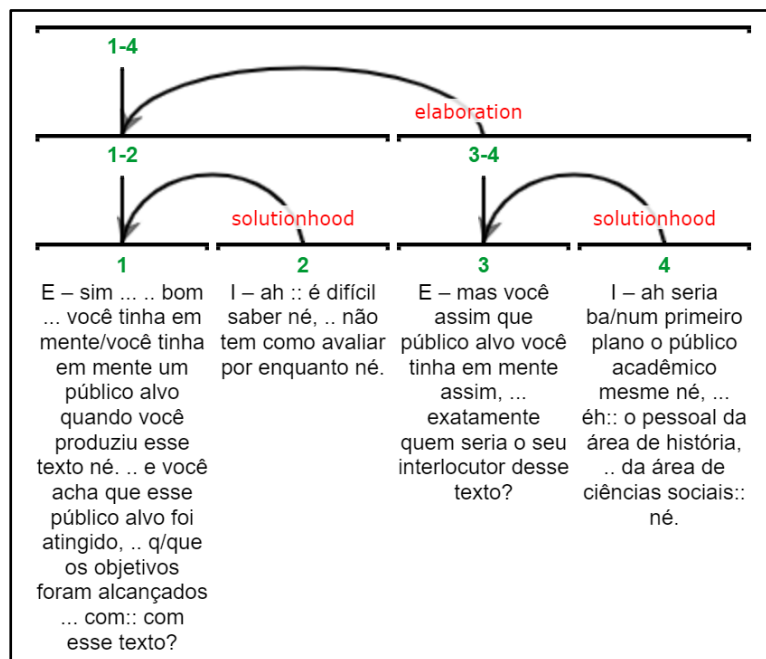
Figura 7 - Relação *Elaboration* introduzindo um subtópico



Fonte: Pecuch (2021, p.51)

Na Figura 7, o *mas* funciona como um sinalizador que introduz um subtópico, marcando a relação *Elaboration*. O segmento 1 está exercendo a função de núcleo e o segmento 2, a função de satélite.

Figura 8 - Relação *Elaboration* na tomada de turno



Fonte: Pecuch (2021, p. 54)

Na Figura 8, tem-se a marcação de duas relações, *Elaboration* e *Solutionhood*. Neste exemplo, o entrevistador faz uma pergunta (1), tem-se o núcleo da relação, ao passo que o entrevistado responde (2), tem-se o satélite da relação *Solutionhood*. Não obtendo a resposta que gostaria, o entrevistador toma a fala (3), marcada pelo *mas*, e faz uma nova pergunta mais elaborada, que funciona ainda como satélite de (1) e seu receptor entende o acréscimo de informações. Em (4), tem-se uma relação de *Solutionhood* de (3).

Tanto Fachada (2019) quanto Pecuch (2021) comprovam que um mesmo marcador, a depender do contexto, pode marcar diferentes relações retóricas. Mesmo concordando que há uma característica base para a conjunção na RST, não impede que ela apareça em outras relações desempenhando uma outra função.

O trabalho de Antonio (2011) foca na relação *Concession* em elocuições formais e em entrevistas orais. Os resultados do pesquisador quanto ao critério dos conectivos mostraram que, em 10 aulas escolhidas para análise, todos os marcadores que apareceram foram os que são propriamente concessivos, *embora*, *apesar de*, *apesar que*, *mesmo*, *mesmo que*, *se bem que*. A conjunção *mas* não chegou a aparecer nos resultados. Esse resultado é interessante pelo fato de o material de análise ser oral, o que, como visto em Neves (1999), na língua corrente, as conjunções adversativas e concessivas estão muito mais mescladas que na formalidade, ainda assim a conjunção *mas* não apareceu e a conjunção *embora* foi a mais recorrente. Uma possível explicação para esse resultado pode estar na metodologia do trabalho, quando Antonio (2011)

explica que os entrevistados são professores universitários e as aulas são no contexto de apresentação de trabalhos; em ambos, há uma maior regulação da linguagem formal, mesmo com as estratégias que o entrevistador utiliza para reduzir a tensão na hora da gravação.

Sob o viés da RST, as relações de oposição têm suas características próprias, todavia isso não impede que elas sejam sinalizadas pelos mesmos conectivos, já que, no final, são conectivos que indicam oposição de ideias e acontecimentos.

2.4 Lições aprendidas

Pelos trabalhos apresentados na subseção 2.1.1, fica evidente a diferença existente em cada relação de oposição. Apesar do seu ponto em comum, a descrição de cada uma delas é diferente, principalmente a da antítese, visto que não há marcadores prototípicos (conjunções) para ela como há na adversativa (contraste) e na concessiva, sendo identificada por meio da antonímia.

Em relação às orações adversativas e concessivas, as gramáticas focam no uso dos MDs, que nesse caso são as conjunções, como identificador dessas orações. Por meio desses trabalhos, pode-se chegar a um consenso quanto ao uso dos marcadores, principalmente o *mas*. Bechara (2009 [1928]) e Cunha e Cintra (2017 [1985]) não negam a possibilidade de haver a mescla entre as conjunções da adversativa e da concessiva, mas apresentam separadamente cada uma delas. Lima (2011) já amplia essa discussão dando margem à possibilidade de os conectivos serem usados tanto em uma quanto em outra, ao passo que Neves (1999) discute de forma ainda mais detalhada a mescla entre eles na língua em uso. Por sua vez, Margarido (2010) apresenta uma perspectiva funcionalista que vai ao encontro de Neves (1999), mostrando que, na funcionalidade do discurso, o valor semântico existente entre as construções adversativas e concessivas se confundem, indo além da sintaxe do texto apresentada pelas gramáticas normativas.

A partir dos trabalhos correlatos, a multifuncionalidade do MD *mas* é evidenciada, ainda que ele seja considerado um marcador contrastivo por natureza, ele pode aparecer em outras relações diferentes.

A RST estuda o texto e as relações de coerência existentes em suas partes. É um modelo teórico que abrange diversas áreas, como a do discurso, da argumentação, a textual, a computacional, entre outras. A cada nível da linguagem estudado pela RST, os desafios para

uma análise textual se tornam cada vez maiores, uma vez que atinge os níveis mais próximos da subjetividade humana, como o semântico, discursivo e pragmático.

As divergências entre analistas no momento de anotação de *corpus* demonstram que essa não é uma tarefa difícil apenas para as máquinas, mas para o próprio ser humano também. No entanto, é uma teoria que se relaciona com diversas outras e promove avanços linguístico-tecnológicos muito úteis para o ser humano e para aplicações computacionais.

Como visto, na RST, as relações *Antithesis* e *Concession* são relações mononucleares, que têm apenas um núcleo e o seu satélite, já a *Contrast* é multinuclear, pois tem mais de um núcleo e nenhum satélite. Também nesse sentido as relações são semelhantes às orações, subordinada, existindo uma relação de dependência e coordenada, com uma relação de independência.

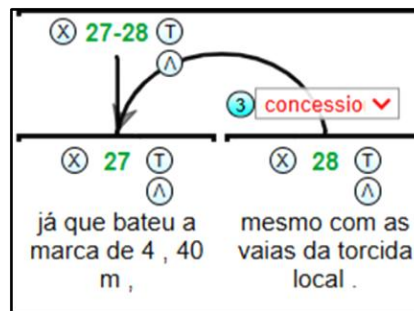
Essas associações facilitam o entendimento sobre as relações retóricas e o estudo de seus conectivos. Contudo, na RST, a seleção do núcleo, informação mais importante, não depende de uma estrutura sintática, como nas orações subordinadas, em que o conectivo que expressa a concessão sempre vem na oração dependente. Usando a Figura 9 como exemplo, a oração principal seria *já que bateu a marca de 4,40m*, como nesse caso também é o núcleo da relação; já *mesmo com as vaias da torcida local* seria a oração subordinada, encabeçada pela locução conjuntiva *mesmo com* e é o satélite da relação *Concession*. Na RST, segundo Taboada e Mann (2006), a especificação do núcleo e do satélite depende de um julgamento plausível do analista, com base no contexto semântico e nas intenções do escritor, o que vai ao encontro de Antonio (2017), que diz que as relações de coerência são de sentido e não de forma. A Figura 9 exemplifica uma relação *Concession* em que o conectivo aparece não no satélite, mas no núcleo.

Figura 9 - Exemplo de relação *Concession*



Fonte: Retirado do *corpus* CSTNews.

Figura 10 - Exemplo da relação *Concession* com núcleo invertido



Fonte: Retirado do *corpus* CSTNews.

Apesar de a RST priorizar o sentido em vez da forma na análise de suas relações, o conhecimento sobre a forma contribui explicando a base e as aplicações semânticas e sintáticas dos conectivos em diferentes situações de uso da língua.

Na próxima seção, serão apresentados os procedimentos metodológicos seguidos para o desenvolvimento da pesquisa, desde o levantamento bibliográfico ao colhimento dos dados do *subcorpus* de análise.

3. PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa é de cunho qualitativo, uma vez que estuda e descreve as relações retóricas de oposição, e quantitativo, ao fazer o levantamento da frequência em que essas relações aparecem no *corpus* de análise, e quais e quantos sinalizadores estão presentes em cada ocorrência.

Para o desenvolvimento desta investigação, algumas tarefas foram seguidas. Primeiro foi feita a leitura prévia de trabalhos relacionados à RST para construção do referencial teórico, auxiliando na delimitação do objeto de estudo.

Em seguida, foi feito um levantamento bibliográfico no *Google Acadêmico* em busca de trabalhos correlatos às relações RST em estudo. Como resultado a esta tarefa, foram selecionados quatro trabalhos para descrever as relações juntamente com as gramáticas. O Quadro 5 demonstra os passos seguidos na metodologia da pesquisa.

Quadro 5 - Resumo do processo metodológico

Divisão teórica e prática	Passos seguidos no processo metodológico		
1ª Parte	Leituras teóricas		Levantamento bibliográfico
2ª Parte	Estudar e conhecer o processo de anotação do <i>Corpus</i> CSTNews (Cardoso et al., 2011)	Estudar e conhecer o processo de anotação do Manual de sinalizadores (Dantas et al., 2024).	Buscar por relações de oposição no <i>corpus</i> para delimitar o <i>subcorpus</i> de análise.

Fonte: Elaborado pela autora.

3.1 A respeito do *corpus* CSTNews

Para esta pesquisa, foi utilizado um *corpus* jornalístico, o CSTNews (Cardoso *et al.*, 2011). Esse *corpus* é composto por 50 conjuntos de textos, denominados *clusters*, sendo que cada um deles possui 2 ou 3 textos que abordam o mesmo assunto. Os textos são do gênero notícia e estão classificados em diversos temas, a saber: esporte, mundo, dinheiro, política, ciências e cotidiano. O gênero notícia, nesse sentido, é muito propício, pois é possível existir vários textos sobre um mesmo assunto.

Essas notícias foram coletadas a partir de diferentes agências jornalísticas *on-line* brasileiras. No total, o *corpus* é constituído por 140 textos, que contabilizam 2.088 sentenças e 47.240 palavras. O CSTNews é um *corpus* multidocumental que contém diversas anotações (manuais e semiautomáticas) linguísticas, incluindo a RST. Essas anotações linguísticas fazem

com que ele seja um *corpus* amplamente utilizado em análises discursivas e em outras aplicações em PLN. O *corpus* foi concebido no âmbito do projeto SUCINTO⁵, que teve por objetivo, investigar e explorar estratégias de sumarização multidocumental de forma mais viável, inteligente e linguisticamente motivada.

Além disso, o referido *corpus* já vem sendo estudado e utilizado em outras pesquisas do projeto interinstitucional ⁶“RST para além dos marcadores discursivos”, o qual busca expandir os estudos e aplicações computacionais da RST em língua portuguesa, não apenas sobre as relações, mas de seus sinalizadores também, além de desenvolver estudos comparativos entre o PB e outras línguas naturais.

Segundo Cardoso *et al.* (2011), tanto para a segmentação quanto para a anotação RST, foi utilizada a ferramenta RSTTool (O'Donnel, 2000), que permitiu a anotação semiautomática do referido modelo teórico. A anotação em RST foi feita por 8 pessoas, 4 delas já tinham um conhecimento teórico aprofundado e tinham experiência com anotação. Todos os anotadores passaram previamente por uma formação e estudo da teoria. O treinamento da segmentação durou 2 semanas. As regras de segmentação passaram por adaptações para o português brasileiro. Em seguida, a anotação foi iniciada e durou dois meses.

A anotação foi feita durante 5 dias por semana, em reuniões de 1 hora, seguindo os seguintes passos, apresentados por Cardoso *et al.* (2011):

- 1) Foi anotado, idealmente, um grupo de textos por reunião;
- 2) Grupos de 2 ou 3 analistas anotaram cada texto do grupo textual, podendo ser considerado até 3 textos por dia;
- 3) As duplas e trios de anotadores foram reorganizados a cada reunião, para evitar preconceitos que conduzissem à reprodução de equívocos da anotação;
- 4) Quando os anotadores de um grupo, a saber as duplas ou trios, não chegavam a um consenso quanto à relação, o problema era apresentado aos demais analistas. Se ainda

⁵ Disponível em: <https://sites.icmc.usp.br/taspardo/sucinto/>. Acesso em 18 de abr. de 2025.

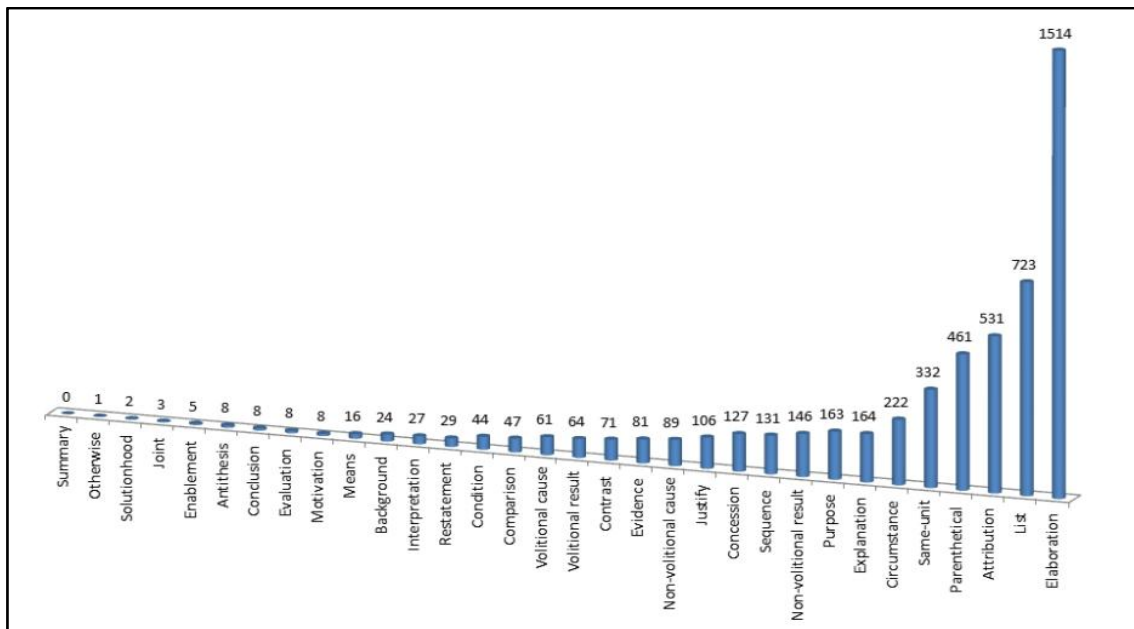
⁶ O projeto POeTiSA faz parte da [iniciativa Processamento de Linguagem Natural \(NLP2\)](#) do [Centro de Inteligência Artificial \(C4AI\)](#) da Universidade de São Paulo, patrocinado pela IBM e FAPESP (processo n° 2019/07665-4). O centro faz parte do [Programa FAPESP de Centros de Pesquisa em Engenharia](#) e está comprometido com a pesquisa de ponta em Inteligência Artificial. A iniciativa POeTiSA também é apoiada pelo Ministério da Ciência, Tecnologia e Inovação, 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 no TIC 13, DOU 01245.010222/2022-44. Furthermore, our project benefits from Scientific Initiation scholarships granted by the **Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)** through calls for proposals from the Programa Institucional de Bolsas de Iniciação Científica (PIBIC), and **Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB)** through the public notice for the "First Projects Program".

persistisse o problema, era escolhida a estratégia de escolher a relação RST mais genérica, caso ela não estivesse clara;

- 5) A cada 10 *clusters* anotados, uma reunião era realizada para chegar a um acordo de anotação. Nesse caso, todos os analistas segmentavam e anotavam o mesmo texto; primeiro, cada grupo passa os dados de sua anotação ao coordenador da sessão de anotação; logo depois, todos discutem e concordam com os resultados da segmentação. Modificações na anotação das relações poderiam ser feitas por meio de acordo quando/se necessário.

Na Figura 11, está o quantitativo de relações de acordo com a ocorrência de cada uma delas no *corpus*.

Figura 11 - Relações RST no *corpus* CSTNews



Fonte: Cardoso *et al.* (2011).

Com relação ao fenômeno da oposição, no total há 206 ocorrências, distribuídas entre 8 em *Antithesis*, 71 em *Contrast* e 127 em *Concession*. Destaca-se que, por se tratar de um *corpus* multidocumento, Cardoso *et al.* (2011) não apresenta como se dá a distribuição das relações do *corpus*: visto que os textos do conjunto tratam do mesmo evento, esse quantitativo pode ser referente a um texto de cada *cluster*, ou ainda envolver todos os textos de todos os *clusters*.

3.2 Proposta taxonômica de sinalizadores discursivos para o português

Desde o início deste trabalho, vem-se abordando o avanço dos estudos da RST para além da definição das relações. O *corpus* CSTNews, por exemplo, que possuía apenas a anotação das relações, foi reanotado, também, com os sinalizadores que evidenciam as relações. A anotação dos sinalizadores foi feita com base na tipologia estabelecida no Manual de Anotação de Dantas *et al.* (2024).

Conforme apresenta Dantas *et al.* (2024)⁷, a anotação dos sinalizadores foi feita considerando válida a pré-anotação do *corpus* com as relações retóricas, não havendo alteração nesse quesito, mas discussão quando, eventualmente, inconsistências eram identificadas. Semelhante à anotação de Cardoso *et al.* (2011), a anotação foi feita por 7 analistas, 2 deles experientes na teoria e na anotação em RST. Antes da anotação, todos os analistas passaram por um treinamento. A anotação durou dois meses.

A ferramenta utilizada para a anotação dos sinalizadores foi a rstWeb (Zeldes, 2016; Gessler; Liu; Zeldes, 2019), que é uma interface baseada em navegador, criada para a anotação em RST e para outras anotações de relações discursivas. Nessa ferramenta, é possível anotar não só as relações, mas os sinalizadores discursivos.

De acordo com Dantas *et al.* (2024), a anotação seguiu os seguintes passos:

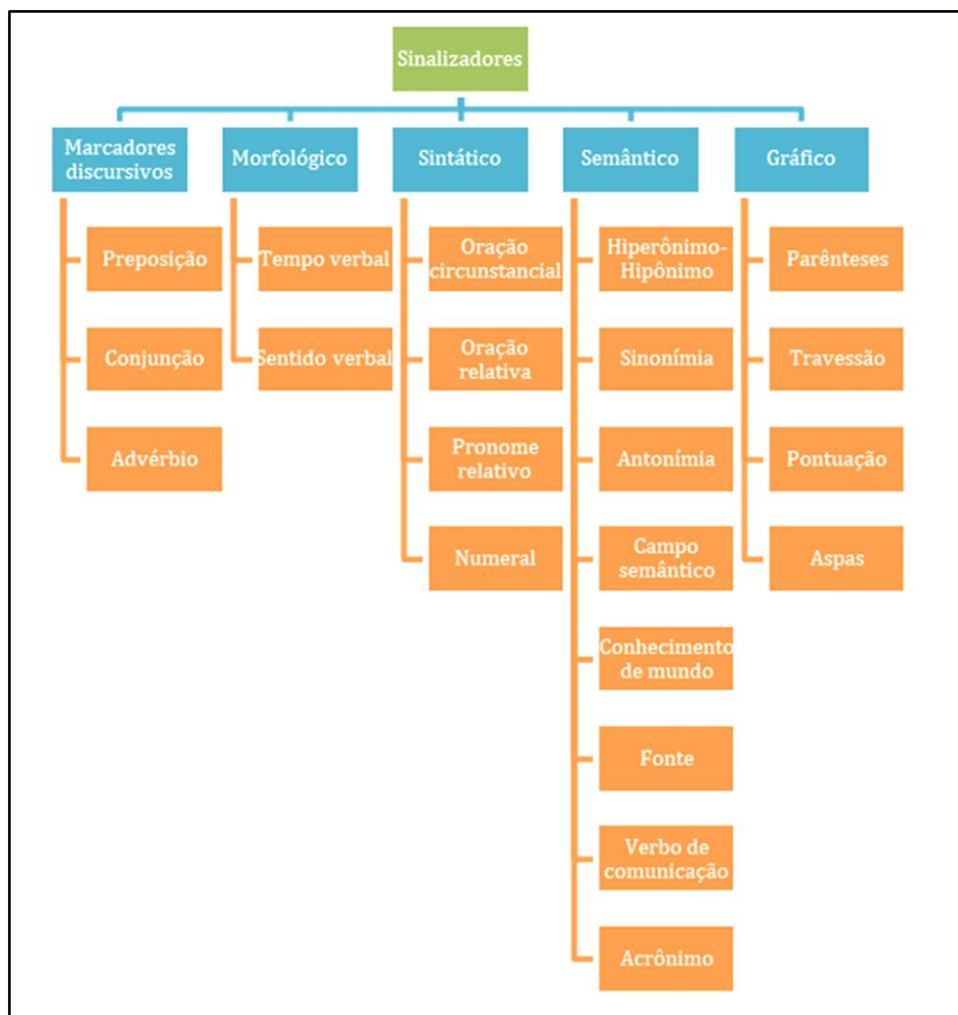
- 1) Os anotadores foram divididos em grupos trios, contando sempre com um anotador mais experiente. Desse modo, os anotadores menos experientes poderiam recorrer ao mais experiente para tirar dúvidas e minimizar, no futuro, discordâncias na anotação;
- 2) A cada texto, os trios eram modificados, para evitar preconceitos que conduzissem à reprodução de equívocos da anotação;
- 3) Todos os integrantes do trio anotavam o mesmo texto para que pudesse ser analisada a concordância entre eles, possibilitando gerar, no futuro, uma versão do texto anotado com a decisão da maioria;
- 4) Periodicamente, todos os analistas anotavam o mesmo texto para obter uma concordância geral do grupo e discutir os casos em reunião;
- 5) A anotação ocorreu apenas a nível intrassentencial, ou seja, de sinalizadores que ocorressem dentro das sentenças;
- 6) A anotação ocorreu de maneira assíncrona, com prazo de entrega de 1 semana para 2 a 3 textos por anotador;

⁷ É importante destacar que a autora desta dissertação não participou da construção do Manual de anotação, nem tampouco da anotação do corpus.

7) Ao todo, foram feitas 1 rodada de treinamento (com três *clusters*) e 8 rodadas de anotação.

Dantas *et al.* (2024) propõe uma tipologia de sinalizadores que se dividem em tipo e subtipo, como ilustrado na Figura 12.

Figura 12 - Tipologia dos sinalizadores em Dantas *et al.* (2024)



Fonte: Dantas *et al.* (2024, p. 9).

A tipologia proposta por Dantas *et al.* (2024) evidencia a abrangência dos sinalizadores que podem justificar uma relação de coerência. A Figura 13 mostra como os sinalizadores estão classificados em diferentes categorias e subcategorias que podem se combinar para garantir o efeito de sentido nos textos, como a combinação entre vírgulas e conjunções contrastivas em relações de oposição, o contraste semântico entre palavras, verbos e construções antitéticas, entre outras. É válido ressaltar que a classe de MD é apenas mais uma dentre os tipos de sinalizadores. Embora os MDs recebessem maior destaque por serem mais claramente e facilmente identificáveis, por se tratar de conectivos, eles não agem de forma exclusiva e

independente, como visto, outras categorias são usadas para identificar e justificar uma relação de coerência.

3.3 *Subcorpus de análise*

Os estudos linguísticos, ancorados em tantas teorias como visto, serviram de base para a análise da anotação dos sinalizadores discursivos no *corpus* CSTNews (Cardoso *et al.*, 2011), considerando a proposta tipológica de Dantas *et al.* (2024). Segundo Hovy e Lavid (2010), a anotação é um processo de enriquecimento de um *corpus* com a adição de várias informações, incluindo a linguística, feito por humanos ou por máquinas para um fim prático ou teórico. Os autores explicam que a tarefa de automatizar a anotação de informações gramaticais, lexicais e sintáticas já foi alcançada com uma precisão razoável (92% a 96%) em sistemas computacionais, porém automatizar a anotação de níveis linguísticos mais altos, como semânticos, pragmáticos e discursivos é mais complexo. Para atingir um nível de precisão satisfatório também em níveis mais altos da língua, primeiro deve ser feita a anotação manual de um pequeno *corpus*, para que seja testada sua concordância e confiabilidade dos resultados e só depois treinar os algoritmos de computador.

Hovy e Lavid (2010) mostram que no processo de anotação de um *corpus* há etapas bem delimitadas que requerem cuidado para se obter um resultado com nível de precisão satisfatória, como a análise da anotação e a avaliação da concordância entre os anotadores, disso depende a continuidade de outras etapas, como a anotação de *corpus* ainda maiores, pois caso haja uma baixa concordância ou discordância total entre os anotadores, indica que não há consistência nem confiabilidade para treinar automatizar o processo por sistemas computacionais.

Com o objetivo de fazer a medição e análise da concordância na anotação dos SDs em relações retóricas, Cruz, Souza e Cardoso (2024) propuseram uma comparação entre duas abordagens distintas, uma mais restrita (*gold*) e uma mais flexível (*silver*). Segundo os autores, a análise *gold* é mais restrita porque implica a concordância estrita dos anotadores sobre um determinado item lexical, enquanto a análise *silver* flexibiliza uma janela arbitrária com escopo entre -5 e +5 em relação ao item lexical alvo. Na *silver*, os anotadores podem escolher um ou outro item lexical, desde que estejam concordando com o escopo estabelecido na janela de 5 itens lexicais. O Quadro 6 exemplifica a comparação entre as duas propostas.

Quadro 6 - Exemplo de comparação entre análises *gold* e *silver*

Texto	Itens lexicais anotados		Item lexical alvo	Concordância	
	Anotador A	Anotador B		<i>Gold</i>	<i>Silver</i>
Eu adoro bolo de chocolate, torta de morango e mousse de limão.	<i>Eu, adoro, chocolate</i>	<i>Eu, adoro, bolo, torta, morango, mousse</i>	<i>Chocolate</i>	<i>Eu, adoro</i>	<i>Eu, bolo, adoro, torta, morango</i>

Fonte: Cruz, Souza e Cardoso (2024, p.3)

Seguindo a proposta *gold*, no exemplo apresentado na tabela, somente os itens lexicais *eu* e *adoro* foram anotados pelos anotadores A e B. Já na *silver*, foram considerados 5 itens lexicais dentro da janela de flexibilização do Item lexical alvo *chocolate* (posição 0), sendo *bolo* (posição -2), *torta* (posição +2) e *morango* (posição +4). Ainda que os anotadores não tenham indicado os mesmos itens lexicais, a flexibilização da janela permite que eles sejam considerados.

Os resultados apresentados pelos autores mostram que a abordagem *silver* tende a produzir um maior nível de concordância quando comparada à abordagem *gold*, considerando a flexibilização que há na janela de concordância proposta pela *silver*. Essa abordagem também permite avaliar pequenas variações nas escolhas lexicais da anotação de cada anotador.

Outra estratégia que visa alcançar um maior nível de padronização e concordância nas informações apresentadas pelos anotadores é a *Adjudicação*, em que um especialista, ou outro anotador que não esteja envolvido propriamente naquela anotação, decidirá em casos residuais em que há discordância. É claro que não há infalibilidade nessa tarefa, uma vez que uma concordância total é muito difícil, ainda mais quando se trata de níveis mais elevados da linguística como apresentado.

Tanto as abordagens *silver/gold* quanto a adjudicação apresentam vantagens no processo de anotação e uma pode dar suporte à outra. Ao se deparar com diferentes anotações de itens lexicais, o adjudicador precisa ponderar sobre o que todos anotaram e julgar dentro da janela de concordância os itens mais adequados para a anotação final. Havendo uma anotação do tipo *gold*, o adjudicador considera como válida a concordância. Neste trabalho, essa estratégia foi utilizada. A adjudicadora considerou como válida a anotação comum e julgou

sobre a mais adequada quando houve discordância ou variação. Para o caso de discordância total ou sugestões, agiu como anotadora.

Como o CSTNews é um *corpus* multidocumento que está estruturado em 50 conjuntos (*clusters*) com três textos que noticiam o mesmo evento, totalizando 150 textos no geral. Metodologicamente foi escolhido o maior texto do conjunto para a análise dos sinalizadores, pois a intenção era conseguir um maior quantitativo de relações para esta análise e evitar a redundância, já que é possível que as relações ocorram em outros textos do conjunto.

Além disso, outro critério utilizado foi a escolha, apenas, de textos em que a ocorrência da relação retórica fosse intrassentencial, ou seja, ocorresse dentro de uma mesma sentença, seguindo a metodologia de Dantas *et al.* (2024), uma vez que o tempo impediria uma nova anotação para as relações interssentenciais também. Para isso, a busca foi feita manualmente no *corpus*, entre os *clusters* 1 e 50, abrindo cada documento na ferramenta rstWeb, para estabelecer o *subcorpus* de análise. Das 100 ocorrências de relações de oposição encontradas (sendo 5 *Antithesis*, 35 *Contrast* e 60 *Concession*), ao aplicar o critério das ocorrências intrassentenciais, o número das relações caiu para 58, distribuídas entre: 2 *Antithesis*, 16 *Contrast* e 40 *Concession*. A Tabela 1 ilustra o quantitativo das relações em cada texto a ser analisado.

Tabela 1 - Relações de *Oposição* nos documentos do *corpus*

Documento	Relação (ões)	Quantidade	Intrassentenciais
D1_C2_Folha ⁸	<i>Contrast</i>	1	0 ⁹
D1_C3_Folha	<i>Contrast</i>	1	0
	<i>Antithesis</i>	1	0
D3_C4_OGlobo	<i>Concession</i>	1	1
D2_C5_GPovo	<i>Antithesis</i>	1	0
	<i>Concession</i>	1	0
D1_C7_Estadão	<i>Concession</i>	2	2
D3_C8_JB	<i>Concession</i>	3	2
D5_C10_GPovo	<i>Concession</i>	1	0
D3_C12_GPovo	<i>Contrast</i>	2	0
	<i>Antithesis</i>	1	1
D3_C13_GPovo	<i>Contrast</i>	4	3

⁸ A identificação do texto se dá pelo código: D1 - primeiro documento do *cluster*; C1 - primeiro *cluster* do *corpus*; Folha - fonte jornalística, por exemplo.

⁹ Os casos encontrados nos documentos que aparecem zerados na quarta coluna foram desconsiderados para a análise, pois se tratam de construções interssentenciais.

D3_C15_OGlobo	<i>Antithesis</i>	1	0
	<i>Contrast</i>	1	0
D2_C16_Estadão	<i>Concession</i>	2	1
	<i>Contrast</i>	2	2
D1_C17_Folha	<i>Concession</i>	3	1
	<i>Contrast</i>	1	0
D2_C18_Estadão	<i>Concession</i>	1	1
D2_C19_Estadão	<i>Concession</i>	1	1
D5_C20_Gpovo	<i>Concession</i>	2	2
D2_C21_Oglobo	<i>Concession</i>	1	1
D4_C22_JB	<i>Concession</i>	1	1
D2_C24_Estadão	<i>Concession</i>	2	2
D1_C25_Folha	<i>Contrast</i>	3	0
	<i>Concession</i>	2	0
D1_C27_Folha	<i>Concession</i>	9	7
D1_C29_Folha	<i>Concession</i>	1	1
	<i>Contrast</i>	2	1
D3_C30_Oglobo	<i>Contrast</i>	1	0
D4_C32_Gpovo	<i>Contrast</i>	1	0
D2_C33_Estadão	<i>Concession</i>	1	1
	<i>Contrast</i>	2	2
D2_C34_Estadão	<i>Contrast</i>	1	0
D1_C35_Folha	<i>Concession</i>	2	2
D4_C36_JB	<i>Concession</i>	4	3
	<i>Contrast</i>	1	1
D1_C37_Oglobo	<i>Contrast</i>	1	1
D5_C41_Gpovo	<i>Concession</i>	1	1
D2_C42_Estadão	<i>Contrast</i>	3	3
	<i>Concession</i>	3	2
D4_C43_JB	<i>Concession</i>	2	0
	<i>Contrast</i>	1	0
D1_C44_Folha	<i>Contrast</i>	3	1
	<i>Concession</i>	3	1
D3_C45_Oglobo	<i>Concession</i>	2	1
	<i>Contrast</i>	2	2
D4_C46_JB	<i>Concession</i>	2	2
D5_C47_Gpovo	<i>Contrast</i>	1	0
D1_C48_Folha	<i>Concession</i>	3	2

D1_C49_Folha	<i>Concession</i>	1	0
D3_C50_Oglobo	<i>Contrast</i>	1	0
	<i>Concession</i>	3	2
	<i>Antithesis</i>	1	1

Fonte: Elaborado pela autora.

Durante o processo de verificação de cada texto, uma planilha foi criada para registrar informações importantes para a análise dos dados. Dentre essas informações, estão o documento de origem, a porção do texto, a quantidade de EDUs de cada cluster, a altura¹⁰ em que ocorre a relação na árvore RST, a indicação de cada anotador separadamente, as concordâncias (*gold* e *silver*) do sinalizador e os tipos e subtipos (de acordo com Dantas et al. (2024) dos sinalizadores. As porções do texto foram retiradas do *corpus*, que dispõe de todos os textos sem anotação. Para isso, acessou-se a ferramenta rstWeb e identificou-se o nome do documento e a porção em que a relação aparecia. Em seguida, o texto foi copiado e colado do *corpus* para a planilha de dados. A Figura 13 ilustra parte dessas informações. Nos apêndices, estão todos os demais casos.

Figura 13 - Ilustração dos dados de análise na planilha

Documento de origem	Porção de texto	Qnt. Total de EDUs	Ocorrência da Concession na EDU	Altura da ocor. da Concession na árv	Documento de origem	Anotador 1
D3_C4_OGlobo	"Amanhã, a temperatura máxima pode chegar a 22 graus, mas na quarta volta a cair"	37	12	9 de 9	D3_C4_OGlobo	mas
					D3_C4_OGlobo	—
D1_C7_Estadão	"O mesmo ocorre com o sexto das aias marrons, corpos pesados o suficiente para gerar calor, mas não o bastante para iniciar uma fusão nuclear como as estrelas legítimas"	38	18/jan	8 de 10	D1_C7_Estadão	mas
					D1_C7_Estadão	—
D1_C7_Estadão	"Ambos os mundos têm massa semelhante à de outros exoplanetas já catalogados, mas não giram em torno de uma estrela"	38	7	7 de 10	D1_C7_Estadão	mas
					D1_C7_Estadão	—
					D1_C7_Estadão	—
D3_C8_JB	"Embalado pela vitória, o Brasil começou o segundo set na frente, mas permitiu a reação dos donos da casa, que chegaram a empatar a parcial em 17"	25	17 - 20	4 de 6	D3_C8_JB	mas
						—

Fonte: Elaborado pela autora.

¹⁰ Infelizmente, até a conclusão deste trabalho não foi possível investigar como a quantidade de EDUs e a altura em que as relações aparecem nas árvores podem contribuir para o entendimento de como as relações de *Oposição* se comportam. Esses dados serão utilizados em investigações futuras.

A organização dos dados em planilha foi escolhida por ser uma maneira simples e que proporciona a aplicação de filtros de busca que facilitam o colhimento de informações específicas, como, por exemplo, a visualização somente de documentos em que o MD *mas* apareceu ou ver o quantitativo ou as combinações com outros sinalizadores.

Para cada relação de oposição, foi feita uma planilha. A Figura 13 mostra uma parte da planilha da relação retórica *Antithesis*. No *documento de origem*, se identifica de onde foi retirada a porção de análise, de qual texto está se tratando. A *porção de texto* deixa visível a ocorrência da relação. Um detalhe sobre esse tópico na planilha é que o satélite da relação foi destacado em vermelho, tanto na *Antithesis* quanto na *Concession*, pois facilita a análise da escolha nuclear dessas relações. A *quantidade de EDUs* mostra o tamanho do texto. Também há a informação da *ocorrência da relação na EDU*, indicando se ela aparece no início, meio ou fim do texto. Em seguida, vem a indicação da *ocorrência na altura da árvore*. Para identificar os sinalizadores, foi necessário acessar a anotação individual de cada anotador e registrar em *Anotador 1*, *Anotador 2* e *Anotador 3*. O tópico *Adjudicadora* indica a concordância, seja a dos anotadores que indicaram o mesmo *sinalizador*, seja da adjudicadora (quando houve discordância ou variação entre os anotadores), papel desempenhado pela autora desta dissertação, que analisou todas as anotações e indicou a mais coerente, com base no que propôs Dantas *et al.* (2024). A indicação do *tipo* e *subtipo* dos sinalizadores segue a tipologia de Dantas *et al.* (2024).

Em suma, a metodologia dessa pesquisa foi dividida em duas partes: teórica e prática. Na teórica, foram feitas leituras e o levantamento bibliográfico do trabalho. Na parte prática, o *corpus* foi descrito em suas duas anotações, a primeira com base em Cardoso *et al.* (2011) e a segunda, dos sinalizadores, em Dantas *et al.* (2024); por fim, o *subcorpus* foi delimitado a partir do colhimento dos fragmentos de textos e elaboração da planilha com informações relevantes para a análise de cada caso, atuando como adjudicadora e como analista dos dados.

4. ANÁLISES E DISCUSSÕES

Nesta seção, o *subcorpus* é analisado e discutido, retomando conceitos importantes do aporte teórico apresentado na seção 2. Seguindo a ordem das relações apresentada na seção 2: *Antithesis*, *Concession* e *Contrast*.

4.1 Análise da relação *Antithesis*

No tocante à relação *Antithesis*, segundo a Tabela 1, verifica-se que ela ocorre, de forma intrassentencial, apenas em dois textos, ambos tratam de política e são analisados nesta seção. Nos Quadros 7 e 8, há a relação de sinalizadores únicos e combinados que aparecem nessa relação.

Quadro 7 - Sinalizadores únicos da relação *Antithesis*

Exemplo	Tipo	Subtipo	Sinalizador único
"que teria se comprometido a elevar o limite de isenção da CPMF para quem tem renda até R\$ 4.340, <i>em vez dos</i> R\$ 1.642 negociados até então"	Marcador discursivo	Preposição	Em vez dos

Fonte: Elaborado pela autora.

Quadro 8 - Sinalizadores combinados na relação *Antithesis*

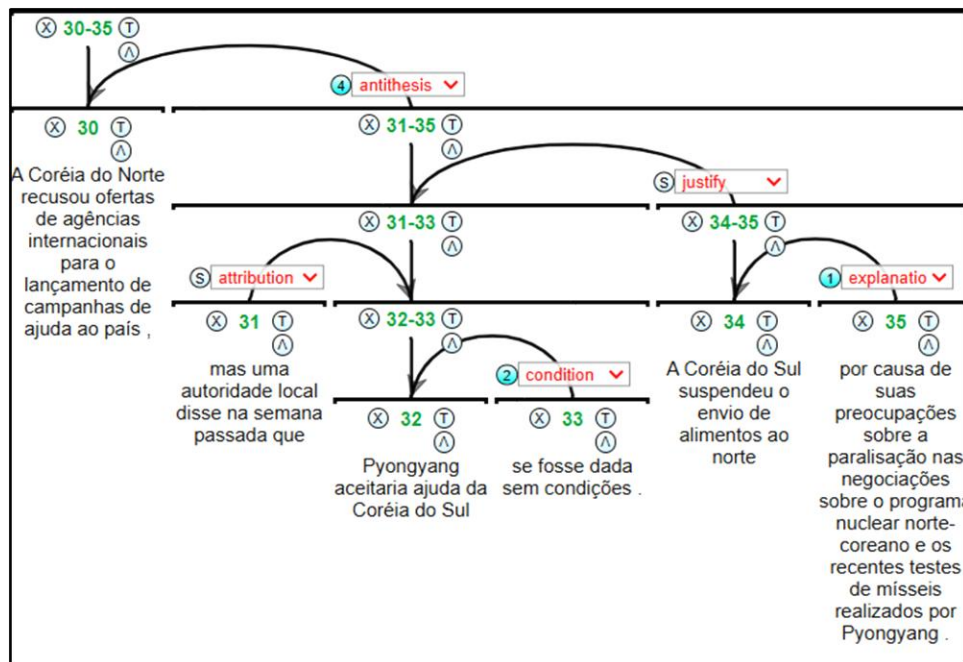
Exemplo	Sinalizadores	Combinações
"A Coreia do Norte <i>recusou</i> ofertas de agências internacionais para o lançamento de campanhas de ajuda ao país, <i>mas</i> uma autoridade local disse na semana passada que Pyongyang <i>aceitaria</i> ajuda da Coreia do Sul se fosse dada sem condições. A Coreia do Sul suspendeu o envio de alimentos ao norte por causa de suas preocupações sobre a paralisação nas negociações sobre o programa nuclear norte-coreano e os recentes testes de mísseis realizados por Pyongyang."	recusou (sentido verbal) + recusou/aceitaria (tempo verbal) + recusou/aceitaria (antonímia) + mas (conjunção) + vírgula (pontuação)	MORF+ MORF+SEM+ MD+GRAF

Fonte: Elaborado pela autora.

Como visto, os sinalizadores geralmente aparecem combinados, contribuindo mutuamente para a sinalização da relação *Antithesis* no modelo RST, considerando o gênero

textual e o registro linguístico utilizado, a saber, formal. Na *Antithesis*, tem-se um caso de sinalizador único e outro caso de sinalizadores combinados.

Figura 14 - D3_C12_GPovo - *Antithesis*



Fonte: Retirado do corpus CSTNews.

Na relação *Antithesis*, segundo o conceito apresentado pela RST (Quadro 4, p.33), o N e o S estão em oposição, havendo uma incompatibilidade entre eles, fazendo com que N seja mais aceito pelo leitor. Lima (2011) defende que a antítese é a contraposição de uma palavra ou frase a outra de significação oposta. Na Figura 15, essa incompatibilidade está evidente pela *Antonímia* entre os verbos *recusou* e *aceitaria*, que mostram como a informação das EDUs 31 a 33 se opõe à informação da EDU 30. Além da contraposição de ideias por meio desses verbos, o MD *mas* mostra como o S está em oposição ao N; no entanto, a informação mais aceita pelo leitor continua sendo a nuclear.

Sobre a anotação dos sinalizadores, o Quadro 9 apresenta os sinalizadores, tipos e subtipos destacados pelos anotadores e o parecer da adjudicadora sobre a concordância entre eles.

Quadro 9 - Anotação D3_C12_GPovo - *Antithesis*

Anotador 1	Anotador 2	Anotador 3	Adjudicadora	Tipo	Subtipo
—	—	Recusou	Recusou	Morfológico	Sentido verbal
—	—	Recusou/ aceitaria	Recusou/ aceitaria	Morfológico	Tempo verbal
—	Recusou/ aceitaria	Recusou/ aceitaria	Recusou/ aceitaria	Semântico	Antonímia
—	—	Mas	Mas	MD	Conjunção

Fonte: Elaborado pela autora

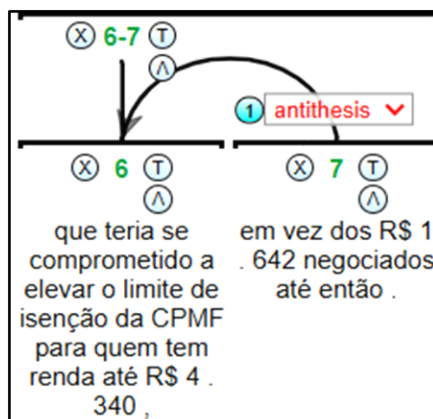
Nessa porção, o anotador 1 não fez nenhum apontamento, deixando a concordância para o que apontaram os anotadores 2 e 3. Considerando as propostas *gold* e *silver*, tem-se uma concordância *gold* no subtipo Antonímia, pois os dois fizeram a mesma anotação. Além disso, o anotador 3 apontou o par de verbos *recusou/aceitaria* como Tempo Verbal, *recusou* como Sentido Verbal e a conjunção *mas*.

Apesar de Dantas *et al.* (2024) apresentar dois subtipos para o tipo morfológico, não há uma diferença relevante, aparente, entre os dois, nem nos sinalizadores apontados, nem na descrição apontada no manual, que, inclusive, é a mesma, se tratam de *Verbos ou locuções verbais que indicam sucessão, coincidência ou ruptura temporal*. No caso desse exemplo em questão (Figura 15), o Tempo Verbal é suficiente para indicar uma oposição entre os tempos verbais apresentados (*pretérito perfeito e futuro do pretérito*).

Em relação ao MD *mas*, ao longo da análise do *subcorpus* feita a partir da planilha, há uma combinação entre os sinalizadores *mas* e a *vírgula*, porém ela não foi apontada em todas as ocorrências por nenhum dos anotadores. Essa combinação se deve às construções adversativas, em que não apenas o *mas*, mas todas as conjunções adversativas (*mas, porém, entretanto, no entanto, todavia*) são antecedidas pelo sinal gráfico da vírgula.

O segundo exemplo da relação *Antithesis* no *subcorpus* é apresentado no texto D3_C50_OGlobo. Nesse exemplo, a oposição entre N e S é percebida pela locução prepositiva *em vez de*, que contrasta uma informação à outra. Esse sinalizador foi apontado por todos os anotadores, também foi o único anotado por todos. Além desse sinalizador, vê-se, neste trabalho, a necessidade de acrescentar o Numeral à anotação dos SDs, pois há uma diferença entre os dados R\$ 4.340 e R\$ 1.642 que são comparados um ao outro na sentença, conforme mostra a Figura 15.

Figura 15 - D3_C50_OGlobo - *Antithesis*



Fonte: Retirado do corpus CSTNews.

Além da combinação entre conjunções adversativas e a *vírgula*, visto na ocorrência anterior, há a recorrência desse sinal gráfico em todas as relações intrassentenciais mononucleares do *subcorpus* – incluindo esta –, separando núcleo do satélite. Essa informação é um indicativo de que esse sinal gráfico por si já é um sinalizador de oposição em relações mononucleares intrassentenciais. É importante considerar que o gênero notícia apresenta, muitas vezes, informações com pontos de vistas diferentes, argumentações e faz uso de uma linguagem que evoca determinados sinalizadores de oposição, a *vírgula*, nesse caso, é um deles.

A relação *Antithesis* foi a que menos teve recorrência no *corpus*, em todo ele, há apenas 8 ocorrências dessa relação. Nos 2 exemplos analisados, há a presença de diferentes sinalizadores, dentre eles, o MD *mas*, mas também há a presença da *vírgula*, comum às ocorrências mononucleares intrassentenciais.

Por sua escassa aparição no *corpus*, essa relação teve apenas uma ocorrência de cada SD, exceto a *vírgula*; no entanto, houve uma variação nos tipos de sinalizadores, não se limitando apenas à Antonímia.

4.2 Análise da relação *Concession*

No tocante à relação *Concession*, segundo a Tabela 1, há um total de 40 ocorrências intrassentenciais no *subcorpus*, e que foram analisadas. Os temas variam entre clima, conhecimentos gerais, política, ciência e esportes. Nos Quadros 10 e 11, há a relação de sinalizadores únicos e uma amostra de sinalizadores combinados. Todos os sinalizadores e combinações estão nos apêndices.

Quadro 10 - Sinalizadores únicos da relação *Concession*

Exemplo	Tipo	Subtipo	Sinalizador único
"Sem citar nominalmente o adversário, Alckmin criticou de novo Lula ao comentar especulações de que o petista, condenado na vitória no primeiro turno, já estaria fazendo planos sobre sua nova equipe ministerial"	Sintático	Oração circunstancial	Sem citar
"Mesmo com esta melhora, ele continuará internado"	MD	Conjunção	Mesmo com
"Sabemos que não temos chance de ganhar [a votação], quanto mais levar aqui, mais difícil fica para o governo no Senado"	Sintático	Valência verbal	sabemos que não temos chance

Fonte: Elaborado pela autora.

Quadro 11 - Sinalizadores combinados na relação *Concession*

Exemplo	Sinalizadores	Combinações
"Amanhã, a temperatura máxima pode chegar a 22 graus, mas na quarta volta a cair"	Amanhã, a temperatura máxima pode chegar a 22 graus (oração circunstancial) + mas (conjunção) + vírgula (pontuação)	SIN+MD+GRAF
"O mesmo ocorre com o sexto das añas marrons, corpos pesados o suficiente para gerar calor, mas não o bastante para iniciar uma fusão nuclear como as estrelas legítimas"	Mas (conjunção) + vírgula (pontuação) + suficiente/ não o bastante (antonímia)	MD+GRAF+SEM
"Ambos os mundos têm massa semelhante à de outros exoplanetas já catalogados, mas não giram em torno de uma estrela"	mas (conjunção) + vírgula (pontuação) + não (advérbio) + giram em torno de uma estrela (conh. de mundo)	MD+GRAF+MD+SEM
"Ex-líder do PP, Pedro Henry (MT) cogitou sair da função, mas teria desistido da ideia"	Mas (conjunção) + vírgula (pontuação) + teria desistido (tempo verbal)	MD+GRAF+MORF

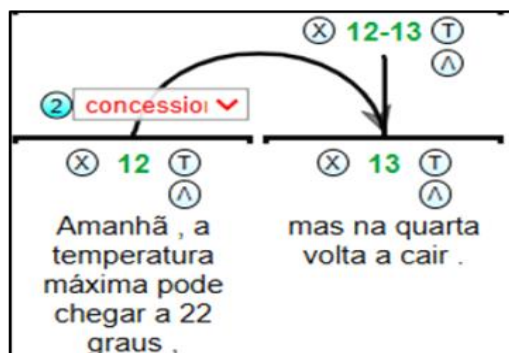
Fonte: Elaborado pela autora.

Dentre as 40 ocorrências, somente 3 foram anotadas com sinalizadores únicos. No entanto, como foi mencionado, a *vírgula* desempenha um importante papel para o sentido de contraste e quebra de expectativa nas relações de oposição, apesar de não terem sido anotadas em todos os casos. Já nas combinações, por serem muitas, foram selecionadas algumas (Quadro 11) para demonstrar como os sinalizadores estão combinados promovendo o sentido de

oposição. As combinações variam, mas é notável que a combinação MD+GRAF+SEM aparece com mais frequência, 8 vezes.

A seguir, alguns exemplos elucidam como esses sinalizadores ocorrem na relação *Concession*.

Figura 16 - D3_C4_OGlobo - *Concession*



Fonte: Retirado do corpus CSTNews.

Na *Concession*, há uma incompatibilidade semântica aparente entre N e S, isso numa definição 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 16, a EDU 12 é o N (informação principal) e a EDU 13 é o S. Nesse caso, o fato de *na quarta a temperatura voltar a cair* não impede que *amanhã a temperatura aumente e chegue a 22 graus*. Há uma incompatibilidade entre as proposições, em que uma fala da temperatura máxima e a outra fala de uma queda dessa temperatura.

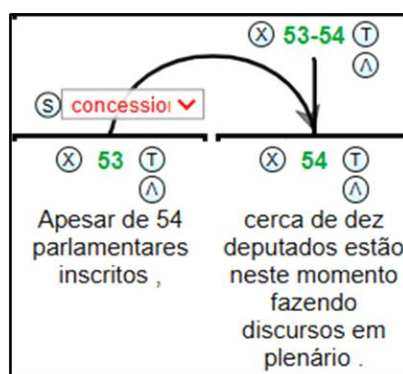
Nessa ocorrência, todos os anotadores indicaram a *conjunção mas*, que, como visto na explanação teórica deste trabalho, é uma conjunção típica das orações adversativas, mas que também pode aparecer em orações concessivas (Lima, 2011; Neves, 1999; Margarido, 2010) e apenas um anotador indicou a Oração Circunstancial no trecho *Amanhã, a temperatura máxima pode chegar a 22 graus*, existindo uma relação nominal-temporal entre *amanhã* e *quarta*.

Além desses sinalizadores, acrescentaria à anotação a oposição entre *máxima* e *cair*, pois ambos os itens lexicais se referem à temperatura, evidenciando uma antonímia entre a temperatura mais alta e a possibilidade de ela baixar. É importante resgatar o conceito de antítese apresentado por Azeredo na seção 2.1.1. O autor defende que a antítese não se apresenta apenas quando há palavras antônimas específicas, mas também em enunciados de conteúdos

opostos, que é o caso da figura 16. A *vírgula* também está presente e faz combinação com o MD *mas*.

O texto D5_C20_GPovo, Figura 17, de tema político, aborda uma votação na Câmara de deputados. É perceptível a incompatibilidade entre as proposições. S traz 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.

Figura 17 - D5_C20_GPovo - *Concession*

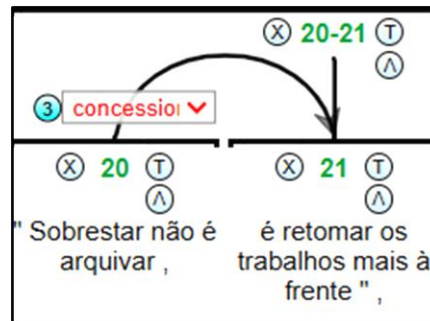


Fonte: Retirado do corpus CSTNews.

Nesse exemplo, 2 anotadores não fizeram anotação alguma, talvez em meio ao exercício exaustivo da tarefa essa ocorrência tenha passado despercebida. O único anotador a indicar os sinalizadores apontou *Apesar de*, mas classificou o subtipo como *Conjunção*. Essa construção tanto existe como locução conjuntiva – *apesar de que* –, que termina em conjunção, quanto como locução prepositiva – *apesar de* –, que termina em preposição. Desse modo, optou-se, no presente trabalho, por tratar por *Preposição* o subtipo. Para diminuir a dificuldade na anotação de conceitos gramaticais, é recomendada a consulta à própria gramática para sanar qualquer dúvida.

Além dos MDs, que sempre ganham destaque nas relações por serem mais característicos, na *Concession* também pode existir oposição por antonímia, como consta na Figura 18.

Figura 18 - D2_C42_Estadão - *Concession*



Fonte: Retirado do corpus CSTNews.

No exemplo da Figura 18, a contrariedade está evidente por meio do advérbio negativo *não*, que faz uma correção de significado nas informações das EDUs 20 e 21. A antonímia percebida nos verbos *arquivar* e *retomar* também mostra a oposição entre as proposições, uma vez que *arquivar* tem sentido de *guardar* e *retomar* tem sentido de *trazer de volta*. O Quadro 12 mostra a anotação desse texto feita pelos anotadores.

Quadro 12 - Anotação D2_C42_Estadão - *Concession*

Anotador 1	Anotador 2	Anotador 3	Adjudicadora	Tipo	Subtipo
Não é	—	Não	Não	MD	Advérbio
—	é retomar os trabalhos mais à frente	—	é retomar os trabalhos mais à frente	Sintático	Oração Circunstancial
—	—	Vírgula	Vírgula	Gráfico	Pontuação
—	—	É arquivar/ é retomar	É arquivar/ é retomar	Semântico	Antonímia

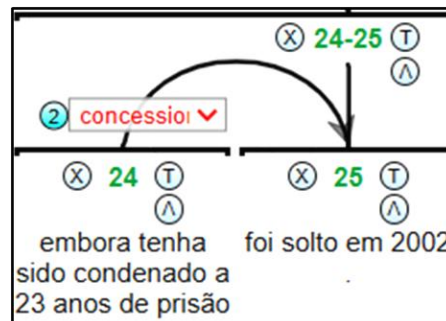
Fonte: Elaborado pela autora.

No primeiro sinalizador, um anotador indicou Preposição como subtipo e outro indicou Advérbio. Optou-se por deixar apenas o sinalizador *não* e classificá-lo como Advérbio. A vírgula foi anotada por apenas um anotador, mas como indicado anteriormente, ela por si já é um sinalizador de oposição intrassentencial em relações mononucleares. Nesse exemplo, praticamente cada anotador fez uma anotação diferente, a qual uma completa a informação da outra.

A partir das proposições, é possível fazer uma associação entre as relações *Concession* e *Contrast*, com as orações concessivas e adversativas, como foi feita na seção 2. Na RST, as relações não têm sinalizadores específicos para cada uma delas, mas é percebido que os conectivos que aparecem são próprios das orações adversativas e/ou concessivas. No exemplo

a seguir, tem-se uma conjunção concessiva *embora*. De acordo com a descrição da RST para a relação *Concession*, a informação do S não é suficiente para impedir a informação do N, e é justamente ela que é mais aceita pelo leitor, é a informação que mais importa na notícia, conforme a Figura 19.

Figura 19 - D1_C35_Folha - *Concession*



Fonte: Retirado do corpus CSTNews.

Nessa anotação, houve uma concordância unânime do sinalizador *embora*, todos os anotadores identificaram a conjunção como um indicador de oposição. A vírgula foi anotada por apenas um anotador. Além do que foi anotado, vale ressaltar a antonímia em *condenado* e *solto*. Sendo assim, a anotação ficaria: *embora* (MD - conjunção) + *vírgula* (sinal gráfico) + antonímia (semântica).

Os *clusters* escolhidos para exemplificar os sinalizadores da relação *Concession* deixam claro que há uma variedade de SDs nessa relação. Percebe-se que os Marcadores Discursivos aparecem em maior quantidade, principalmente as conjunções. Dentre as 37 ocorrências de conjunções, 26 delas são do conectivo *mas*. Assim como, dentre as 11 ocorrências de advérbios, 8 são do advérbio negativo *não*. O tipo Semântico aparece em seguida com um grande quantitativo de ocorrências, variando entre os seus subtipos. Esses dados comprovam a recorrência das combinações MD+GRAF+SEM. Os tipos Sintático e Morfológico aparecem em menor quantidade.

4.3 Análise da relação *Contrast*

No tocante à relação *Contrast*, segundo a Tabela 1, verifica-se que ela ocorre 16 vezes de forma intrassentencial. Os temas abordados variam entre política e cotidiano. No Quadro 13,

há a relação dos sinalizadores únicos que foram anotados, já no Quadro 14, são os sinalizadores combinados.

Quadro 13 - Sinalizadores únicos da relação *Contrast*

Exemplo	Tipo	Subtipo	Sinalizador único
"que vê o Brasil indo tão bem em tantas áreas - índices nunca antes alcançados, capital estrangeiro entrando, notoriedade no mundo todo - e muito mal em questões básicas que a gente não consegue resolver, como a questão da violência."	Sintático	Oração circunstancial	que vê o Brasil indo tão bem em tantas áreas

Fonte: Elaborado pela autora.

Quadro 14 - Sinalizadores combinados da relação *Contrast*

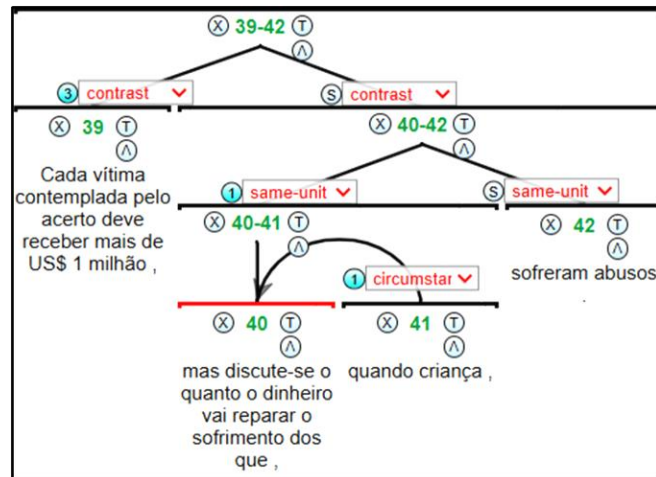
Exemplo	Sinalizadores	Combinações
"o grupo pretendia enviar uma equipe à região, mas foi impedido por soldados"	pretendia/ foi (tempo verbal) + mas (conjunção) + vírgula (pontuação)	MORF+MD+GRAF
"Uma página na internet pró-Tigres Tâmeis responsabilizou o governo pelas mortes, mas o exército rejeitou a acusação"	responsabilizou/ rejeitou (tempo verbal) + mas (conjunção) + vírgula (pontuação) + responsabilizou/ rejeitou a acusação (conh. de mundo)	MORF+MD+GRAF+SEM
"vou instaurar todos os processos juntos, mas a ideia é que os 15 casos mais graves, que têm provas contundentes, sejam julgados na frente"	todos os processos/ os 15 casos mais graves (hiper e hiponímia) + mas (conjunção) + vírgula (pontuação)	SEM+MD+GRAF
"Cada vítima contemplada pelo acerto deve receber mais de US\$ 1 milhão, mas discute-se o quanto o dinheiro vai reparar o sofrimento dos que, quando criança, sofreram abusos."	Mas (conjunção) + vírgula (pontuação) + US\$ 1 milhão /dinheiro (campo semântico) + receber/reparar (antonímia)	MD+GRAF+SEM+SEM

Fonte: Elaborado pela autora.

Para as combinações da relação *Contrast* também foram selecionadas algumas, estando as demais nos apêndices do trabalho. As combinações da *Contrast* se assemelham às da *Concession*, variadas, mas os tipos Marcadores Discursivos e o Semântico aparecem relacionados com mais frequência. A única ocorrência de sinalizador único feita pelos anotadores também não demonstra ser única, há outros sinalizadores que estão se combinando, mas esse exemplo será discutido mais à frente.

A seguir, alguns exemplos serão comentados, para mostrar um pouco como os SD aparecem nas relações.

Figura 20 - D1_C29_Folha – Contrast



Fonte: Retirado do corpus CSTNews.

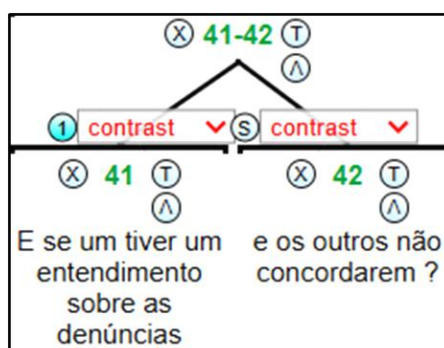
Na relação *Contrast*, há dois núcleos tidos como igualmente importantes e independentes, que contrastam entre si. As EDUs 39 e 40-42 estão em oposição. O texto apresenta um acordo feito entre a Igreja Católica e as vítimas de abuso sexual por parte dos padres. Na EDU 39, é dito o valor recebido por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

Esse é mais um exemplo em que dois dos três anotadores não fizeram nenhuma anotação. Assim, os sinalizadores apontados serão de apenas um dos anotadores. No exemplo ilustrado, foram anotados 3 sinalizadores segundo a tipologia apresentada por Dantas *et al.* (2023). De forma prototípica, tem-se o MD *mas* combinado com a *vírgula*, apresentando uma oposição entre as duas informações. Além dele, há dois sinalizadores Semânticos, um de subtipo Campo Semântico, apresentado em *US\$ 1 milhão /dinheiro*, e outro em forma de Antonímia, por meio dos verbos *receber* e *reparar*.

Como não há a anotação dos outros anotadores para comparar a concordância entre eles, a adjudicadora não pode avaliar concordância, e sim sugerir a anotação de outros sinalizadores se for o caso. A Antonímia indicada nos verbos *receber* e *reparar* não parece acontecer de fato, pois o *mas* introduz o argumento de que o dinheiro recebido, não importando o quão alto seja o valor, não repara o dano feito às vítimas; há uma oposição nas proposições, mas não é por meio dos verbos apontados, uma vez que o oposto de *receber* é *dar*, e o oposto de *reparar* é *danificar*.

No exemplo anterior, há o *mas* como MD, evidenciando o contraste entre as proposições. Mas além desse sinalizador, há outros MDs, como no texto D1_C44_Folha. Esse exemplo, ilustrado na Figura 21, aborda os processos contra Renan Calheiros, contando com o julgamento de diferentes relatores.

Figura 21 - D1_C44_Folha – *Contrast*



Fonte: Retirado do corpus CSTNews.

No exemplo da Figura 21, foram anotados três sinalizadores que identificam o contraste de ideias, como a conjunção *e se/e* nas duas proposições, comparando-as; o advérbio de negação *não* que evidencia a possível divergência no entendimento dos relatores, bem como a antonímia em *tiver um entendimento/não concordarem*. Essa descrição é a visão geral da anotação. Abaixo, no Quadro 15, está a anotação de cada anotador.

Quadro 15 - Anotação D1_C44_Folha – *Contrast*

Anotador 1	Anotador 2	Anotador 3	Adjudicadora	Tipo	Subtipo
Não	Não	CPD	Não	MD	Advérbio
-	E se/e		E se/e	MD	Conjunção
-	Tiver um entendimento/não concordarem		Tiver um entendimento/não concordarem	Semântico	Antonímia

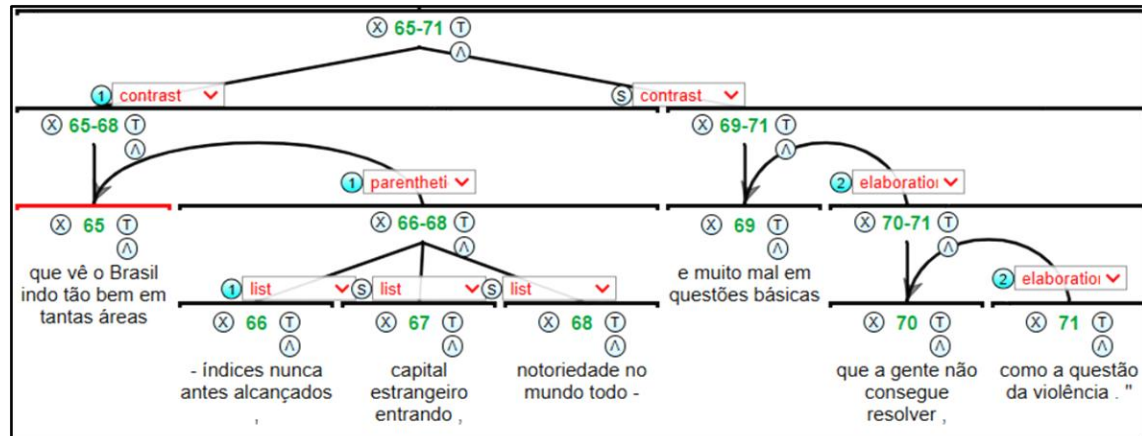
Fonte: Elaborado pela autora.

Como apresenta o Quadro 15, houve uma concordância entre dois anotadores em um dos sinalizadores. Já o anotador 3 usou a tag *CPD*, que significa *casos para discussão*, ou seja, quando há dúvida por parte do anotador na catalogação do SD. Ademais desses sinalizadores, acrescentaria à anotação o contraste entre os itens lexicais *um* e *outros*.

Além de exemplos com a presença de MDs, há também exemplos somente com SDs, como é o caso do texto D3_C45_OGlobo, trecho ilustrado na Figura 22. Esse documento relata

um roubo que o apresentador Luciano Huck sofreu enquanto estava parado em um semáforo de trânsito.

Figura 22 - D3_C45_OGlobo - *Contrast*



Fonte: Retirado do corpus CSTNews.

Nesse exemplo, dos três anotadores, dois fizeram a mesma anotação, que, por sinal, foi a única feita nesse caso. Foi sinalizado apenas a Oração Circunstancial em *que vê o Brasil indo tão bem em tantas áreas*. Mas o sinalizador do tipo Semântico poderia ser anotado, identificado por meio da antonímia entre *tão bem* e *muito mal*. Existe uma oposição de ideias ao mostrar o crescimento do país em tantas áreas, mas que ainda não consegue resolver coisas básicas.

Os três exemplos da relação *Contrast* apresentam diferentes tipos de sinalizadores, de acordo com Dantas *et al.* (2024), desde os MDs (conjunções, preposições e advérbios) a outros SDs de diversos tipos. Como pode ser visto, existe um grande quantitativo de MD, sendo 13 conjunções *mas* das 15 ocorrências de MD, em que os outros dois foram os advérbios *não* e *nunca*. No caso das Concession, a variedade de MDs foi maior (*embora, mesmo com, apesar de, não, porém, dificilmente*). As conjunções em *Contrast* estão sempre combinadas com o sinalizador Gráfico, a *vírgula*, ainda que esse SD não tenha sido anotado em todas as ocorrências do *mas*, apenas em 9 de 13. Depois dos MDs, o tipo Semântico é o mais recorrente, se apresentando, principalmente, por meio da Antonímia.

4.4 Reflexões e contribuições à anotação dos sinalizadores

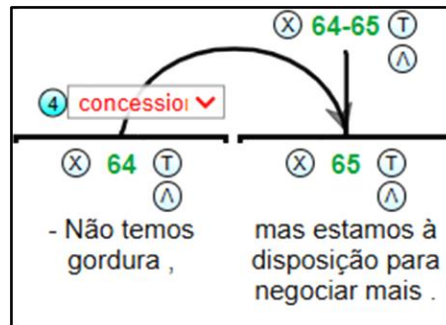
As relações *Antithesis*, *Contrast* e *Concession* compartilham uma mesma característica, que é a ideia de oposição existente em suas construções. Desse modo, um dos objetivos deste trabalho é mostrar a diferença entre elas, já que tratam do mesmo fenômeno.

Na seção 4, da análise dos dados, ficou evidente que não apenas os MDs são usados na identificação da relação, do sentido que ela carrega, apesar de ser muito frequente, outros sinalizadores também podem determiná-la. Entretanto, há tipos de sinalizadores que podem ser compartilhados por uma variedade de relações, como o Morfológico, o Sintático, e o Gráfico, e outros tipos e subtipos que são mais específicos, como os MDs, o Semântico, como a Antonímia. Os resultados mostraram que os sinalizadores mais específicos tiveram mais recorrência e um peso maior na identificação dessas relações, que nesse caso, são as de oposição. Outro ponto sobre os MDs, mais especificamente as conjunções, é que, apesar de o *mas* ter sido o conectivo mais abundante nas três relações, há conectivos concessivos como *apesar de*, *embora* e *mesmo com* que só apareceram nas relações *Concession*, claro que essa é uma observação a partir dos dados desse *subcorpus*, não é uma informação taxativa.

Uma vez que o *mas* foi um ponto crucial para o desenvolvimento desta pesquisa e se mostrou muito importante também no resultado dela, é válido discorrer um pouco mais sobre esse marcador.

A conjunção *mas*, segundo autores como Bechara (2009 [1928]) e Lima (2011), é uma conjunção adversativa por excelência. Apesar disso, Lima (2011) e Neves (1999) também defendem que ela pode aparecer em casos de concessivas. Para caracterizar a função do *mas* nesses dois tipos de orações, Neves (1999) fez uma correlação entre elas, conforme descrito na subseção 2.1.2. A autora apresenta o fator argumentativo como base para essa diferença; enquanto o *mas* na concessão refuta uma objeção, na adversativa, ele admite um fato. Essa argumentação de Neves (1999) vai ao encontro da descrição em RST acerca das relações *Concession* e *Contrast*, enquanto na primeira há um efeito pretendido no leitor (Figura 23), na segunda há apenas o reconhecimento do fato (Figura 24).

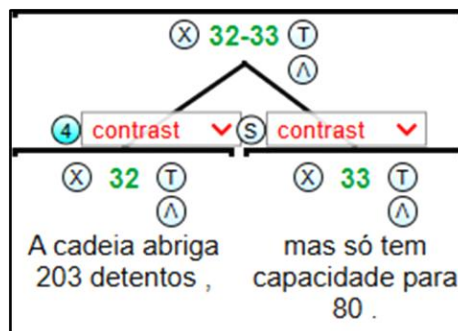
Figura 23 - Mas na relação *Concession*



Fonte: Retirado do corpus CSTNews.

Nesse exemplo, do texto D3_C50_OGlobo, da relação *Concession*, há um argumento usado por parte do escritor/falante ao seu leitor/ouvinte, de que, mesmo não tendo gordura para negociar, há a disposição para negociar outras coisas, ou seja, há uma tentativa de convencimento.

Figura 24 - Mas na relação *Contrast*.



Fonte: Retirado do corpus CSTNews.

Já nesse exemplo, do texto D1_C37_OGlobo, da relação *Contrast*, a oposição é feita entre duas informações que estão apenas apresentando um fato, não há um argumento ou tentativa de convencimento para que a informação seja aceita por parte do leitor/ouvinte, como é característico de notícias jornalísticas.

Na Figura 14, texto D3_C12_GPovo, da relação *Antithesis*, anteriormente comentada na seção 4.1, o *mas* introduz um argumento que se opõe à informação de N. A partir da compreensão desse argumento expresso em S, o leitor tem a tendência de aceitar melhor N.

Esses e outros casos em que o MD *mas* aparece nas relações de oposição evidenciam que essa conjunção cumpre com sua proposta básica de contraste, por isso ela aparece nas três

relações. No entanto, o contexto em que ela for empregada é que vai dizer se se trata de uma relação *Contrast*, *Concession* ou *Antithesis*.

Em relação à anotação dos sinalizadores, baseados em Dantas *et al.* (2024), seria mais seguro abordar a anotação única depois da concordância, assim veria se os pontos de falhas encontrados foram solucionados. Por outro lado, dá a oportunidade de os anotadores e pesquisadores revisarem o material produzido.

Durante a análise da anotação, a única categoria que despertou confusão no entendimento tanto do anotador como da adjudicadora foi o Sentido do Verbo, uma vez que a descrição feita no Manual de Anotação não deixa clara a diferença entre ele e o subtipo Tempo Verbal. Sendo assim, a descrição feita precisa ser reformulada para deixar mais clara a sua função na taxonomia dos sinalizadores.

Como visto em dois momentos da análise, alguns textos foram anotados por apenas um dos três anotadores, como foram os casos de *D5_C20_Folha*, na *Concession*, e *D1_C29_Folha*, na *Contrast*. Mas além desses, há os textos *D2_C16_ Estadão*, *D1_C27_Folha* que foram anotados por apenas um anotador. Há também 13 textos que um dos três anotadores não fizeram a anotação. É compreensível que a tarefa é lenta e cansativa, demanda muito do anotador para a análise, por isso podem ter passado despercebido. De todo modo, é um ponto muito importante a ser revisto, para que realmente possa ser obtida a concordância da anotação e o material final ser confiável para a reprodução e ensinamento da máquina. Além disso, um adjudicador para a etapa final é uma sugestão válida, alguém que não esteja envolvido na anotação, mas só na revisão. Assim, sinalizadores que passaram despercebidos, podem ser identificados e discutidos, como feito em alguns momentos da análise desta pesquisa, que, inclusive, podem ser considerados para uma nova anotação. No mais, houve algumas discordâncias menores na catalogação dos sinalizadores, como em *apesar de*, que um anotador colocou como conjunção e outro colocou como preposição; e outros casos que podem ser visualizados na planilha.

5. CONSIDERAÇÕES FINAIS

Nesta pesquisa, a posição em relações RST foi o objeto de estudo, na qual se construiu um arcabouço teórico acerca do fenômeno, desde a teoria RST a conceitos gramaticais e linguísticos sobre ele. Esse conhecimento teórico serviu para caracterizar as relações *Antithesis*, *Concession* e *Contrast*, entender sua estrutura e os sinalizadores que lhes dão sentido.

Na seção 2, foi construída uma fundamentação teórica que buscou relacionar a teoria RST com conceitos da gramática, respeitando seus pontos de encontro e desencontro. Uma vez que, como dito, as proposições das relações RST se aproximarem às orações gramaticais em sua estrutura e também no uso das conjunções, um dos principais marcadores discursivos das relações de oposição, essa associação foi feita, entendendo que são teorias muito distintas.

Na subseção 2.1.2, houve uma progressão do que dizem as gramáticas mais tradicionais (Bechara (2009 [1928]) e Cunha e Cintra (2017 [1985]), Lima (2011)), que consideram a estrutura da Língua, e gramáticas funcionalistas, como a de Neves (1999) e autores que também caminham pela funcionalidade da Língua, como apresentado no trabalho de Margarido (2010) em que os primeiros enquadram as orações e suas conjunções respectivas, aos passo que os últimos consideram as construções reais dos falantes na Língua corrente. Essa diferença de percepção mostra que o falante entende o conceito de oposição e faz uso das estruturas, sejam adversativas ou concessivas, contudo não regulam que cada construção (adversativa e concessiva) tem conjunções próprias, mas fazem uso delas de forma mesclada, como apresentou Margarido (2010). Com isso, a RST, que considera o discurso, apresenta, em suas relações de oposição, uma mescla das conjunções nos três tipos de relações.

Ao longo deste trabalho, buscou-se discutir três questões da pesquisa. A primeira é: *Fazendo parte do mesmo fenômeno de Oposição, o que caracteriza as relações Antithesis, Contrast e Concession?* Viu-se que a RST em si não apresenta tantos detalhes sobre a diferença entre elas, mas a informação dada corrobora o que diz Neves (1999) sobre a diferença entre orações adversativas e concessivas, que o fator argumentativo é o que as diferencia, uma vez que nas adversativas apenas se apresenta versões opostas, ao passo que nas concessivas há uma forma argumentativa de convencimento sobre o que se apresenta. A RST também apresenta o efeito pretendido no leitor/ouvinte como um fator crucial que as caracteriza na categoria de relações mononucleares ou multinucleares. Na *Contrast*, os contrastes são identificados, comparados e aceitos. Na *Concesion*, os contrastes são identificados, mas o efeito pretendido é que o leitor/ouvinte aceite melhor o núcleo, o mesmo acontece com a *Antithesis*. A teoria não apresenta mais detalhes de como diferenciar essas relações, como quais sinalizadores marcam

cada uma delas. Pardo (2005, p. 143) apresenta um apêndice com marcadores superficiais de cada relação, e o resultado encontrado é que esses marcadores são, em sua grande maioria, conjunções adversativas, concessivas, preposições ou locuções prepositivas que indicam oposição, o que aproxima da gramática, bem como a divisão dos núcleos e da dependência entre orações. Contudo, tanto na RST quanto na gramática, trata-se de estrutura estabelecida, o que pode se mostrar diferente na língua em uso.

A segunda questão de pesquisa é: *Sendo possível indicar mais de uma relação, o que caracteriza a conjunção mas?* Margarido (2010), indo além do que dizem os gramáticos sobre a excelência da conjunção *mas* nas adversativas, defende que esse conectivo tanto pode expressar uma oposição entre dois conteúdos, como uma quebra de expectativa criada pela primeira proposição, o que lhe concede valor concessivo. Em todas as ocorrências analisadas, o *mas* aparece combinado com a vírgula, isso se dá pelo fato de conjunções adversativas serem antecedidas por vírgula. Apesar de ser apontado pela gramática como uma conjunção adversativa, a análise do *subcorpus* sobre esse MD mostra que ele pode ser empregado em qualquer uma das três relações, a depender do contexto aplicado, tanto pode apenas apresentar ideias contrárias, como pode exercer função argumentativa. Desse modo, é necessário olhar para além do SD para a identificação da relação retórica.

A necessidade de caracterizar as relações retóricas para além dos SDs levanta uma questão que já foi apresentada aqui, tanto a RST quanto a gramática apresentam uma estrutura estabelecida; no entanto, as relações, ao serem representadas no discurso, não apresentam uma diferença estrutural tão bem delimitada assim, pois o falante não considera de forma intencional a estrutura dessas teorias ao fazer uso da linguagem. Um exemplo são as relações *Antithesis* e *Contrast*, que, estruturalmente, são distintas, uma é mononuclear e a outra é multinuclear, mas no discurso, essa diferença não é perceptível.

A terceira questão de pesquisa é sobre os demais sinalizadores atuantes nessas relações. *Quais outros sinalizadores são usados para marcar essas relações?* O resultado da análise mostrou que os marcadores discursivos são predominantes, pois aparecem com mais frequência e se mostram como principal sinalizador opositor, principalmente a conjunção *mas*. Entretanto, outros sinalizadores como a pontuação, a antonímia, o conhecimento de mundo, os tempos verbais, entre outros, têm forte papel na identificação da oposição das relações *Antithesis*, *Concession* e *Contrast*.

O *subcorpus* analisado ainda é pequeno, nem metade do quantitativo de relações de oposição do *corpus*, há mais para ser investigado e desenvolvido, como a anotação dos

sinalizadores nas relações interssentenciais, observando se há detalhes característicos nas relações entre sentenças como foi percebida a função contrastiva da *vírgula* em relações mononucleares intrassentenciais. A relação *Antithesis* também precisa ser mais investigada, inclusive em outros *corpora*, já que o quantitativo dela é baixo no CSTNews com apenas 8 ocorrências no *corpus* inteiro.

O desenvolvimento dessa pesquisa contribui para o avanço dos estudos da RST, proporcionando análises mais detalhadas e específicas sobre essas relações para serem copiadas, reproduzidas e aplicadas a outros contextos, colaborando com o objetivo do POeTISA de produzir recursos linguísticos do PB aplicáveis ao PLN. Também colabora para o desenvolvimento do projeto *RST para além dos marcadores discursivos*, que busca avançar os estudos da RST com outros tipos de sinalizadores, não apenas os marcadores discursivos. Além disso, o projeto busca não só investigar os sinalizadores, mas aplicá-los em anotações manuais, como feito por Dantas *et. al* (2024) e continuará sendo feito pelo grupo de pesquisadores, com base, também, nesses dados apresentados.

Na apresentação deste trabalho, uma das motivações para investigar a RST foi o vínculo empregatício como corretora de redações ENEM. A competência 3 trata da capacidade do estudante de selecionar, interpretar e relacionar informações para defesa de uma tese. A competência 4 aborda o uso adequado e estratégico dos conectivos para um bom encadeamento de ideias e argumentos. Nessas competências, é possível perceber como as relações retóricas estão aplicadas, tanto que já existem estudos (Filho e Nogueira, 2020; Cândido e Webber, 2021) específicos e ferramentas privadas de correção automática para esses textos. Essa é mais uma das aplicações da RST que mostra a importância dessa teoria para os estudos linguísticos aplicados ao ensino e ao PLN.

REFERÊNCIAS

- ANTONIO, Juliano Desiderato. Expressão da relação retórica de concessão em elocuções formais e entrevistas orais. *Revista de estudos da linguagem*. V. 19, n. 1, p. 24, 2011. DOI:10.17851/2237-2083.19.1.143-166. Disponível em: <http://periodicos.letras.ufmg.br/index.php/relin/article/view/2556>. Acesso em: 25 set. 2023.
- ANTONIO, Juliano Desiderato. Mecanismos utilizados pelos destinatários do discurso para identificação de relações de coerência não sinalizadas por conectores. *DELTA. Documentação de Estudos em Linguística Teórica e Aplicada (Online)*, v. 33, p. 79-108, 2017. DOI:10.1590/0102-445025798334674077. Disponível em: <https://www.researchgate.net/publication/315588736/>. Acesso em: 20 mai. 2023.
- AZEREDO, José Carlos de. **Gramática Houaiss da língua portuguesa**. São Paulo: Publifolha. 2008.
- BECHARA, Evanildo. **Moderna gramática portuguesa**. Rio de Janeiro: Nova Fronteira, 2009.
- CÂNDIDO, T. G.; WEBBER, C. J.; Avaliação da Coesão Textual: Desafios para Automatizar a Correção de Redações. *Renote*, Porto Alegre, v. 16, n. 1, 2018. DOI: 10.22456/1679-1916.86013. Disponível em: <https://seer.ufrgs.br/index.php/renote/article/view/86013>. Acesso em: 29 abr. 2025.
- CARDOSO, P. C. F.; MAZIERO, E. G.; JORGE, M. L. R. C.; SENO, E. M. R.; FELIPPO, A.; RINO, L. H. M.; NUNES, M. G. V.; PARDO, T. A. S.; **CSTNews - A Discourse-Annotated 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, 2011.
- CASELI, Helena de Medeiros; NUNES, Maria das Graças Volpe; PAGANO, Adriana. **O que é PNL?**. In: *Processamento de Linguagem Natural: conceitos, técnicas e aplicações em português*. p. 8-13. 2010. Disponível em: https://brasileiraspln.com/livro-pln/1a-edicao/Livro_PLN.pdf. Acesso em: 05 jan. 2024.
- CEGALLA, Domingos Paschoal. **Novíssima gramática da língua portuguesa**. Companhia Editora Nacional, 2010.
- CRUZ, G. S. B.; SOUZA, J. W. C.; CARDOSO, P. C. F.. **Estratégias Automáticas para Análise da Concordância da Anotação de Sinalizadores Discursivos**. In: SIMPÓSIO BRASILEIRO DE TECNOLOGIA DA INFORMAÇÃO E DA LINGUAGEM HUMANA (STIL), 15, 2024, Belém/PA. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2024. p. 440-444. DOI: <https://doi.org/10.5753/stil.2024.245159>.
- CUNHA, Celso; CINTRA, Lindley. **Nova gramática do português contemporâneo**. Rio de Janeiro: Lexikon, 2017. 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.; CARDOSO, P. C. F.; **Manual de anotação de sinalizadores discursivos**. Série de Relatórios Técnicos do NILC. 2024.
- EGG, Markus; DAS, Debopam. **Signalling conditional relations**. *Linguistics Vanguard*, vol. 8, no. s4, 2022, p. 383-392. DOI: 10.1515/lingvan-2021-0027. Disponível em: <https://www.degruyter.com/document/doi/10.1515/lingvan-2021-0027/>. Acesso em: 22 mar. 2024.

EMÍLIO, Aline C. K. **Antítese: do discurso jornalístico a um sentido amplo**. 1999. p.53-63. Instituto Brasileiro de Informação em Ciência e Tecnologia. 1999. Disponível em: <https://periodicos.ufsc.br/index.php/workingpapers/article/view/2310>. Acesso em: 12 abr. 2025.

FACHADA, Beatriz. *Mas em artigos de opinião: valores e relações retóricas*. ElingUP: Revista eletrônica de linguística dos Estudantes da Universidade de Porto. v. 8, n. 1, 2019. Disponível em: <https://ojs.letras.up.pt/index.php/elingUP/article/view/6406>. Acesso em: 25 set. 2023.

FILHO, L. A. C.; NOGUEIRA, M. T.; A Estrutura Retórica De Uma Redação Do Enem: Uma Análise Funcionalista. Revista Con (textos) Linguísticos, v. 14, n. 28. 2020. Disponível em: <https://periodicos.ufes.br/contextoslinguisticos/article/view/31039>. Acesso em: 28 abr. 2025.

FIORIN, José Luiz. As Figuras De Pensamento: Estratégia Do Enunciador Para Persuadir O Enunciatário. ALFA: Revista de Linguística, v. 32, p. 1-15. 1988. Disponível em: <http://hdl.handle.net/11449/107633>. Acesso em: 11 abr. 2025.

GESSLER, Luke; LIU, Yang Janet; ZELDES, Amir. **A discourse signal annotation system for RST trees**. In: Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019. 2019. p. 56-61.

HOU, Shengluan; ZHANG, Shuhan; FEI, Chaoqun. Rhetorical structure theory: A comprehensive review of theory, parsing methods and applications. Expert Systems with Applications, v. 157, p. 113421, 2020.

HOVY, Eduard; LAVID, Julia. **Towards a ‘Science’ of Corpus Annotation: A New Methodological Challenge for Corpus Linguistics**. International Journal Of Translation v. 22, n. 1, 2010.

KOCH, Ingedore Grunfeld Villaça; TRAVAGLIA, Luiz Carlos. **Texto e Coerência**. São Paulo: Cortez, 2005.

LIMA, Rocha. **Gramática normativa da língua portuguesa**. Rio de janeiro: José Olympio, 2011.

LIU, Yang; ZELDES, Amir. **Discourse relations and signaling information: Anchoring discourse signals in RST-DT**. Society for Computation in Linguistics, v. 2, n. 1, 2019.

MANN, William C.; THOMPSON, Sandra A. **Rhetorical structure theory: A theory of text organization**. Los Angeles: University of Southern California, Information Sciences Institute, 1987.

MANN, Willian C. TABOADA, Maite. RSTWeb, 2010. Disponível em: <https://www.sfu.ca/rst/03research/index.html>. Acesso em 20 out. de 2023.

MARGARIDO, Renata. **Construções (coordenadas) adversativas e construções (subordinadas) adverbiais concessivas em português: pontos de contato e contraste na língua em função**. 2010. p.193. Dissertação (Mestrado em Estudos Linguísticos). Universidade Presbiteriana Mackenzie, São Paulo, 2010. Disponível em: <https://dspace.mackenzie.br/items/ed4ec197-cc80-4f56-8898-04b7cfe95b19>. Acesso em: 30 jan. 2024.

NEVES, Maria Helena de Moura. **Gramática do Português Falado**. São Paulo: Humanitas/FFLCH/USP; Campinas: Editora da Unicamp. V. 7, 1999.

PARDO, T. A. S. **Métodos para análise discursiva automática**. 2025. p. 211. Tese de Doutorado. São Carlos: Universidade de São Paulo, 2005. DOI: 10.11606/T.55.2005.tde-29082005-172336. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-29082005-172336/pt-br.php/>. Acesso em: 12 mai. 2023.

PECUCH, Gabriele. A relação retórica de Elaboração sinalizada pelo marcador discursivo *mas* em aulas e em entrevistas orais. *Letras Escreve*. Macapá. v. 11, n. 1, 2021. Disponível em: <https://periodicos.unifap.br/index.php/letras>. Acesso em: 25 set. 2023.

RODRIGUES, Roana; SOUZA, Jackson Wilke da Cruz; CARDOSO, Paula Christina Figueira. **Sinalizadores retórico-discursivos: revisitando a anotação RST no corpus CSTNews**. In: Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, 2023, Brasil. Anais do XIV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL 2023). p. 249.

SOUZA, Jackson Wilke da Cruz; CARDOSO, Paula Christina Figueira; RODRIGUES, Roana. Systematic Review of Studies on Rhetorical Structure Theory (RST). **REVISTA DE ESTUDOS DA LINGUAGEM**, v. 31, n. 3, p. 1643-1675. DOI: <http://dx.doi.org/10.17851/2237-2083.31.3.1643-1675>

TABOADA, Maite; MANN, William C. Rhetorical Structure Theory: looking back and moving ahead. *Discourse studies*, v. 8, n. 3, p. 423-459, 2006. DOI: <https://doi.org/10.1177/1461445606061881>. Disponível em: <https://journals.sagepub.com/doi/10.1177/1461445606061881/>. Acesso em: 28 jun. 2023

APÊNDICES

Tabela 2 - Planilha com dados da relação *Antithesis*

Doc. de origem	Porção de texto	Qnt. Total de EDUs no texto	Ocor. da Antithesis na EDU	Altura da ocor. de Antithesis na árv	Doc. de origem	Anotador 1	Anotador 2	Anotador 3	Adjudicadora	Tipo	Subtipo	Combinados
D3_C12_GPovo	“A Coreia do Norte recusou ofertas de agências internacionais para o lançamento de campanhas de ajuda ao país, mas uma autoridade local disse na semana passada que Pyongyang aceitaria ajuda da Coreia do Sul se fosse dada sem condições. A Coreia do Sul suspendeu o envio de alimentos ao norte por causa de suas preocupações sobre a	35	31 - 35	3 de 8	D3_C12_GPovo	–	–	recusou	recusou	Morfológico	Sentido verbal	MD+GRAF+MORF+SEM+MORF
					D3_C12_GPovo	–	–	mas	mas	Marcador Discursivo	Conjunção	

	paralisação nas negociações sobre o programa nuclear norte-coreano e os recentes testes											
	de mísseis											
	realizados por Pyongyang.”											
D3_C50_Oglobo	"que teria se comprometido a elevar o limite de isenção da CPMF para quem tem renda até R\$ 4.340, em vez dos R\$ 1.642 negociados até então"	84	6 a 7	10 de 14	D3_C50_Oglobo	em vez	em vez dos	em vez dos	em vez dos	Marcador Discursivo	Preposição	IND

Tabela 3 - Planilha com os dados da relação *Concession*

Doc. de origem	Porção de texto	Qnt. Total de EDUs	Ocor. da Concession na EDU	Alt. da ocor. da Concession na árv	Doc. de origem	Anotador 1	Anotador 2	Anotador 3	Adjudicadora	Tipo	Subtipo	Combinados
D3_C4_O Globo	“Amanhã, a temperatura máxima pode chegar a 22 graus, mas na quarta volta a cair”	37	12	9 de 9	D3_C4_O Globo	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SIN
					D3_C4_O Globo	–	Amanhã, a temperatura máxima pode chegar a 22 graus	–	Amanhã, a temperatura máxima pode chegar a 22 graus	Sintático	Oração Circunstancial	
D1_C7_E stadão	“O mesmo ocorre com o sexto das añas marrons, corpos pesados o suficiente para gerar calor, mas não o bastante para iniciar uma fusão nuclear como as estrelas legítimas”	38	18/jan	8 de 10	D1_C7_E stadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM
					D1_C7_E stadão	–	–	não o bastante	suficiente/ não o bastante	Semântico	Antonímia	
D1_C7_E stadão	“Ambos os mundos têm massa semelhante à de outros exoplanetas já catalogados, mas não giram	38	7	7 de 10	D1_C7_E stadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MD+SEM
					D1_C7_E stadão	–	–	–	não	Marcador Discursivo	Advérbio	
					D1_C7_E stadão	–	–	não giram em torno	giram em torno de uma estrela	Semântico	Conh. Mundo	

	em torno de uma estrela”							de uma estrela				
D3_C8_J B	"Embalado pela vitória, o Brasil começou o segundo set na frente, mas permitiu a reação dos donos da casa, que chegaram a empatar a parcial em 17"	25	17 - 20	4 de 6	D3_C8_J B	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM
						–	permitiu a reação dos donos da casa/ chegaram a empatar a parcial em 17	o Brasil começou o segundo set na frente/ donos da casa/ empatar a parcial	o Brasil começou o segundo set na frente/ donos da casa/ empatar a parcial	Semântico	Conh. Mundo	
D3_C8_J B	"Na última parcial, a Finlândia novamente emparelhou o jogo com o Brasil até a metade do set, mas depois de uma sequência de pontos brasileiros não conseguiu reagir e perdeu por 25 a 21"	25	23 - 25	3 de 6	D3_C8_J B	–	a Finlândia novamente e emparelhou o jogo com o Brasil	emparelhou o jogo/ perdeu	emparelhou o jogo/ perdeu	Semântico	Conh. Mundo	MD+GRAF+SEM
						mas	–	mas	mas	Marcador Discursivo	Conjunção	
D2_C16_ Estadão	"Ex-líder do PP, Pedro Henry (MT) cogitou sair da função, mas teria desistido da ideia"	78	21 - 24	9 de 11	D2_C16_ Estadão			teria desistido	teria desistido	Morfológico	Tempo-verbal	MD+GRAF+MOR F
					D2_C16_ Estadão			mas	mas	Marcador Discursivo	Conjunção	

D1_C17_Folha	"Sem citar nominalmente o adversário, Alckmin criticou de novo Lula ao comentar especulações de que o petista, condenado na vitória no primeiro turno, já estaria fazendo planos sobre sua nova equipe ministerial"	77	23 - 25	7 de 13	D1_C17_Folha	Sem citar nominalmente o adversário	Sem citar	–	Sem citar	Sintático	Oração Circunstancial	IND
D1_C17_Folha	"Não sou nenhum expert, mas gosto de dançar."	77	72 - 73	5 de 13	D1_C17_Folha	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF
D2_C18_Estadão	"nas últimas duas semanas houveram ameaças de bomba no local, mas ainda não se sabe se elas têm alguma relação com o tiroteio"	85	69 - 70	5 de 11	D2_C18_Estadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM
					D2_C18_Estadão	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D2_C18_Estadão	–	–	ameaças de bomba/tiroteio	ameaças de bomba/tiroteio	Semântico	Conh. Mundo	
D2_C19_Estadão	"Mesmo com esta melhora, ele continuará internado"	25	09 e 10	7 de	D2_C19_Estadão	mesmo	mesmo	Mesmo com	mesmo com	Marcador Discursivo	Conjunção	IND

D5_C20_ Gpovo	"Sabemos que não temos chance de ganhar [a votação], quanto mais levar aqui, mais difícil fica para o governo no Senado"	62	47 - 50	6 de 9	D5_C20_ Gpovo	–	sabemos que não temos chance	–	sabemos que não temos chance	Sintático	Val. Verbal	IND
D5_C20_ Gpovo	"Apesar de 54 parlamentares inscritos, cerca de dez deputados estão neste momento fazendo discursos em plenário"	62	53 - 54	4 de 9	D5_C20_ Gpovo	–	Apesar de	–	Apesar de	Marcador Discursivo	Conjunção	MD+SIN
						–	54/dez	–	54/dez	Sintático	Numeral	
D2_C21_ Oglobo	"Mesmo com esta melhora, ele continuará internado"	25	09 a 10	7 de 7	D2_C21_ Oglobo	mesmo com	mesmo	mesmo com	mesmo com	Marcador Discursivo	Conjunção	MD+SEM
					D2_C21_ Oglobo	–	com esta melhora	–	com esta melhora/internado	Semântico	Campo-Semântico	
D4_C22_ JB	"O aeroporto Internacional Tom Jobim, na Ilha do Governador, Zona Norte do Rio, está aberto para pouso e decolagens, mas operando por instrumentos na manhã desta terça-feira"	34	20 - 21	4 de 5	D4_C22_ JB	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF

D2_C24_ Estádio	"já que bateu a marca de 4,40m, mesmo com as vaias da torcida local "	37	27 - 28	4 de 5	D2_C24_ Estádio	vírgula	vírgula	vírgula	vírgula	Gráfico	Pontuação	MD+GRAF+SIN
					D2_C24_ Estádio	–	mesmo	mesmo com	mesmo com	Marcador Discursivo	Conjunção	
					D2_C24_ Estádio	mesmo com as vaias da torcida local	–	as vaias da torcida local	as vaias da torcida local	Sintático	Oração Circunstancial	
D2_C24_ Estádio	"Fabiana Murer, no entanto, não parecia incomodada, mas errou sua primeira tentativa de alcançar 4,50 m"	37	29 - 30	3 de 5	D2_C24_ Estádio	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MOR F
					D2_C24_ Estádio	Vírgula	–	Vírgula	Vírgula	Gráfico	Pontuação	
					D2_C24_ Estádio	–	no entanto	–	–	Marcador Discursivo	Conjunção	
					D2_C24_ Estádio	–	–	errou	errou	Morfológico	Sentido verbal	
D1_C27_ Folha	"o Brasil chegou a ser vaiado pelo torcedor, mas deslanchou no fim"	65	2 a 3	6 de 9	D1_C27_ Folha	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF
					D1_C27_ Folha	Vírgula	–	–	vírgula	Gráfico	Pontuação	
D1_C27_ Folha	"A equipe tocava a bola no meio-campo, mas não atacava com muito perigo "	65	12 a 13	7 de 9	D1_C27_ Folha	mas	mas	mas	mas	Marcador discursivo	Conjunção	MD+GRAF+MD+ SEM
					D1_C27_ Folha	Vírgula	–	–	vírgula	Gráfico	Pontuação	
					D1_C27_ Folha	–	–	não	não	Marcador discursivo	Advérbio	
					D1_C27_ Folha	–	–	não atacava	não atacava	Semântico	Conh. Mundo	
D1_C27_ Folha	"Robinho, Ronaldinho e	65	27 - 28	5 de 9	D1_C27_ Folha	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MD

	Kaká se esforçaram, mas não criaram muitas jogadas"				D1_C27_Folha	–	–	não	não	Marcador Discursivo	Advérbio	
D1_C27_Folha	"Na etapa final, o treinador equatoriano colocou o centroavante Tenorio no lugar do meio-campista Quiróz. A substituição tornou o time mais ofensivo, porém os meias brasileiros tiveram mais espaços para criar."	65	29 - 31	4 de 9	D1_C27_Folha	–	–	porém	porém	Marcador Discursivo	Conjunção	MD+GRAF+SEM
					D1_C27_Folha	–	–	time mais ofensivo/meias/ espaços para criar	time mais ofensivo/meias/ espaços para criar	Semântico	Campo-semântico	
D1_C27_Folha	"Castillo bateu forte, mas a bola passou por cima do gol"	65	33 - 34	7 de 9	D1_C27_Folha	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM
					D1_C27_Folha	–	–	a bola passou por cima do gol	a bola passou por cima do gol	Semântico	Conh. Mundo	
D1_C27_Folha	"Mesmo com a decisão de Dunga de trocar Love pelo meio-campista Elano, a equipe não perdeu o poder ofensivo"	65	52 - 53	6 de 9	D1_C27_Folha	–	mesmo com	mesmo com	mesmo com	Marcador Discursivo	Conjunção	MD+GRAF+SIN+MD
					D1_C27_Folha	Mesmo com a decisão de Dunga de trocar Love pelo meio-	–	–	Mesmo com a decisão de Dunga de trocar Love pelo meio-campista Elano	Sintático	Oração Circunstancial	

						campista Elano						
					D1_C27_Folha	vírgula	–	–	vírgula	Gráfico	Pontuação	
					D1_C27_Folha	–	–	não	não	Marcador Discursivo	Advérbio	
D1_C27_Folha	"A bola foi fraca, mas o goleiro Daniel Viteri falhou"	65	59 - 60	6 de 9	D1_C27_Folha	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MOR F
					D1_C27_Folha	vírgula	–	–	vírgula	Gráfico	Pontuação	
					D1_C27_Folha	–	–	falhou	falhou	Morfológico	Sentido verbal	
D1_C29_Folha	"O acerto divulgado evita que Mahony seja chamado a depor e explicar se a igreja tinha conhecimento dos abusos - cometidos entre as décadas de 40 e 90-, porém exige que documentos e arquivos sejam revelados"		23 - 26	7 de 11	D1_C29_Folha	–	porém	porém	porém	Marcador Discursivo	Conjunção	MD+GRAF
D2_C33_Estação	"O problema da fome no planeta não	105	12 a 14	8 de 11	D2_C33_Estação	–	–	não	não	Marcador Discursivo	Advérbio	MD+GRAF+MD

	decorre da falta de alimentos, mas da falta de renda que golpeia quase um bilhão de homens, mulheres e crianças"				D2_C33_Estadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	
D1_C35_Folha	"Em 1996, quando Abadia foi entregue à polícia colombiana, os Estados Unidos pediram a extradição dele pelo "envolvimento presumível" com o cartel, mas não foram atendidos"	49	15 - 18	6 de 8	D1_C35_Folha	–	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM
					D1_C35_Folha	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D1_C35_Folha	–	não foram atendidos	–	não foram atendidos	Semântico	Campo-semântico	
D1_C35_Folha	"embora tenha sido condenado a 23 anos de prisão, foi solto em 2002"	49	24 - 25	8 de 8	D1_C35_Folha	embora	embora	embora	embora	Marcador Discursivo	Conjunção	MD+GRAF
					D1_C35_Folha	–	–	vírgula	vírgula	Gráfico	Pontuação	
D4_C36_JB	"Tal capacidade de mutação fez escola, mas dificilmente as criaturas saberão superar o criador"	86	16 - 17	6 de 10	D4_C36_JB	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MD
					D4_C36_JB	–	vírgula	–	vírgula	Gráfico	Pontuação	
					D4_C36_JB	–	–	dificilmente	dificilmente	Marcador Discursivo	Advérbio	

D4_C36_JB	"Governista por vocação, ensaiou uma aliança até com o presidente Luiz Inácio Lula da Silva, mas o namoro político não resistiu aos primeiros meses do governo petista"	86	18 - 19	6 de 7	D4_C36_JB	–	vírgula	–	vírgula	Gráfico	Pontuação	MD+GRAF
					D4_C36_JB	mas	mas	mas	mas	Marcador Discursivo	Conjunção	
D4_C36_JB	"Antonio Carlos Peixoto de Magalhães formou-se médico, mas encontrou na política a sua profissão"	86	21 - 22	5 de 7	D4_C36_JB	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF
					D4_C36_JB	–	vírgula	vírgula	vírgula	Gráfico	Pontuação	
D5_C41_GPovo	"O brasileiro dominou quase toda a prova, mas, nos metros finais, perdeu o ritmo e acabou apenas na quarta colocação, com o tempo de 1m49s49"	60	20 - 22	6 de 7	D5_C41_GPovo	–	–	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM+MORF
					D5_C41_GPovo	dominou quase toda prova/ perdeu o ritmo/ acabou apenas na quarta colocação	–	–	dominou quase toda prova/ perdeu o ritmo/ acabou apenas na quarta colocação	Semântico	Campo-semântico	
					D5_C41_GPovo	–	–	acabou	acabou	Morfológico	Sentido verbal	
D2_C42_Estação	"Sobrestar não é arquivar, é	79	20 - 21	9 de 12	D2_C42_Estação	não é	–	não	não	Marcador Discursivo	Advérbio	MD+GRAF+SIN+SEM

	retomar os trabalhos mais à frente"				D2_C42_Estadão	–	é retomar os trabalhos mais à frente	–	é retomar os trabalhos mais à frente	Sintático	Oração Circunstancial	
					D2_C42_Estadão	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D2_C42_Estadão	–	–	é arquivar/ é retomar	é arquivar/ é retomar	Semântico	Antonímia	
D2_C42_Estadão	"Depois de a Schincariol ter comprado uma fábrica de refrigerantes da família Calheiros por R\$ 27 milhões, quando não valeria mais do que R\$ 10 milhões, segundo avaliações de mercado"	79	69 - 70	10 de 12	D2_C42_Estadão	quando não	quando	–	quando	Marcador Discursivo	Advérbio	MD+GRAF+SIN+MORF+MD
					D2_C42_Estadão	–	–	R\$ 27 milhões/ R\$ 10 milhões	R\$ 27 milhões/ R\$ 10 milhões	Sintático	Numeral	
					D2_C42_Estadão	valeria	–	–	valeria	Morfológico	Tempo-verbal	
					D2_C42_Estadão	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D2_C42_Estadão	mais do que	–	–	mais do que	Marcador Discursivo	Advérbio	
D1_C44_Folha	"mesmo se a unificação dos processos para aprovação pelo conselho, vai defender relatores diferentes para cada uma das denúncias contra Renan"	65	37 - 38	6 de 10	D1_C44_Folha	mesmo se	–	mesmo se	mesmo se	Marcador Discursivo	Conjunção	MD+SEM+MORF
					D1_C44_Folha	–	unificação /diferentes	–	unificação /diferentes	Semântico	Antonímia	
					D1_C44_Folha	–	–	for/vai	for/vai	Morfológico	Sentido verbal	

D3_C45_ Oglobo	"Não estou reclamando como o apresentador Luciano Huck, mas como um cidadão que vê o Brasil indo tão bem em tantas áreas - índices nunca antes alcançados, capital estrangeiro entrando, notoriedade no mundo todo - e muito mal em questões básicas que a gente não consegue resolver, como a questão da violência"	71	64 - 71	6 de 10	D3_C45_ Oglobo	não	—	não	não	Marcador Discursivo	Advérbio	MD+GRAF
					D3_C45_ Oglobo	—	—	mas	mas	Marcador Discursivo	Conjunção	
D4_C46_ JB	"e os reatores de usinas nucleares em Niigata desligaram-se automaticamente para verificações, embora não haja relatos de vazamento de radiação"	30	6 a 8	7 de 8	D4_C46_ JB	—	embora	embora	embora	Marcador Discursivo	Conjunção	MD+SEM
					D4_C46_ JB	—	—	reatores de usinas nucleares/ radiação	reatores de usinas nucleares/ radiação	Semântico	Campo-semântico	

D4_C46_ JB	"Um incêndio em um transformador elétrico na usina nuclear de Kashiwazaki Kariwa rapidamente foi extinto, mas ainda não está claro quando a companhia elétrica de Tóquio vai religar três unidades no complexo"	30	24 - 25	6 de 8	D4_C46_ JB	mas	mas	mas ainda não	mas	Marcador Discursivo	Conjunção	MD+SEM
					D4_C46_ JB	–	–	transformador elétrico/ companhia elétrica/ religar	transformador elétrico/ companhia elétrica/ religar	Semântico	Campo-semântico	
D1_C48_ Folha	"Mesmo com altos e baixos, a equipe brasileira não teve muito trabalho para superar o fraco time do Canadá por 3 sets a 0 (25/18, 25/19 e 25/17), em uma hora e dez minutos"	70	23 - 26	5 de 12	D1_C48_ Folha	–	–	vírgula	vírgula	Gráfico	Pontuação	MD+GRAF
					D1_C48_ Folha	–	mesmo com	mesmo com	mesmo com	Marcador Discursivo	Preposição	
D1_C48_ Folha	"O Ricardinho faz falta, mas a responsabilidade foi passada para o Marcelinho e o Bruno, e a	70	39 - 41	7 de 12	D1_C48_ Folha	–	o Ricardinho faz falta	–	o Ricardinho faz falta	Sintático	Oração Circunstancial	MD+GRAF+SIN
					D1_C48_ Folha	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D1_C48_ Folha	–	–	mas	mas	Marcador Discursivo	Conjunção	

	torcida tem que apoiá-los"											
D3_C50_Oglobo	"Uma reunião da executiva do partido, que daria a palavra final sobre o acordo, estava marcada para a noite desta terça-feira, mas foi cancelada,"	84	29 - 32	7 de 14	D3_C50_Oglobo	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM
					D3_C50_Oglobo	–	–	marcada, noite de terça-feira/ cancelada	marcada, noite de terça-feira/ cancelada	Semântico	Campo-semântico	
D3_C50_Oglobo	"Não temos gordura, mas estamos à disposição para negociar mais"	84	64 - 65	5 de 14	D3_C50_Oglobo	–	–	Não	não	Marcador Discursivo	Advérbio	MD+GRAF+MD+SEM
					D3_C50_Oglobo	mas	mas	mas	mas	Marcador Discursivo	Conjunção	
					D3_C50_Oglobo	–	–	não temos gordura/ à disposição para negociar mais	não temos gordura/ à disposição para negociar mais	Semântico	Campo-semântico	

Fonte: Elaborado pela autora.

Tabela 4 - Planilha com dados da relação *Contrast*

Doc. de origem	Porção de texto	Qnt. Total de EDUs	Ocor. do Contraste na EDU	Alt. da ocor. de Contrast na árv	Doc. de origem	Anotador 1	Anotador 2	Anotador 3	Adjudicadora	Tipo	Subtipo	Combinados
D3_C13_Gpovo	“o grupo pretendia enviar uma equipe à região, mas foi impedido por soldados”	36	14 e 15	8 de 9	D3_C13_Gpovo	pretendia enviar/foi impedido	–	pretendia/foi	pretendia/foi	Morfológico	Tempo-verbal	MORF+MD+GRA F
					D3_C13_Gpovo	–	mas	mas	mas	Marcador Discursivo	Conjunção	
					D3_C13_Gpovo	–	vírgula	vírgula	vírgula	Gráfico	Pontuação	
D3_C13_GPovo	“Uma página na internet pró-Tigres Tâmeis responsabilizou o governo pelas mortes, mas o exército rejeitou a acusação”	36	20 e 21	4 de 9	D3_C13_Gpovo	Responsabilizou/rejeitou	–	responsabilizou/rejeitou	responsabilizou / rejeitou	Morfológico	Tempo-verbal	MORF+MD+GRA F+SEM
					D3_C13_GPovo	–	mas	mas	mas	Marcador Discursivo	Conjunção	
					D3_C13_GPovo	–	vírgula	vírgula	vírgula	Gráfico	Pontuação	
					D3_C13_GPovo	–	–	responsabilizou/rejeitou a acusação	responsabilizou / rejeitou a acusação	Semântico	Conh. Mundo	
D3_C13_GPovo	“O governo diz estar comprometido com a trégua, mas a situação política com os rebeldes, que lutam pela independência do norte e do leste do país permanece em um impasse”	36	31 - 34	3 de 9	D3_C13_Gpovo	Diz	–	–	Diz	Morfológico	Tempo-verbal	MORF+SEM+MD +GRAF
					D3_C13_Gpovo	–	–	trégua/impasse	trégua/impasse	Semântico	Conh. Mundo	
					D3_C13_Gpovo	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D3_C13_Gpovo	–	–	mas	mas	Marcador Discursivo	Conjunção	

D2_C16_ Estadão	“vou instaurar todos os processos juntos, mas a ideia é que os 15 casos mais graves, que têm provas contundentes, sejam julgados na frente”	78	40 - 43	8 de 11	D2_C16_ Estadão				todos os processos/ os 15 casos mais graves	Semântico	Hiper e Hiponímia	MD+GRAF+SEM
					D2_C16_ Estadão				mas	Marcador Discursivo	Conjunção	
D3_C16_ Estadão	“Dois dos 69 deputados acusados são da mesa diretora, mas se afastaram das funções”	78	56 e 57	5 de 11	D2_C16_ Estadão				mesa diretora/ afastaram das funções	Semântico	Conh. Mundo	MD+GRAF+SEM
					D2_C16_ Estadão				mas	Marcador Discursivo	Conjunção	
D1_C29_ Folha	"Cada vítima contemplada pelo acerto deve receber mais de US\$ 1 milhão, mas discute-se o quanto o dinheiro vai reparar o sofrimento dos que, quando criança, sofreram abusos."	57	39 - 42	5 de 11	D1_C29_ Folha	—	—	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM +SEM
					D1_C29_ Folha	—	—	US\$ 1 milhão /dinheiro	US\$ 1 milhão /dinheiro	Semântico	Campo semântico	
					D1_C29_ Folha	—	—	receber/reparar	receber/reparar	Semântico	Antonímia	
D2_C33_ Estadão	"Cada um de nós deve assumir sua parte nesta tarefa, mas não	105	35 - 36	8 de 11	D2_C33_ Estadão	vírgula	—	—	vírgula	Gráfico	Pontuação	MD+GRAF

	é admissível que o ônus maior da imprevidência dos privilegiados recaia sobre os despossuídos da terra".				D2_C33_Estadão	mas	mas	–	mas	Marcador Discursivo	Conjunção	
D2_C33_Estadão	"A mensagem do artista é singela, mas poderosa"	105	90 - 91	8 de 11	D2_C33_Estadão	vírgula	vírgula	–	vírgula	Gráfico	Pontuação	MD+GRAF+SEM
					D2_C33_Estadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	
					D2_C33_Estadão	–	–	singela/poderosa	singela/poderosa	Semântico	Antonímia	
D4_C36_JB	"Antonio Carlos Peixoto de Magalhães formou-se médico, mas encontrou na política a sua profissão e nunca exerceu a medicina."	86	21 - 23	4 de 10	D4_C36_JB	formou-se médico/nunca exerceu a medicina	–	formou-se médico/nunca exerceu a medicina	formou-se médico/nunca exerceu a medicina	Semântico	Antonímia	SEM+MD+MD
					D4_C36_JB	–	–	nunca	nunca	Marcador Discursivo	Advérbio	
					D4_C36_JB	–	–	e	e	Marcador Discursivo	Conjunção	
D1_C37_OGlobo	"A cadeia abriga 203 detentos, mas só tem capacidade para 80."	38	32 - 33	4 de 11	D1_C37_OGlobo	cadeia/capacidade	–	–	cadeia/capacidade	Semântico	Hiper e Hiponímia	SEM+SIN+MD+GRAF+SEM
					D1_C37_OGlobo	203/80	–	203/80	203/80	Sintático	Numeral	
					D1_C37_OGlobo	mas só	mas	mas	mas	Marcador Discursivo	Conjunção	
					D1_C37_OGlobo	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D1_C37_OGlobo	–	–	cadeia abriga 203/capacidade para 80	cadeia abriga 203/capacidade para 80	Semântico	Antonímia	

D1_C44_Folha	"E se um tiver um entendimento sobre as denúncias e os outros não concordarem?"	65	41 - 42	8 de 10	D1_C44_Folha	não	não	–	não	Marcador Discursivo	Advérbio	MD+MD+SEM
					D1_C44_Folha	–	e se/ e	–	e se/ e	Marcador Discursivo	Conjunção	
					D1_C44_Folha	–	tiver um entendimento/não concordarem	–	tiver um entendimento/não concordarem	Semântico	Antonímia	
D2_C42_Estadão	"O petista havia dado indícios de que iria pedir o arquivamento do caso, mas a idéia agora é esperar até que as investigações sejam concluídas na Câmara dos Deputados"	78	08 a 11	7 de 12	D2_C42_Estadão	mas	–	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+SEM+MORF
					D2_C42_Estadão	–	–	vírgula	vírgula	Gráfico	Pontuação	
					D2_C42_Estadão	–	–	pedir o arquivamento do caso/esperar até que as investigações sejam concluídas	pedir o arquivamento do caso/esperar até que as investigações sejam concluídas	Semântico	Antonímia	
					D2_C42_Estadão	–	–	iria/ é	iria/ é	Morfológico	Tempo-verbal	
D2_C42_Estadão	"A posição de Quintanilha tem o apoio de alguns senadores, mas não é bem vista pela oposição"	78	56 - 57	9 de 12	D2_C42_Estadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MORF+MD+SEM
					D2_C42_Estadão	–	vírgula	vírgula	vírgula	Gráfico	Pontuação	
					D2_C42_Estadão	–	–	tem/é	tem/é	Morfológico	Tempo-verbal	
					D2_C42_Estadão	–	–	não	não	Marcador Discursivo	Advérbio	
					D2_C42_Estadão	–	–	tem apoio de alguns senadores/	tem apoio de alguns senadores/não é	Semântico	Antonímia	

								não é bem vista pela oposição	bem vista pela oposição			
D2_C42_Estadão	"A posição de Quintanilha tem o apoio de alguns senadores, mas não é bem vista pela oposição"	78	71 - 72	5 de 12	D2_C42_Estadão	mas	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MORF+MD+SEM
					D2_C42_Estadão	–	vírgula	vírgula	vírgula	Gráfico	Pontuação	
					D2_C42_Estadão	–	–	tem/é	tem/é	Morfológico	Tempo-verbal	
					D2_C42_Estadão	–	–	não	não	Marcador Discursivo	Advérbio	
					D2_C42_Estadão	–	–	tem apoio de alguns senadores/não é bem vista pela oposição	tem apoio de alguns senadores/não é bem vista pela oposição	Semântico	Antonímia	
D3_C45_Oglobo	"A polícia afirma que chamou Luciano Huck para fazer o reconhecimento do criminoso, mas a assessoria do apresentador informou que ele está gravando no Rio de Janeiro e não poderá ir até a delegacia."	71	16 - 20	7 de 10	D3_C45_Oglobo	–	mas	mas	mas	Marcador Discursivo	Conjunção	MD+GRAF+MD
					D3_C45_Oglobo	–	não	–	não	Marcador Discursivo	Advérbio	

D3_C45_ Oglobo	"que vê o Brasil indo tão bem em tantas áreas - índices nunca antes alcançados, capital estrangeiro entrando, notoriedade no mundo todo - e muito mal em questões básicas que a gente não consegue resolver, como a questão da violência."	71	65 - 71	de 10	D3_C45_ Oglobo	—	que vê o Brasil indo tão bem em tantas áreas	que vê o Brasil indo tão bem em tantas áreas	que vê o Brasil indo tão bem em tantas áreas	Sintático	Oração Circunstancial	IND
-------------------	--	----	---------	-------	-------------------	---	--	--	--	-----------	-----------------------	-----

Fonte: Elaborado pela autora.

ANEXO A

Tabela 5 - Lista de relações RST

Relações	Multinuclear	Natureza das relações
ANTITHESIS	Não	Intencional
ATtribution	Não	Estrutural
BACKGROUND	Não	Intencional
CIRCUMSTANCE	Não	Semântica
COMPARISON	Não	Semântica
CONCESSION	Não	Intencional
CONCLUSION	Não	Semântica
CONDITION	Não	Semântica
ELABORATION	Não	Semântica
ENABLEMENT	Não	Intencional
EVALUATION	Não	Semântica
EVIDENCE	Não	Intencional
EXPLANATION	Não	Semântica
INTERPRETATION	Não	Semântica
JUSTIFY	Não	Intencional
MEANS	Não	Semântica
MOTIVATION	Não	Intencional
NON-VOLITIONAL CAUSE	Não	Semântica
NON-VOLITIONAL RESULT	Não	Semântica
OTHERWISE	Não	Semântica
PARENTHETICAL	Não	Estrutural
PURPOSE	Não	Semântica
RESTATEMENT	Não	Semântica
SOLUTIONHOOD	Não	Semântica
SUMMARY	Não	Semântica
VOLITIONAL CAUSE	Não	Semântica
VOLITIONAL RESULT	Não	Semântica
CONTRAST	Sim	Semântica
JOINT	Sim	Semântica
LIST	Sim	Semântica
SAME-UNIT	Sim	Estrutural
SEQUENCE	Sim	Semântica

Fonte: Pardo (2005, p. 135)