

Relações Retóricas de *List* e *Sequence* em textos jornalísticos

Tobias J. A. Almeida¹, Patrícia V. Almeida¹, Paula C.F. Cardoso²

¹FAELCH/ Departamento de Estudos da Linguagem – Universidade Federal de Lavras

²Faculdade de Computação – Universidade Federal do Pará (UFPA)

{tobias.almeida@estudante.ufla.br, patriciaalmeida@ufla.br,
pcardoso@ufpa.br}

Abstract. This study investigates the rhetorical relations of *List* and *Sequence* in journalistic texts to identify discourse signals (DS) beyond conventional discourse markers (MDs). Utilizing a corpus of 87 texts from CSTNews, the research considers various elements, including morphological and syntactic aspects. Results reveal that other discourse signals significantly contribute to identifying these relations. The study concludes that integrating diverse discourse signals enhances discourse analysis and suggests new research directions.

Resumo. Esta pesquisa investiga as relações retóricas *List* e *Sequence* em textos jornalísticos, visando identificar sinalizadores discursivos (SDs) além dos marcadores discursivos (MDs) convencionais. Utilizando um corpus de 87 textos do CSTNews, a análise abrange elementos morfológicos, semânticos, sintáticos e gráficos. Os resultados mostram que outros SDs são essenciais para a identificação dessas relações. Conclui-se que a combinação de diferentes SDs enriquece a análise discursiva e sugere novas direções de pesquisa.

1. Introdução

A *Rhetorical Structure Theory* (RST), ou Teoria da Estrutura Retórica em português, é uma teoria linguístico-descritiva que tem como foco a organização textual, caracterizando as relações estabelecidas entre as partes de um texto para formar uma estrutura discursiva coesa e interligada, frequentemente representada em forma de árvore (Mann e Thompson, 1987). As relações retóricas descritas pela RST indicam como essas partes se conectam para compor um todo coerente e significativo. A teoria concentra-se nas interações entre as unidades de informação que constituem um texto, evidenciando como sua organização contribui tanto para a expressão da intenção comunicativa do autor quanto para orientar a interpretação do leitor.

Comumente, essas relações são identificadas por meio de marcadores discursivos (MDs), que são elementos coesivos compostos por uma ou mais palavras que tornam explícita a relação entre as partes de um texto (Koch e Travaglia, 2002). No entanto, a presença de MDs nem sempre é clara ou suficiente para definir a relação retórica existente. Nesse contexto, Taboada e Das (2013; 2018) sugerem que a análise deve ir além dos MDs, incorporando outros elementos formais conhecidos como

sinalizadores discursivos (SD) e discutem a necessidade de identificar outros elementos, uma vez que nem todas as relações retóricas apresentam um MD de forma explícita.

Diante do exposto, este trabalho investiga as relações retóricas *List* e *Sequence*, que, de acordo com Mann e Thompson (1987), são multinucleares, ou seja, compostas por múltiplos núcleos de igual importância. Na relação *List*, os elementos apresentados são comparáveis e conectados por uma estrutura de lista, em que o leitor reconhece a possibilidade de comparação entre os elementos apresentados. Por outro lado, a relação *Sequence*, apresenta uma sucessão temporal de eventos, descritos de forma sequencial em que o leitor percebe na construção das proposições a sequência em que os fatos apresentados ocorreram. Nesse sentido, com base na noção de SDs e no estudo de Rodrigues, Souza e Cardoso (2023), o objetivo desta pesquisa é examinar quais SDs, restritos às proposições intrasentenciais, indicam a presença das relações *List* e *Sequence* em textos jornalísticos do português brasileiro (PB). Na seção de resultados, são apresentados exemplos de alguns dos SDs observados até o momento para essas relações.

2. Trabalhos relacionados

Nesta seção, destacam-se três trabalhos relacionados aos objetivos desta pesquisa.

Antônio e Rúbio (2012) analisaram um corpus de elocuções formais em PB, concluindo que as relações entre orações podem ser descritas pela intenção comunicativa do autor e sua avaliação do destinatário. Eles observaram que, na relação *List*, o conectivo “e” é usado para estabelecer a relação, sem subsequência temporal, e os elementos são comparáveis. Na relação *Sequence*, observaram o uso do conectivo “e” com conjugação verbal, indicando sequência temporal, além de conectivos de sequenciadores temporais.

Antônio (2017), usando um corpus oral, argumentou que as relações RST são baseadas em sentido, não apenas na forma, e são interpretadas independentemente de serem marcadas explicitamente por conectivos. Ele identificou diversos SD, incluindo pontuação, correlação modo-temporal, conteúdo textual, paralelismo sintático, paráfrase, inserção parentética, repetição e apresentação de evidências.

Para a língua inglesa, Das e Taboada (2018) analisaram os SD no corpus *RST Discourse Treebank*. Eles identificaram que uma relação pode ser marcada por um sinalizador único (como MD, referências pessoais, orações relativas) ou por um sinalizador combinado (como vírgula + oração no particípio passado, construção sintática paralela + cadeia lexical).

3. Metodologia

Para realizar esta pesquisa, foi utilizado um corpus de textos jornalísticos em PB, anotado manualmente em diversos aspectos da organização do discurso, como a RST. O corpus utilizado foi o CSTNews¹ (Cardoso et al., 2011), composto por 50 conjuntos de textos jornalísticos coletados em 2007, que abrangem diversos tópicos de agências de notícias online populares no Brasil.

Com foco nas relações discursivas de *List* e *Sequence*, foram selecionados 31 *clusters*, compostos por 2 ou 3 textos cada, resultando em um total de 87 textos. O

¹Disponível em: <http://nilc.icmc.usp.br/CSTNews/login/?next=/CSTNews/>

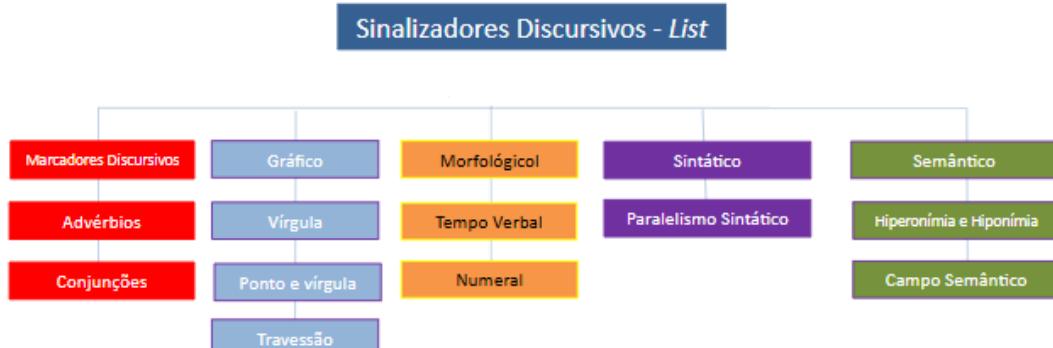
anotador, a partir da leitura e da investigação dos elementos presentes nos textos, identificou e anotou os SDs encontrados nas proposições. Baseando-se em estudos prévios, supramencionados, a tarefa consistia em identificar outros possíveis sinalizadores que ocorressem nas relações analisadas, considerando elementos morfológicos, sintáticos, semânticos e gráficos. Para isso, foi empregada a taxonomia de SDs catalogada no estudo de Dantas et al. (2024), com o objetivo de identificar as classificações existentes ou propor novas categorias a partir dessa taxonomia. A ferramenta *rstWeb* (Zeldes, 2016) foi utilizada para visualizar os textos em formato de árvore, facilitando a anotação dos SDs. Concluída essa etapa, foi realizada uma reflexão e análise dos elementos formais que constituem os sinalizadores discursivos, a fim de atingir os objetivos da investigação.

4. Discussão e resultados parciais

Nesta seção, são apresentados os resultados parciais obtidos a partir da análise geral dos dados coletados, com ênfase nas relações *List* e *Sequence*. A análise dos textos anotados considerou, além dos MDs, outros elementos formais como SDs, incluindo aspectos morfológicos, semânticos, sintáticos e gráficos, por serem pistas linguísticas e estruturais com grande potencial para indicar as relações propostas pela RST.

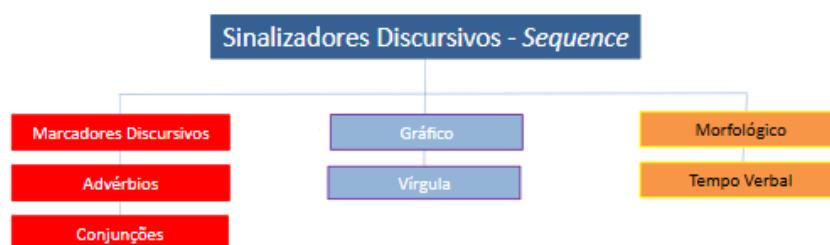
Na Figura 1, é apresentada uma proposta de categorização dos elementos encontrados para a relação de *List*.

Figura 1. Sinalizadores Discursivos de List



Na Figura 2, é apresentada uma proposta de categorização dos elementos encontrados para a relação de *Sequence*.

Figura 2. Sinalizadores Discursivos de Sequence



Na Figura 3, são apresentados alguns exemplos, selecionados aleatoriamente dentro do conjunto de 87 textos, de SDs encontrados nos textos analisados.

Figura 3. Exemplo de *List* e *Sequence*

Relação	Sinalizadores	Exemplos
List	MD (advérbio) + gráfico (vírgula) + morfológico (sentido verbal) + paralelismo sintático + MD (conjunção) sintático (numeral) sintático (numeral) + morfológico (sentido verbal) + MD (conjunção)	(1) Neste cenário, Lula sobe de 27% para 31%, e Alckmin passa de 4% para 14%, Heloísa também cresce, de 1% para 6%. (2) ... 57 pessoas morreram e 128 ficaram feridas
Sequence	gráfico (vírgula) + morfológico (tempo verbal) + MD (advérbio) + MD (advérbio) + MD (conjunção) + morfológico (tempo verbal)	(5) Na política partidária, iniciou na extinta UDN, tendo sido eleito para um mandato de deputado estadual e dois de federal até 1967, até ser nomeado prefeito de Salvador. (7) ... chegou na semana passada à ilha e manteve diferentes reuniões com as duas partes.

Os resultados apresentados na Figura 3, foram selecionados aleatoriamente a partir do com foco nas construções que continham o maior número de SDs. A seleção aleatória foi realizada com o intuito de exemplificar uma amostra representativa das construções que mais se destacavam pela presença de SDs. Tais resultados revelam que para ambas relações analisadas os SDs se constituem na combinação de MDs prototípicos já identificados pela literatura e outros elementos formais supramencionados. Embora a conjunção "e" seja frequentemente utilizada como um MD para ambas as relações, outros sinalizadores ajudam a distingui-las. Nas relações de *List*, observa-se uma igualdade entre os itens apresentados, que nem sempre seguem uma ordem temporal, sendo geralmente construídos em paralelismo sintático. Já nas relações de *Sequence*, a sucessão temporal dos eventos é enfatizada, frequentemente indicada por sinalizadores morfológicos, como o tempo verbal, e por MDs, como advérbios temporais, que detalham a ordem cronológica dos fatos.

Além disso, foi observado que a relação de *List*, nos exemplos analisados neste estudo (embora nem todos foram exemplificados aqui), é simétrica, ou seja, os elementos das orações podem ser trocados sem alterar o sentido. Em contraste, a relação de *Sequence* é assimétrica, pois indica uma sucessão temporal, o que impede a troca dos elementos sem modificar o sentido.

5. Considerações Finais

Neste estudo preliminar, o objetivo foi explorar pistas adicionais que pudesse indicar as relações de *List* e *Sequence* no modelo teórico da RST, utilizando um corpus previamente anotado. Os resultados ressaltam a importância de considerar sinalizadores discursivos além dos MDs, extrapolando a análise superficial a fim de ir mais a fundo na compreensão das relações analisadas.

Para pesquisas futuras, planeja-se ampliar a amostra e expandir as análises para incluir relações intersentenciais. Além disso, será relevante investigar outros gêneros textuais, como textos não estruturados (por exemplo, tweets), dado que o corpus utilizado nesta pesquisa é composto exclusivamente por textos jornalísticos.

Agradecimentos

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

6. Referências

- Antonio, Juliano Desiderato. **Mecanismos utilizados pelos destinatários do discurso para identificação de relações de coerência não sinalizadas por conectores.** DELTA: Documentação e Estudos em Linguística Teórica e Aplicada, v. 33, n. 1, 2017.
- Antonio, Juliano Desiderato; RÚBIO, Marília Gabriela. **Relações retóricas de lista e de sequência estabelecidas por orações paratáticas aditivas e por orações paratáticas justapostas: uma investigação funcionalista.** Signum: Estudos da Linguagem, v. 15, n. 1, p. 33-56, 2012
- Das, D. e Taboada, M. (2018) **RST Signalling Corpus: A corpus of signals of coherence relations.** Language Resources and Evaluation, Vol 52, N. 1, pp. 149-184
- DANTAS, Ewerson et al. **Manual de anotação de sinalizadores discursivos em textos jornalísticos.** São Carlos: Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. Disponível em: <https://repositorio.usp.br/item/003207370>. Acesso em: 25 ago. 2024. , 2024
- Koch, I.V. e TRAVAGLIA, L.C. (2002). **A Coerência Textual.** Editora Contexto.
- Mann, W. C.; Thompson, S. A. **Rhetorical structure theory: Toward a functional theory of text organization.** Text-interdisciplinary Journal for the Study of Discourse, v. 8, n. 3, p. 243–281, 1988.
- Rodrigues, R.; Souza, J.W.C; Cardoso, P.C.F.. **Sinalizadores retórico-discursivos: revisitando a anotação RST no córpus CSTNews.** In: Anais do XIV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana. SBC, 2023. p. 249-257
- Souza, J.W.C.; Cardoso, P.C.F.; Rodrigues, R. **Systematic Review of Studies on Rhetorical Structure Theory (RST).** Revista de Estudos da Linguagem, [s.l.], v. 31, n. 3, p. 1643-1675, 2023.
- Taboada, Maite; DAS, Debopam. **Annotation upon annotation: Adding signalling information to a corpus of discourse relations.** Dialogue & Discourse, v. 4, n. 2, p. 249-281, 2013.