



DISCURSO – PARTE 1

SCC5869 Tópicos em Processamento de Língua Natural

Thiago A. S. Pardo

DISCURSO VS. DISCURSOS

- A discourse is an **instance of language use** whose type can be classified on the basis of such factors as **grammatical and lexical choices** and their distribution in
 - **main versus supportive materials**
 - **Theme**
 - **Style, and**
 - The framework of **knowledge** and **expectations** within which the addressee interprets the discourse

2

DISCURSO VS. DISCURSOS

- A discourse is an **extended sequence of sentences** produced by one or more people with the **aim of conveying or exchanging information**.

Mitkov, R. (2004). *The Oxford Handbook of Computational Linguistics*.

3

DISCURSO VS. DISCURSOS

- A discourse is any string of language – usually one that is **more than one sentence long**. Text-books, novels, weather reports and conversations are all discourses.

Russel, S. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*.

4

DISCURSO VS. DISCURSOS

- **Diversas perspectivas** de análise
 - Texto **vs.** discurso
 - Texto **&** discurso
 - Texto-discurso
- **Linha** europeia vs. americana
- Etc.

5

DISCURSO

- **Um texto é mais do que uma simples seqüência de sentenças justapostas**
 - Estrutura altamente elaborada
 - Diversas propriedades em níveis variados
 - Coesão
 - Coerência
 - Tessitura
 - Progressão temática
 - Tópicos
 - Intenções
 - Estilo
 - Etc.
- Há uma “sintaxe textual”

6

COESÃO VS. COERÊNCIA

- **Coesão**: uso de mecanismos lingüísticos para conectar unidades textuais → “cola” textual
 - Palavras (mesma palavra, sinônimos, hiperônimos, etc.)
 - Expressões referenciais
 - Marcadores discursivos, frases e palavras indicativas
 - Pontuação
 - Paralelismo sintático
 - Etc.
 - Superfície textual
- **Coerência**: relação de significado entre unidades textuais, como o significado do discurso pode ser inferido a partir das partes textuais
 - Conteúdo textual

7

COESÃO VS. COERÊNCIA

- **Fatos**
 - Há textos coesos e coerentes
 - Há textos não coesos e não coerentes
 - Apesar da capacidade humana de dar sentido a quase tudo!

João escondeu as chaves do carro de Pedro. Ele estava bêbado.

João escondeu as chaves do carro de Pedro. Ele gosta de espinafre.
- **Questões**
 - Há textos coesos e não coerentes?
 - Há textos não coesos e coerentes?

8

EXEMPLO: ANÁLISE DISCURSIVA

- *The Tin Woodman went to the Emerald City to see the Wizard of Oz and ask for a heart. After he asked for it, the Woodman waited for the Wizard's response.*
 - *he = Tin Woodman (e não Wizard of Oz)*
 - *it = the heart (e não Emerald City)*
 - *the Wizard = the Wizard of Oz*
 - *the Woodman = the Tin Woodman*
 - *to see the Wizard of Oz and ask for a heart = razão para The Tin Woodman went to the Emerald City*
 - Seqüência temporal entre as sentenças
 - Marcadores explícitos: *after*
 - Tópico = *asking for a heart to the wizard*
 - Etc.

9

EXEMPLO: ANÁLISE DISCURSIVA

- Gracie: Oh, yeah... and then Mr. And Mrs. Jones were having matrimonial trouble, and my brother was hired to watch Mrs. Jones.
 - George: Well, I imagine she was a very attractive woman.
 - Gracie: She was, and my brother watched her day and night for six months.
 - George: Well, what happened?
 - Gracie: She finally got a divorce.
 - George: Mrs. Jones?
 - Gracie: No, my brother's wife.
- E esse caso?

10

EXEMPLO: ANÁLISE DISCURSIVA

- Gracie: Oh, yeah... and then Mr. And Mrs. Jones were having matrimonial trouble, and my brother was hired to watch Mrs. Jones.
 - George: Well, I imagine she was a very attractive woman.
 - Gracie: She was, and my brother watched her day and night for six months.
 - George: Well, what happened?
 - Gracie: She finally got a divorce.
 - George: Mrs. Jones?
 - Gracie: No, my brother's wife.
- Referentes de pronomes e nomes, resolução do *my*, mecanismos de "continuidade" (*well*), turnos, expectativas, humor, progressão temática, intenção de manter uma conversação/desabafar, etc.

11

DISCURSO

- Tipos
 - "Monólogo"
 - Discurso em um único sentido, do escritor para o leitor
 - Diálogo
 - Turnos
 - Humano-humano, humano-máquina (tema de estudo de IHC)
- Material
 - Monodocumento: um único texto
 - Multidocumento: vários textos

12

DISCURSO

o Nível

- Mais refinado: micro-análise
 - o Por exemplo, RST (Mann e Thompson, 1987), Jordan (1992), etc.
- Blocos textuais: macro-análise
 - o Por exemplo, modelo Problema-Solução (Winter, 1976; Jordan, 1980), CARS - *Create a Research Space* (Swales, 1990), etc.
 - o Complementares

o Escopo

- Local: contexto local de um segmento textual
 - o Entidades
- Global: textual, relacional
 - o Sentido global
 - o Relacionamento entre partes textuais

13

NÍVEIS DE ANÁLISE DA LÍNGUA

o Entre os níveis mais abstratos

- Mas considerado por muitos menos complexo do que níveis inferiores
- Forte interação (nebulosa!) com semântica e pragmática
- Influência em todos os níveis de análise

Abstração &
complexidade



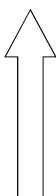
Pragmática/**Discurso**
Semântica
Sintaxe
Morfologia
Fonética/Fonologia

14

NÍVEIS DE ANÁLISE DA LÍNGUA

- Entre os níveis mais abstratos
 - Envolve a situação de comunicação (Koch e Travaglia, 2002): escritor/falante e leitor/ouvinte

Abstração &
complexidade



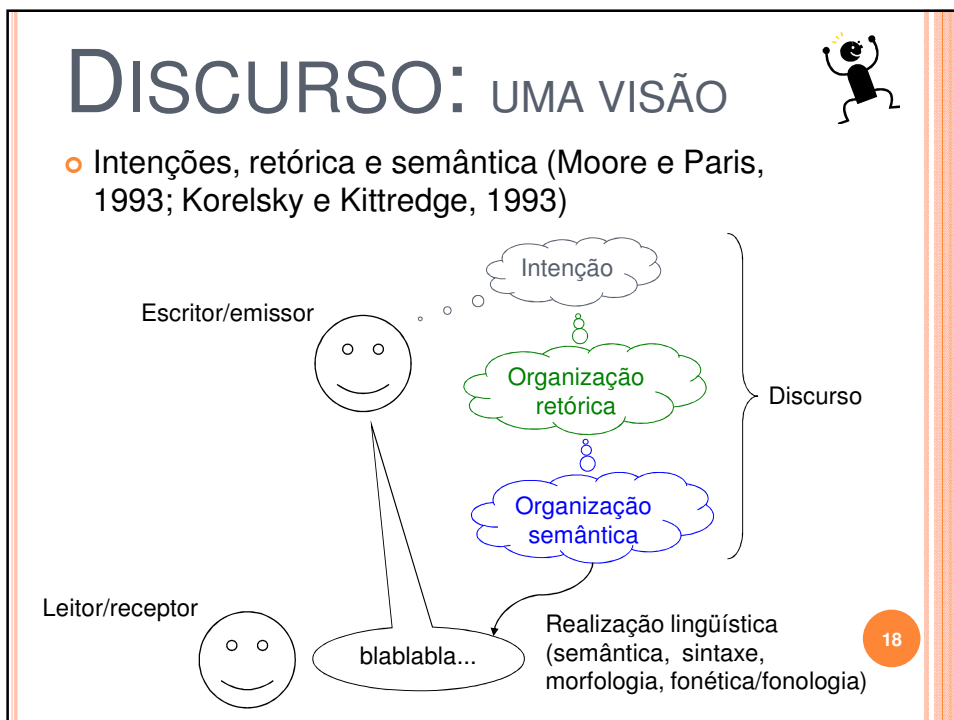
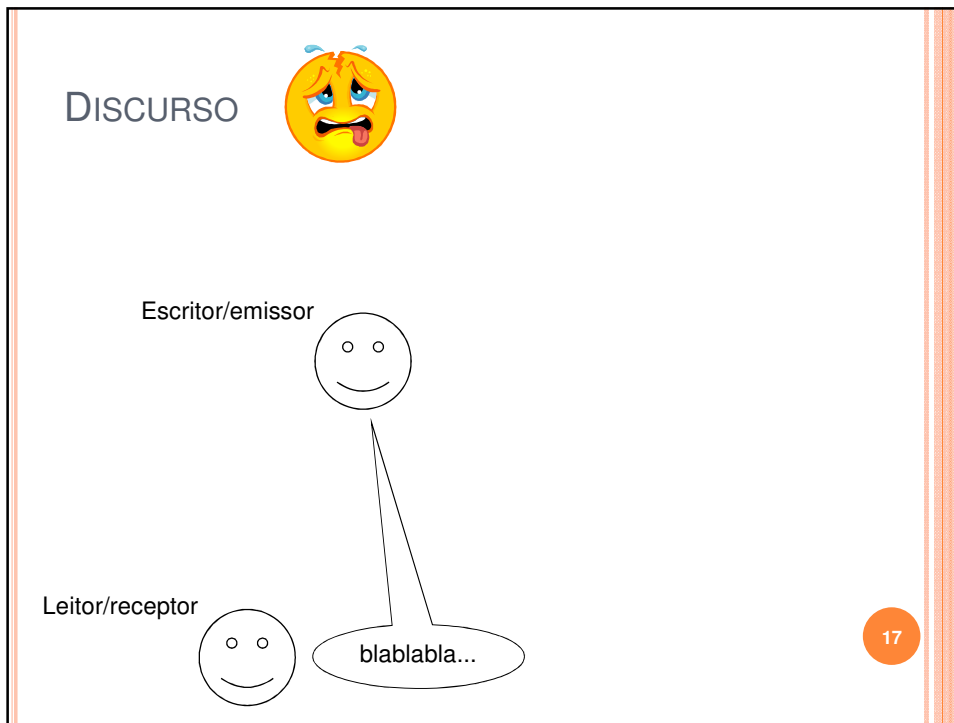
Pragmática/**Discurso**
Semântica
Sintaxe
Morfologia
Fonética/Fonologia

15

TEORIAS DISCURSIVAS (OU MODELOS?)

- Grosz e Sidner (1986): *attentions, intentions and the structure of discourse*
 - 2 relações principais
- Mann e Thompson (1987): *Rhetorical Structure Theory* (RST)
 - 23 relações
 - Uma das principais teorias discursivas
- Jordan (1992) e Kehler (2002): relações semânticas, esquemas textuais
 - Jordan: ~70 relações
 - Kehler: 3 relações
 - Relação entre fatos, desprovida de força argumentativa
- Moore e Pollack (1992), Moore e Paris (1993), Korelsky e Kittredge (1993), Moser e Moore (1996), Rino (1996) e Marcu (1999, 2000), entre outros: mapeamentos entre os níveis do discurso

16



TEORIAS DISCURSIVAS (OU MODELOS?)

- Hobbs (1979): relações de coerência
- Kamp (1981)
 - *Discourse Representation Theory* (DRT)
- Grosz et al. (1995)
 - *Centering*
- Cristea et al. (1998)
 - *Veins Theory*

19

TEORIAS DISCURSIVAS (OU MODELOS?)

- Radev (2000): a popularização multidocumento
 - *Cross-document Structure Theory*
- Asher e Lascarides (2003)
 - *Segmented Discourse Representation Theory* (SDRT)
- Wolf e Gibson (2005): retórica mais livre
- Afantenos et al. (2008)
 - Relações multidocumento sincrônicas e diacrônicas
- Etc.



20

DISCURSO

o **Propósitos** variados

- Descrição textual
- Estudo de fenômenos lingüísticos particulares
- Geração de língua natural
- Aplicações computacionais

o **Origens** variadas

- Lingüística
- Computação
- Filosofia

21

DISCURSO E PLN

o Duas grandes linhas de trabalho em PLN

- **Relacionamento** entre partes textuais
 - o Questões de coerência, propósito do texto, etc.
- **Entidades** textuais
 - o Correferências, anáforas, etc.

22

DISCURSO E PLN

o Muitos usos

- Sumarização automática
- Tradução automática
- Avaliação automática de redações
- Extração de informação
- Auxílio à escrita
- Perguntas e respostas
- Geração textual
- Resolução anafórica
- Etc.

o Como é usado?

23



RHETORICAL STRUCTURE THEORY

RST - *RHETORICAL STRUCTURED THEORY* (MANN E THOMPSON, 1987)

- Recupera a **estrutura retórica de textos**
 - Descreve relações entre o conteúdo de segmentos discursivos em termos funcionais
 - Proposições

- Premissa
 - Todo texto tem uma estrutura retórica subjacente que permite **recuperar o objetivo comunicativo** que o escritor do texto pretendia atingir ao escrevê-lo

25

RST

- **Retórica**: parte “palpável” da pragmática (Hovy, 1988)

- Meio pelo qual um texto é organizado para satisfazer um objetivo comunicativo
 - Intenção

- Organização funcional do texto
 - Função de suas partes para o sucesso da comunicação

26

PRINCIPAIS ELEMENTOS

- Estruturas retóricas (ou estrutura RST)
 - Compostas por **segmentos discursivos inter-relacionados** por meio de relações retóricas
 - **Segmentos discursivos**: unidades mínimas de significado que compõem um texto

27

PRINCIPAIS ELEMENTOS (...CONTINUAÇÃO)

- **Relação retórica**
 - Relações de significado existentes entre dois ou mais segmentos discursivos
 - Conjunto original: 23 relações
- Cada segmento discursivo representa um núcleo (N) ou satélite (S)
 - **Núcleo**, ou unidade nuclear, expressa a informação principal sendo, portanto, mais relevante do que o satélite
 - **Satélite** apresenta informação adicional, a qual exerce influência na interpretação do leitor sobre a informação apresentada em N

28

RST: EXEMPLO

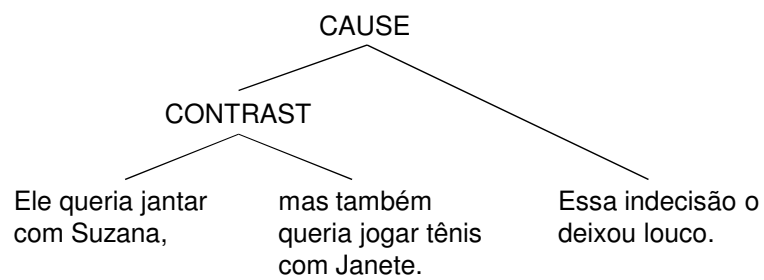
Ele queria jantar
com Suzana,

mas também
queria jogar tênis
com Janete.

Essa indecisão o
deixou louco.

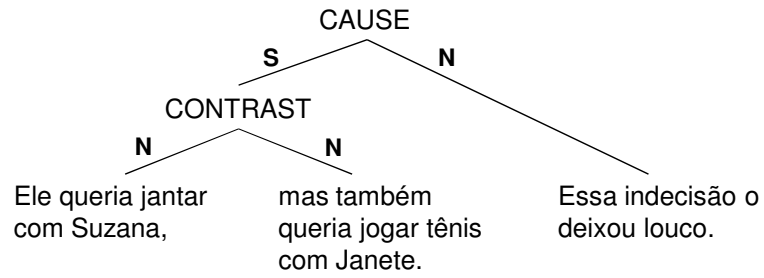
29

RST: EXEMPLO



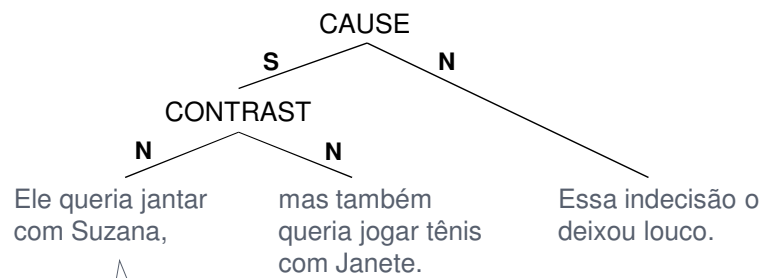
30

RST: EXEMPLO



31

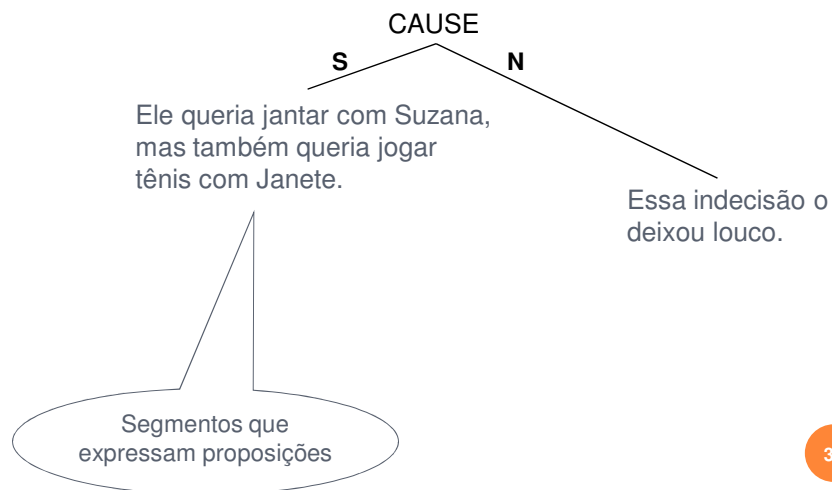
RST: EXEMPLO



Segmentos que expressam proposições

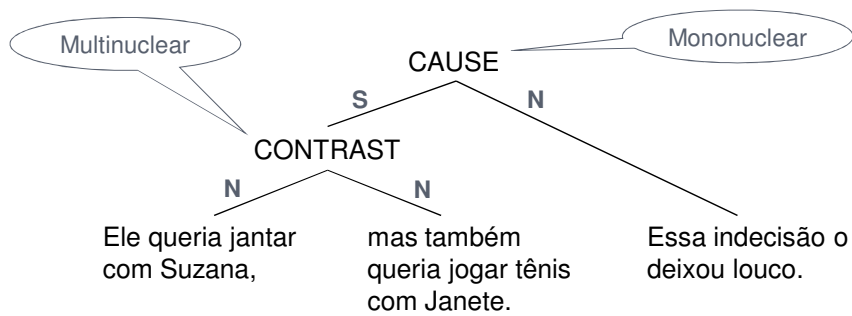
32

RST: EXEMPLO



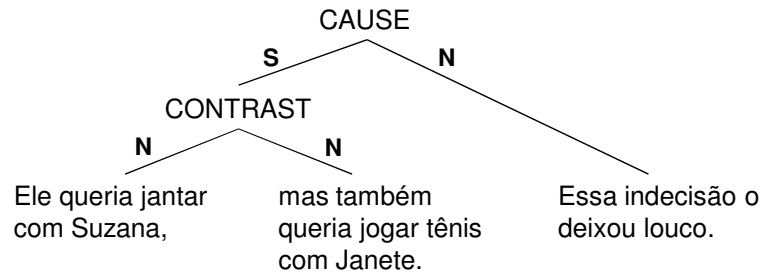
33

RST: EXEMPLO



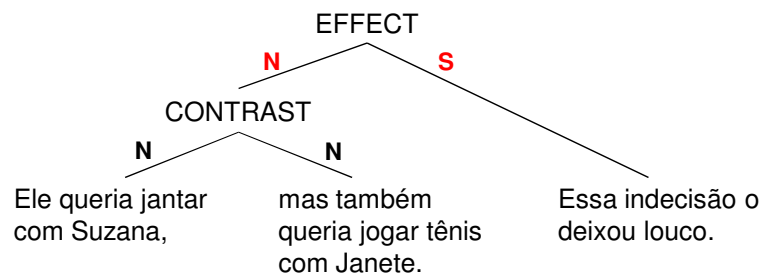
34

RST: EXEMPLO



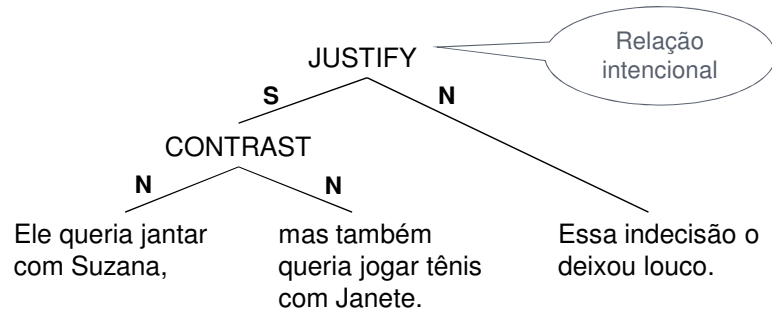
35

RST: EXEMPLO



36

RST: EXEMPLO



- Como decidir por uma relação?

37

RST

- Estilo tradicional de anotação

RELAÇÕES RETÓRICAS

○ Organização das relações

- Relações de Causa
 - *Volitional Cause*
 - *Volitional Result*
 - *Non-Volitional Cause*
 - *Non-Volitional Result*
 - *Purpose*
- Relações de Evidência e Justificativa
 - *Evidence*
 - *Justify*
- Relações de Motivação
 - *Enablement*
 - *Motivation*

39

RELAÇÕES RETÓRICAS (...CONTINUAÇÃO)

- Relações de Antítese e Concessão
 - *Antithesis*
 - *Concession*
- Relações de Condição
 - *Condition*
 - *Otherwise*
- Relações de Interpretação e Avaliação
 - *Interpretation*
 - *Evaluation*
- Relações de Reiteração e Sumário
 - *Restatement*
 - *Summary*

40

RELAÇÕES RETÓRICAS (...CONTINUAÇÃO)

- Outras
 - *Background*
 - *Circumstance*
 - *Contrast*
 - *Elaboration*
 - *Sequence*
 - *Solutionhood*
 - *Joint*

41

RST

◦ **Relações**

- **Intencionais:** alteram a inclinação do leitor para algo
- **Informativas:** informam o leitor sobre algo
 - Podem co-ocorrer? É natural que isso aconteça? Ou é um problema da teoria?

42

RST

- Exemplo de **ambigüidade**

*George Bush supports big business.
He's sure to veto House Bill 1711.*

- Relação de **evidência** (intencional) ou de **causa** (informativa)?
- O que fazer?

43

RST

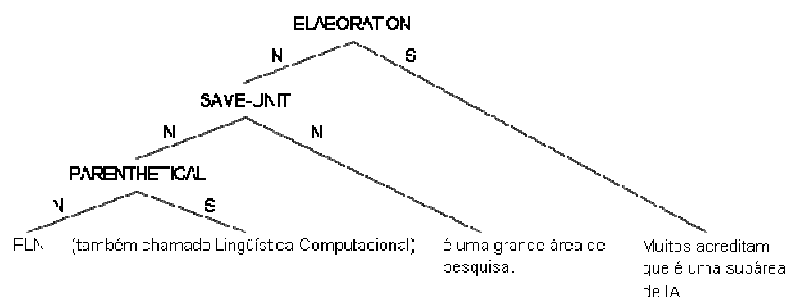
- **Relações**

- **Estruturais**: sem significado, apenas mantêm a integridade da estrutura textual
 - Um terceiro tipo, introduzido posteriormente

44

RST

- Exemplos de **relações estruturais**



45

DETERMINAÇÃO DE UMA RELAÇÃO ENTRE DUAS PROPOSIÇÕES

- Definição com 4 campos
 - Restrições sobre o núcleo (N)
 - Restrições sobre o satélite (S)
 - Restrições sobre o núcleo e o satélite
 - Efeito
 - Locus of effect?*
- Verificação das restrições para as proposições envolvidas
 - Escrita ou leitura de um texto

46

DEFINIÇÃO DE RELAÇÕES

Relação: JUSTIFY**Restrições sobre N:** não há**Restrições sobre S:** não há**Restrições sobre N+S:** a compreensão de S pelo leitor aumenta sua prontidão para aceitar o direito do escritor de apresentar N**Efeito:** a prontidão do leitor para aceitar o direito do escritor de apresentar N aumenta

47

DEFINIÇÃO DE RELAÇÕES

Relação: CAUSE**Restrições sobre N:** apresenta uma ação volitiva ou uma situação que poderia surgir de uma ação volitiva**Restrições sobre S:** não há**Restrições sobre N+S:** S apresenta uma situação que pode ter acarretado o fato do agente da ação volitiva em N ter realizado a ação; sem S, o leitor poderia não reconhecer a motivação da ação; N é mais central para a satisfação do objetivo do escritor do que S**Efeito:** o leitor reconhece que a situação apresentada em S como a causa da ação apresentada em N

48

DEFINIÇÃO DE RELAÇÕES

Relação: EVIDENCE

Restrições sobre N: o leitor pode não acreditar em N de forma satisfatória para o escritor

Restrições sobre S: o leitor acredita em S ou o acha válido

Restrições sobre N+S: a compreensão de S pelo leitor aumenta sua crença em N

Efeito: a crença do leitor em N aumenta

49

DEFINIÇÃO DE RELAÇÕES

Relação: CONCESSION

Restrições sobre N: o escritor julga N válido

Restrições sobre S: o escritor não afirma que S pode não ser válido

Restrições sobre N+S: 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

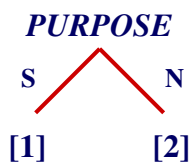
Efeito: o leitor aceita melhor N

50

TIPOS DE RELAÇÕES

- Relações Mononucleares X Relações Multinucleares
- Exemplo (Mononuclear): **PURPOSE**

[1] A fim de passar no vestibular de medicina, [2] Carla se dedicou aos estudos durante o ano todo.

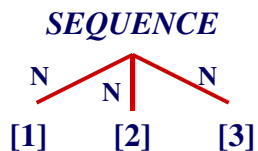


51

TIPOS DE RELAÇÕES (...CONTINUAÇÃO)

- Exemplo (Multinuclear): **SEQUENCE**

[1] A menina caiu da bicicleta, [2] quebrou o braço [3] e foi levada para o hospital.



52

ALGUMAS LIMITAÇÕES

- Subjetividade da análise e ambigüidade
 - Determinação dos segmentos
 - Orações, sentenças ou parágrafos?
 - Fragmentos com significado?
 - Determinação da nuclearidade
 - NS, SN ou NN?
 - Escolha das relações
 - CAUSE, EFFECT ou JUSTIFY?
 - CONCESSION, CONTRAST, ANTITHESIS ou OTHERWISE?
 - Relações de naturezas diferentes
 - Critérios de plausibilidade
 - Forma de representação da análise
 - Árvore ou grafo (Wolf e Gibson, 2005)?

53

EXTENSÕES DA RST

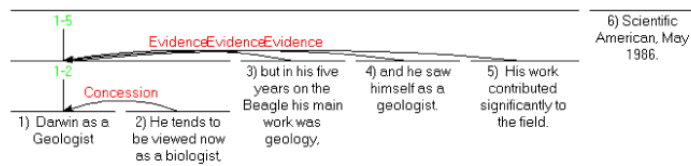
- Extensão do conjunto de relações: diversas possibilidades
 - Por exemplo, Marcu (1997)
 - Relações encaixadas (*embedded*)
- Propostas de regras de segmentação
 - Por exemplo, Carlson e Marcu (2001)
 - Manual de segmentação para textos escritos em inglês:
- Manuais com instruções para anotação humana de textos
 - Por exemplo, Marcu (1999)

54

ANÁLISE RETÓRICA

○ Restrições na análise

- Adjacência de segmentos, com algumas exceções

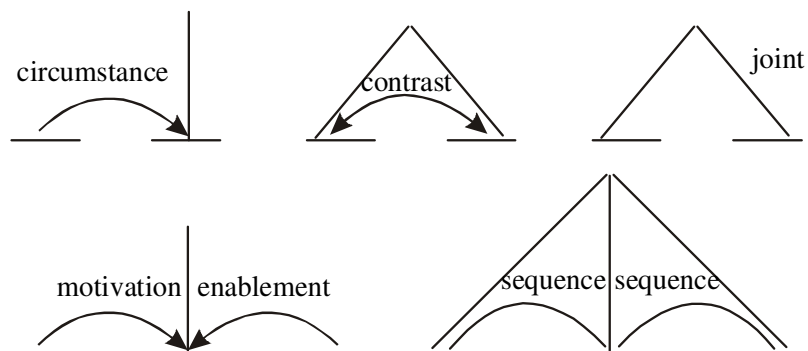


- Não compartilhamento de satélites
- Não sobreposição de segmentos

55

ANÁLISE RETÓRICA

○ Esquemas possíveis de análise



56

ALGUMAS REGRAS PARA SEGMENTAÇÃO (CARLSON E MARCU, 2001)

- Segmentar
 - Orações principais
 - Orações sinalizadas por marcadores discursivos como, por exemplo, *Porque, Apesar de, Conforme, Segundo, Em consequência de*, entre outros
 - Orações subordinadas introduzidas por marcadores discursivos
 - Orações coordenadas
 - Orações relativas, apositivas e parênteses como segmentos encaixadas

57

ALGUMAS REGRAS PARA SEGMENTAÇÃO (...CONTINUAÇÃO)

- Não segmentar
 - Orações complementares, exceto quando introduzirem complemento de um verbo de atribuição
 - Por exemplo: [1] *A companhia disse que* [2] *fechará a fábrica*
 - Orações subordinadas substantivas e objetivas

58

CONSTRUÇÃO DA ESTRUTURA RETÓRICA

o Construção composicional

- Formação de subestruturas: relacionamento de dois ou mais segmentos discursivos como, por exemplo, orações simples
- Formação da estrutura completa: relacionamento de segmentos discursivos mais elaborados (por exemplo, sentenças, parágrafos)
 - o Critério de composicionalidade: relações devem conectar elementos mais nucleares de subárvores

o Forma de representação mais comum: árvores

- Nós-folha: segmentos discursivos
- Nós intermediários: relações retóricas

59

EXEMPLO DE ANÁLISE RETÓRICA

o Texto: Mundo 7 (Corpus Rhetalho)

Uma explosão neste domingo em Pristina, capital do Kosovo, deixou três crianças feridas e causou estragos no prédio onde fica a sede do partido de oposição ORA, informaram policiais e testemunhas.

Um porta-voz da polícia do país disse que as crianças, que vivem no andar inferior ao do escritório do partido, sofreram ferimentos leves por vidros estilhaçados.

"Foram levadas a um hospital, mas já foram liberadas", disse Refki Morina. Segundo ele, a explosão causou grandes estragos no prédio.

O local foi isolado e membros da força de paz liderada pela Otan (Organização do Tratado do Atlântico Norte) se dirigiram ao local.

60

SEGMENTAÇÃO (ORACIONAL)

PARÁGRAFO 1: [1] Uma explosão neste domingo em Pristina, capital do Kosovo, deixou três crianças feridas e causou estragos no prédio [2] onde fica a sede do partido de oposição ORA, [3] informaram policiais e testemunhas.

PARÁGRAFO 2: [4] Um porta-voz da polícia do país disse que [5] as crianças, [6] que vivem no andar inferior ao do escritório do partido, [7] sofreram ferimentos leves por vidros estilhaçados.

PARÁGRAFO 3: [8] "Foram levadas a um hospital, mas já foram liberadas", [9] disse Refki Morina. [10] Segundo ele, a explosão causou grandes estragos no prédio.

PARÁGRAFO 4: [11] O local foi isolado, [12] e membros da força de paz liderada pela Otan (Organização do Tratado do Atlântico Norte) se dirigiram ao local.

61

CONSTRUÇÃO DA ESTRUTURA RETÓRICA

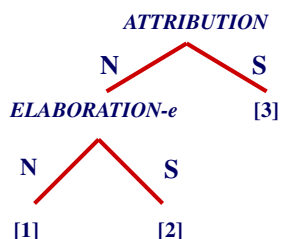
- Possível procedimento para análise: **análise incremental**
 - Relacionamento de todos os segmentos discursivos que compõem uma sentença, dois a dois
 - Relacionamento de todas as sentenças que compõem um parágrafo, duas a duas
 - Relacionamento de todos os parágrafos que compõem o texto, dois a dois

62

CONSTRUÇÃO DA ESTRUTURA RETÓRICA (...CONTINUAÇÃO)

○ PARÁGRAFO 1: **uma sentença**

[1] Uma explosão neste domingo em Pristina, capital do Kosovo, deixou três crianças feridas e causou estragos no prédio [2] onde fica a sede do partido de oposição ORA, [3] informaram policiais e testemunhas.

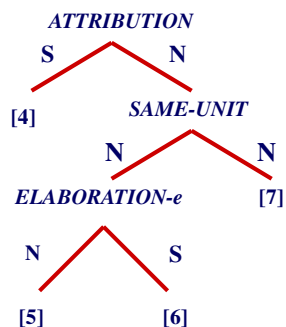


63

CONSTRUÇÃO DA ESTRUTURA RETÓRICA (...CONTINUAÇÃO)

○ PARÁGRAFO 2: **uma sentença**

[4] Um porta-voz da polícia do país disse que [5] as crianças, [6] que vivem no andar inferior ao do escritório do partido, [7] sofreram ferimentos leves por vidros estilhaçados.



64

CONSTRUÇÃO DA ESTRUTURA RETÓRICA (...CONTINUAÇÃO)

○ PARÁGRAFO 3: duas sentenças

• Sentença 1

[8] "Foram levadas a um hospital, mas já foram liberadas", [9] disse Refki Morina.



• Sentença 2

[10] Segundo ele, a explosão causou grandes estragos no prédio.

65

CONSTRUÇÃO DA ESTRUTURA RETÓRICA (...CONTINUAÇÃO)

○ PARÁGRAFO 3: duas sentenças

• Relacionamento: sentença 1 e sentença 2



[10]

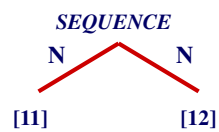
Relacionar ao
parágrafo 3 ou 4?

66

CONSTRUÇÃO DA ESTRUTURA RETÓRICA (...CONTINUAÇÃO)

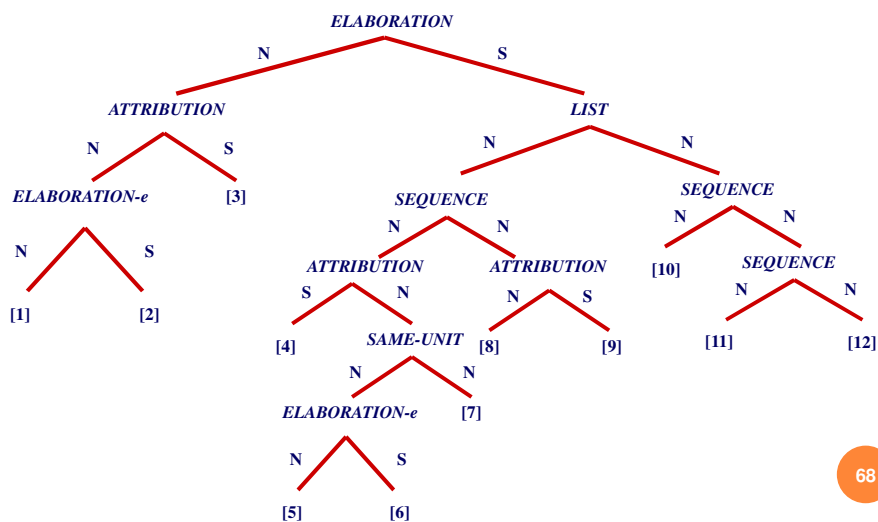
○ PARÁGRAFO 4: **uma sentença**

[11] O local foi isolado, [12] e membros da força de paz liderada pela Otan (Organização do Tratado do Atlântico Norte) se dirigiram ao local.



67

ESTRUTURA RETÓRICA DO TEXTO

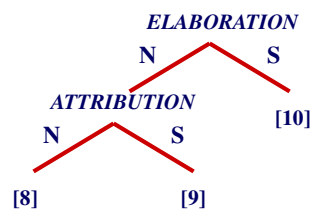


68

OUTRA ESTRUTURA POSSÍVEL

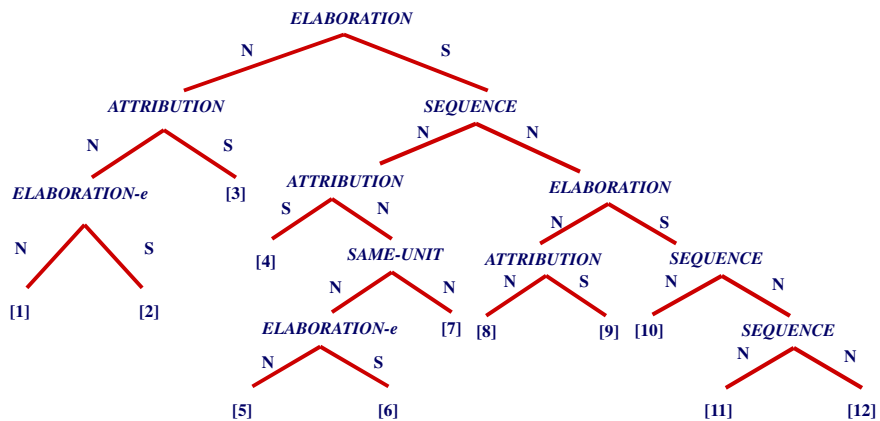
- PARÁGRAFO 3: duas sentenças

[8] "Foram levadas a um hospital, mas já foram liberadas", [9] disse Refki Morina. [10] Segundo ele, a explosão causou grandes estragos no prédio.



69

OUTRA ESTRUTURA POSSÍVEL (...CONTINUAÇÃO)



70