



## DISCURSO – PARTE 1

*SCC5908 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 sequência 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 linguísticos 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*
  - Sequência 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

### ○ Nível

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

### ○ Escopo

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

13

## NÍVEIS DE ANÁLISE DA LÍNGUA

### ○ 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

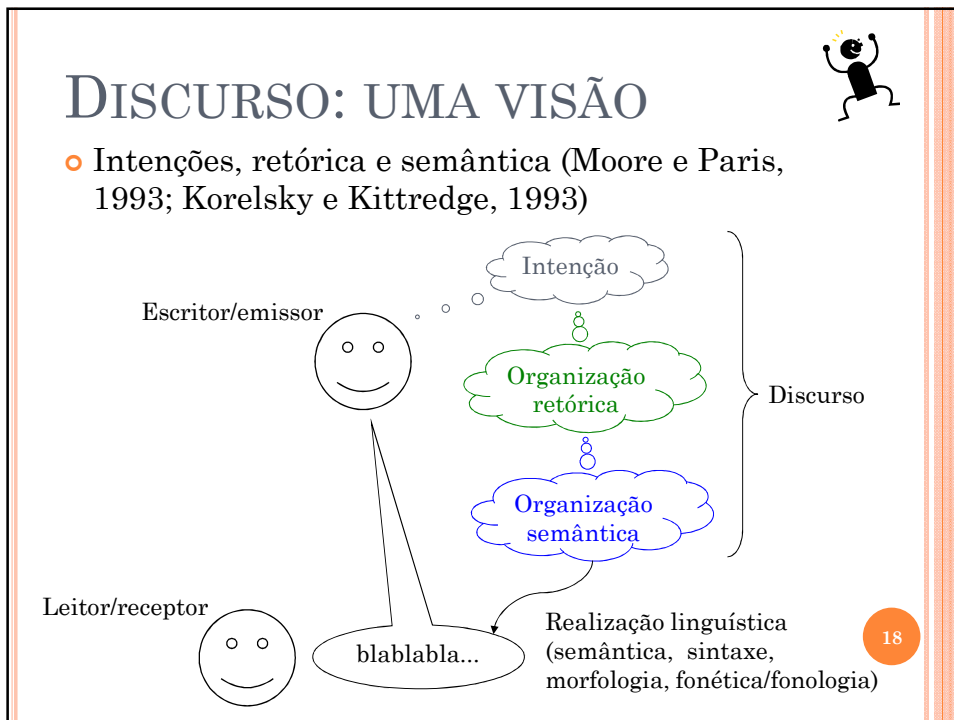
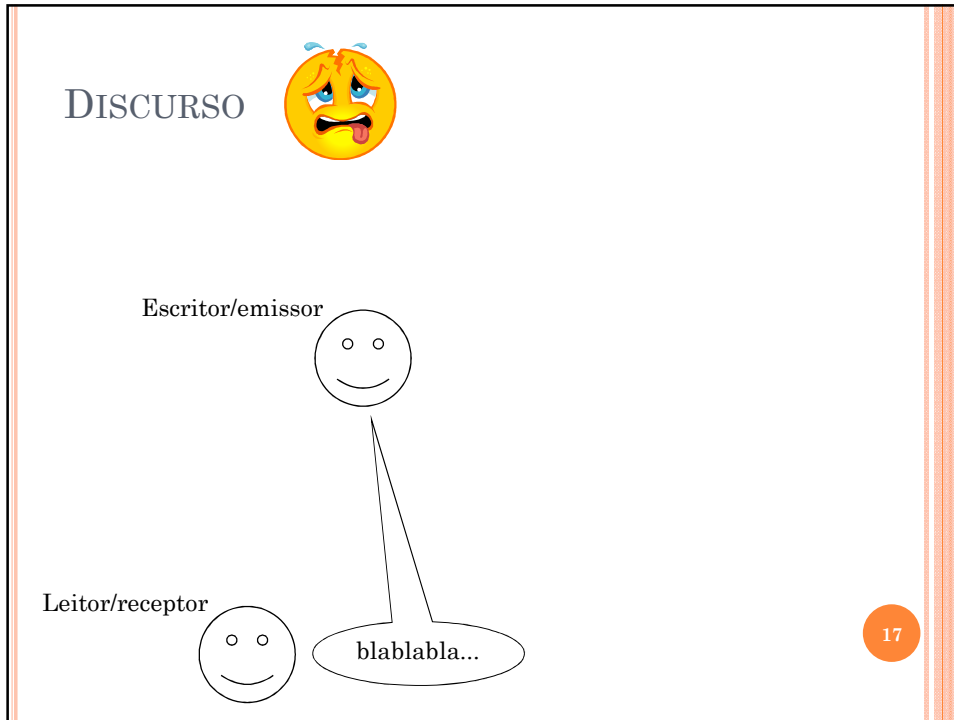
15

## 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

### ◦ **Propósitos** variados

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

### ◦ **Origens** variadas

- Linguística
- Computação

21

## DISCURSO E PLN

### ◦ Duas grandes linhas de trabalho em PLN

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

22

## DISCURSO E PLN

### ○ Muitos usos

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

○ Como é usado?

23



## *RHETORICAL STRUCTURE THEORY*

(MANN E THOMPSON, 1987)

## RST - *RHETORICAL STRUCTURE THEORY*

- 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

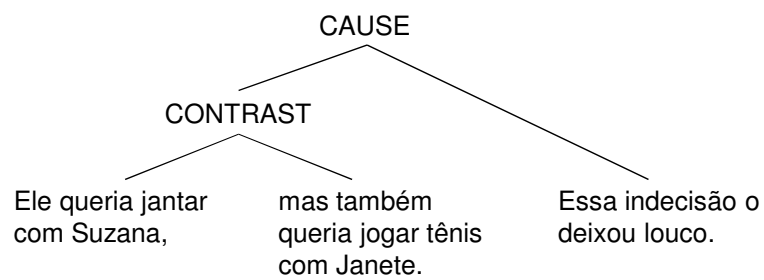
Ele queria jantar  
com Suzana,

mas também  
queria jogar tênis  
com Janete.

Essa indecisão o  
deixou louco.

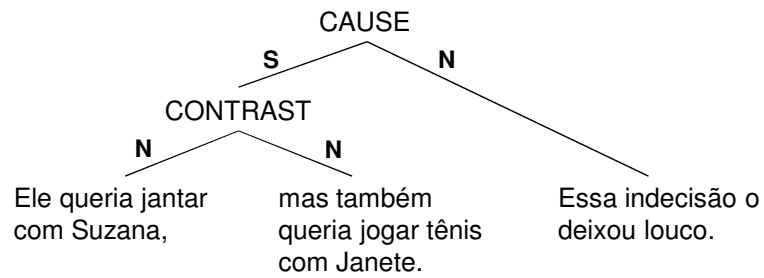
29

## RST: EXEMPLO



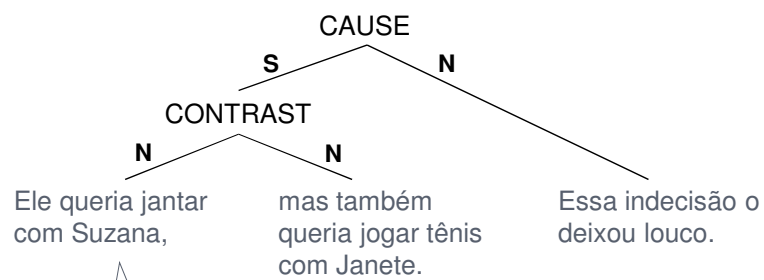
30

## RST: EXEMPLO



31

## RST: EXEMPLO

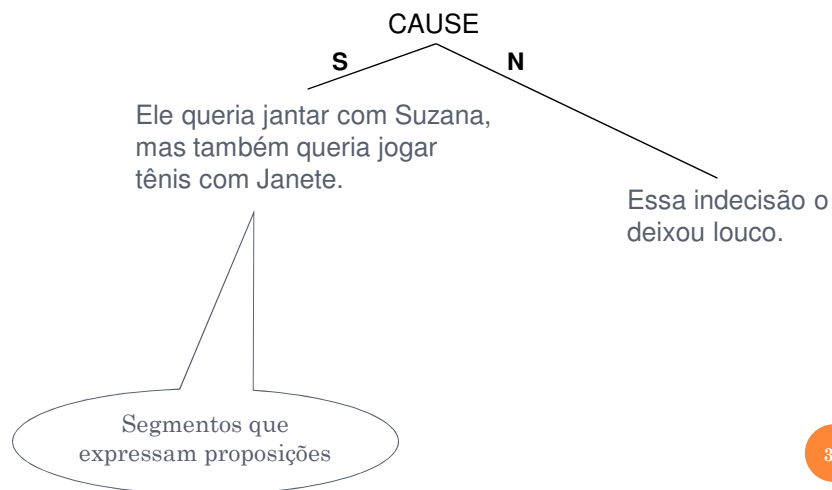


Segmentos que expressam proposições

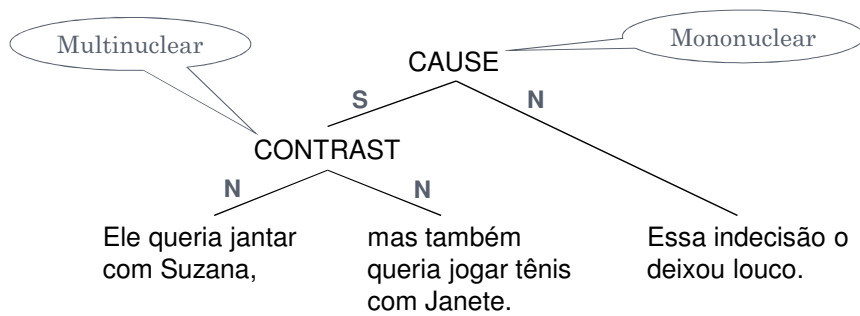
32



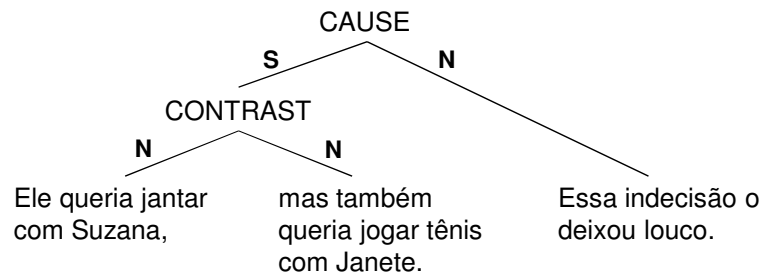
## RST: EXEMPLO



## RST: EXEMPLO

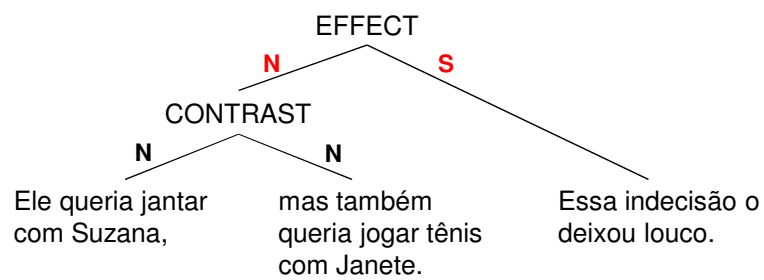


## RST: EXEMPLO



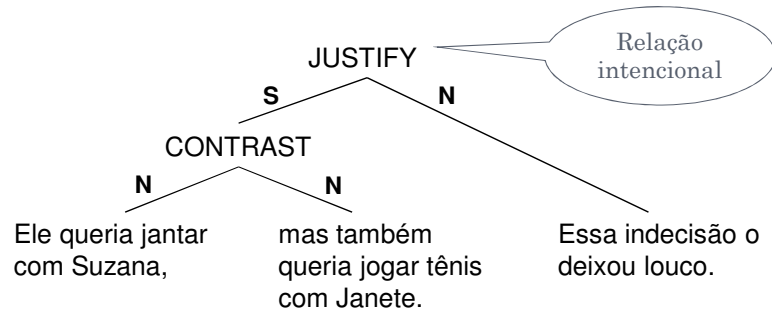
35

## RST: EXEMPLO



36

## RST: EXEMPLO



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 **ambiguidade**

*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

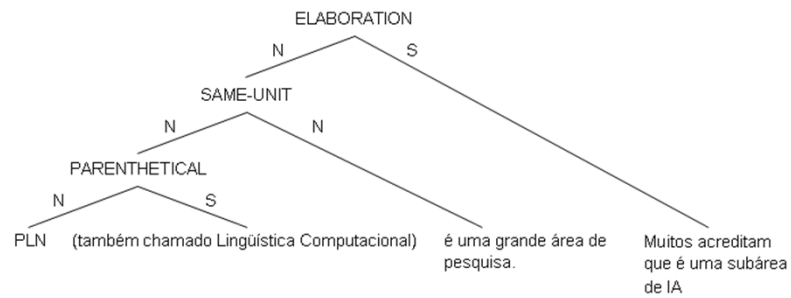
- **Relações**

- **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

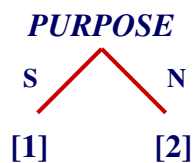
50

## 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.

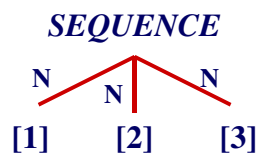


51

## 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 ambiguidade
  - 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

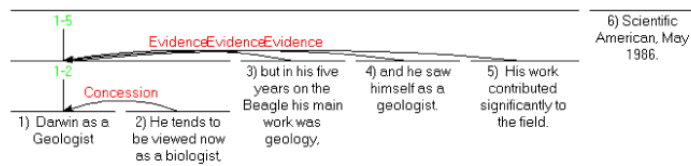
- Extensão do **conjunto de relações**: diversas possibilidades
  - Por exemplo, Marcu (1997)
    - Relações encaixadas (*embedded*)
- Propostas de **regras de segmentação**
  - Por exemplo, Carlson e Marcu (2001)
    - Manual de segmentação para textos escritos em inglês
- Manuais com **instruções para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

54

## ANÁLISE RETÓRICA

### ○ Restrições na análise

- Adjacência de segmentos, com algumas exceções

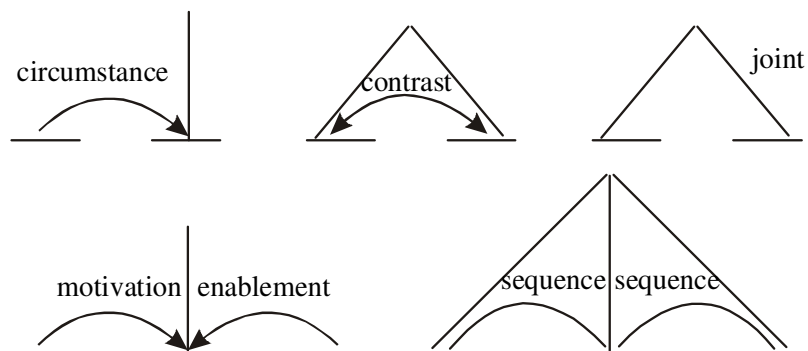


- Não compartilhamento de satélites
- Não sobreposição de segmentos

55

## ANÁLISE RETÓRICA

### ○ Esquemas possíveis de análise



56

## ALGUMAS REGRAS PARA SEGMENTAÇÃO (CARLSON E MARCU, 2001)

### o Segmentar

- Orações principais
- Orações sinalizadas por marcadores discursivos como *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 encaixados

57

## ALGUMAS REGRAS PARA SEGMENTAÇÃO (...CONTINUAÇÃO)

### o Não segmentar

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

58

## CONSTRUÇÃO DA ESTRUTURA RETÓRICA

### ○ Construção composicional

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

### ○ Forma de representação mais comum: **árvores**

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

59

## EXEMPLO DE ANÁLISE RETÓRICA

### ○ 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

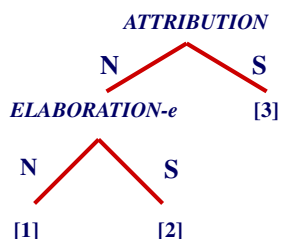
- Possível procedimento para análise: **análise incremental**
  - Relacionamento de todos os segmentos discursivos que compõem uma sentença, dois a dois
  - Relacionamento de todas as sentenças que compõem um parágrafo, duas a duas
  - Relacionamento de todos os parágrafos que compõem o texto, dois a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

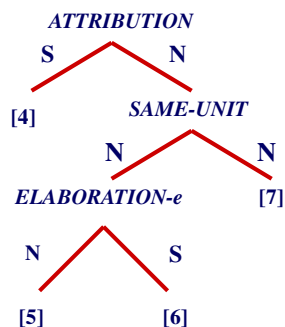


63

## 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?

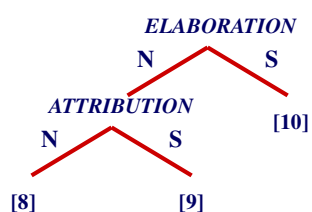
66



## OUTRA ESTRUTURA POSSÍVEL

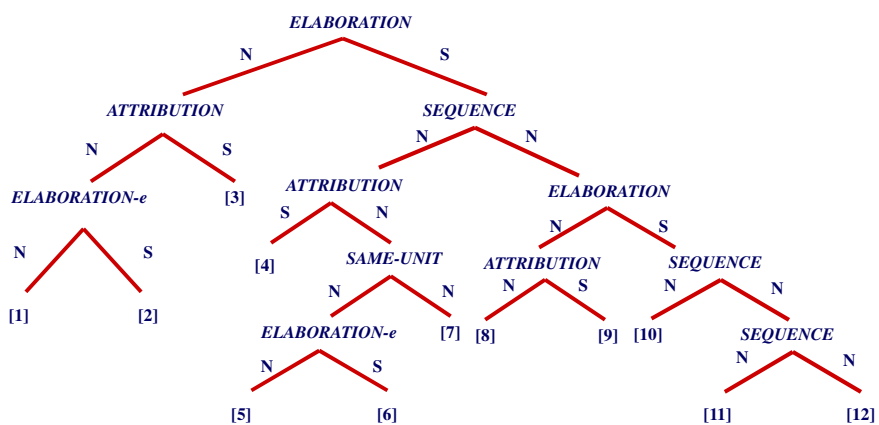
### ○ PARÁGRAFO 3: duas sentenças

[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

## EXERCÍCIO

- Parte 1 – análise de um conjunto de relações

ANTITHESIS	CONTRAST	JUSTIFY	PURPOSE
ATTRIBUTION	ELABORATION	LIST	RESTATEMENT
BACKGROUND	ENABLEMENT	MEANS	SAME-UNIT
CIRCUMSTANCE	EVALUATION	MOTIVATION	SEQUENCE
COMPARISON	EVIDENCE	NON-VOL-CAUSE	SOLUTIONHOOD
CONCESSION	EXPLANATION	NON-VOL-RESULT	SUMMARY
CONCLUSION	INTERPRETATION	OTHERWISE	VOL-CAUSE
CONDITION	JOINT	PARENTHETICAL	VOL-RESULT

71

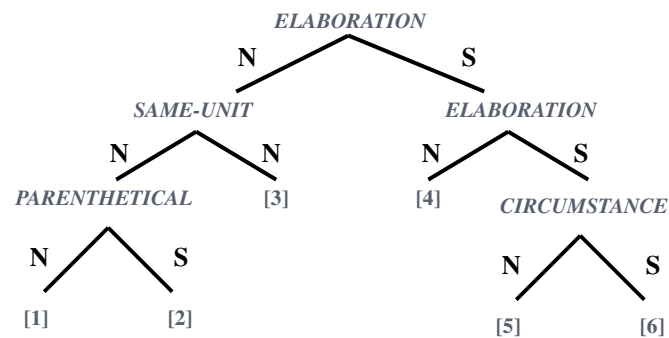
## EXERCÍCIO

- Parte 2 – análise do texto abaixo

A empresa Produtos Pirata Indústria e Comércio Ltda., de Contagem (na região metropolitana de Belo Horizonte), deverá registrar este ano um crescimento de produtividade nas suas áreas comercial e industrial de 11% e 17%, respectivamente. Os ganhos são atribuídos pela diretoria da fábrica à nova filosofia que vem sendo implantada na empresa desde outubro do ano passado, quando a Pirata se iniciou no Programa Sebrae de Qualidade Total.

72

[1] A empresa Produtos Pirata Indústria e Comércio Ltda., de Contagem [2] (na região metropolitana de Belo Horizonte), [3] deverá registrar este ano um crescimento de produtividade nas suas áreas comercial e industrial de 11% e 17%, respectivamente. [4] Os ganhos são atribuídos pela diretoria da fábrica à nova filosofia [5] que vem sendo implantada na empresa desde outubro do ano passado, [6] quando a Pirata se iniciou no Programa Sebrae de Qualidade Total.



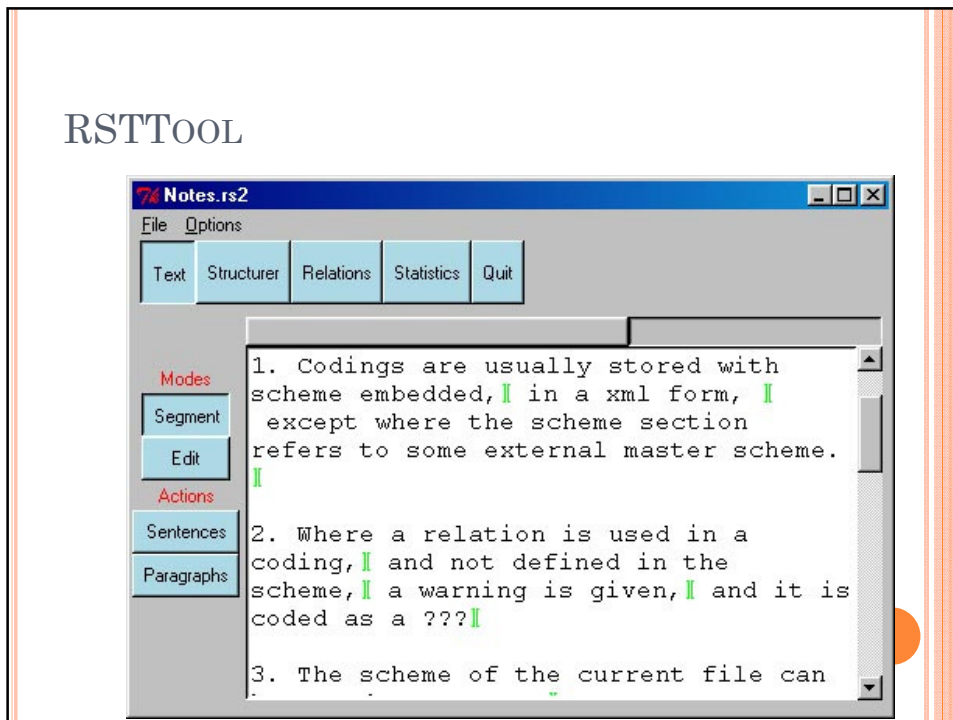
73

## FERRAMENTAS DE AUXÍLIO

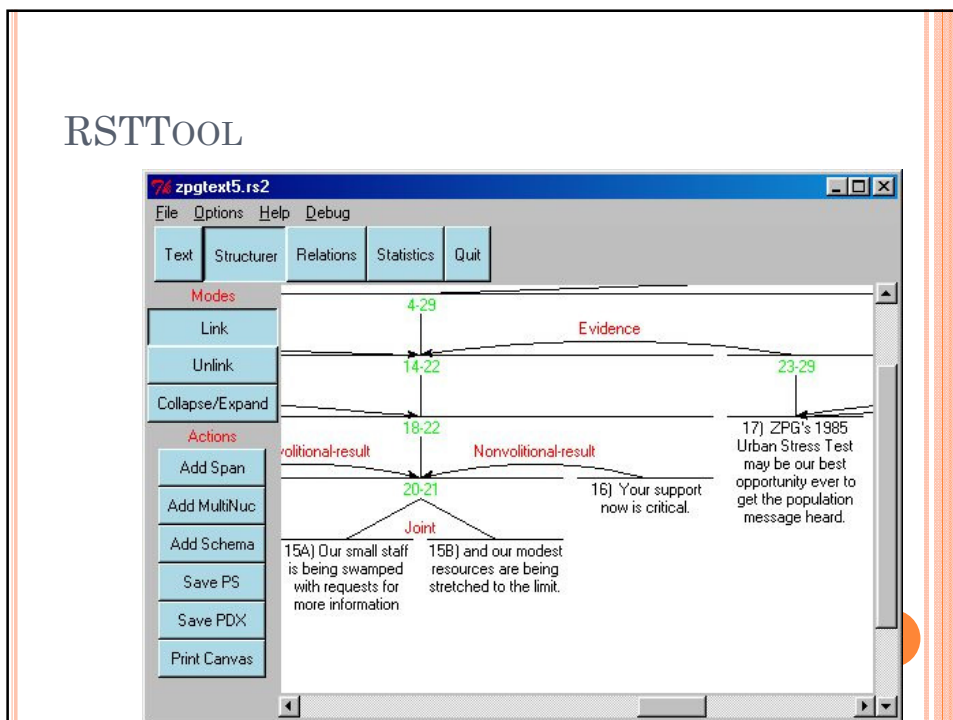
- *RSTTool* (O' Donnel, 1997)
- *RST Annotation Tool*: extensão da *RSTTool* (Marcu, 1997)
- *RST Toolkit* (NILC)

74

## RSTTOOL



## RSTTOOL



## CÓRPUS ANOTADOS

### ○ Português

- TeMário (Pardo e Rino, 2003; Seno, 2005)
  - 40 textos (jornalísticos)
- CorpusTCC (Pardo, 2005)
  - 100 introduções de teses e dissertações
- Retalho (Pardo e Seno, 2005)
  - Corpus de referência
  - 50 textos (20 jornalísticos e 30 científicos)
  - Protocolo de anotação retórica
- CSTNews (Aleixo e Pardo, 2008)
  - 140 textos jornalísticos

### ○ Inglês

- *RST Discourse Treebank, Penn Discourse Treebank, Discourse Relations Reference Corpus, etc.*

77

## PARSERS DISCURSIVOS

### ○ Português do Brasil

- DiZer (Pardo e Nunes, 2008)

### ○ Inglês

- Marcu (1997) e Soricut e Marcu (2003)
- Corston-Oliver (1998) e Microsoft
- E muitos outros vieram depois

### ○ Japonês

- Sumita et. al (1992)

### ○ Francês, grego, espanhol, etc.

78

## PARSING DISCURSIVO

- Tarefas envolvidas
  - **Determinação dos segmentos**
    - Deve ser sistemática
    - Uso de indicações lexicais, morfossintáticas, sintáticas, aprendizado de máquina
  - **Determinação das relações**
    - Marcadores discursivos, palavras e frases indicativas
    - Relações lexicais
    - Sintaxe
    - Aprendizado de máquina
  - **Construção de árvores RST válidas**
    - Critérios da teoria

79

## PARSING DISCURSIVO

- Determinação das relações
  - **Marcadores discursivos**
    - Vs. **marcadores sentenciais**

*Ele correu para casa e ligou para a mãe* (relação de sequência)  
*Ele e ela correram para casa* (--)
    - **Ambíguos**: não há mapeamento de um-para-um entre marcadores e relações
      - Várias listagens para várias línguas
      - Exemplo
        - *mas* → contraste, concessão, etc.
        - contraste ← *mas, entretanto, no entanto*
    - Marcadores **fracos**: *e*
    - Muitas relações **não são marcadas!!!**

80



## PARSING DISCURSIVO

- Exemplo para o CorpusTCC

Relação	Nro. rel. marcadas	% de rel. marcadas
ANTITHESIS	20	95,2
ATTRIBUTION	185	100
BACKGROUND	47	41,5
CAUSE	147	98,6
CIRCUMSTANCE	138	90,0
COMPARISON	11	100
CONCESSION	67	94,3
CONCLUSION	12	85,7
CONDITION	20	100
ELABORATION	1.010	60,0
ENABLEMENT	47	88,6
EVALUATION	14	93,3
EVIDENCE	3	20,0
EXPLANATION	23	76,6
INTERPRETATION	12	85,7
JUSTIFY	91	94,7
MEANS	60	90,9
MOTIVATION	16	84,2
OTHERWISE	2	100
PURPOSE	450	98,4
RESTATEMENT	17	85,0
RESULT	129	96,9
SOLUTIONHOOD	49	98,0
SUMMARY	4	100
CONTRAST	83	93,2
LIST	256	46,5
SEQUENCE	51	72,8

81

## PARSING DISCURSIVO

- Determinação das relações

- Palavras e frases indicativas

- Dependentes do gênero e tipo textual
  - Os resultados da pesquisa são...
  - O objetivo desse método é...
  - ... se consagrou campeão mundial.

- Relações lexicais, sintaxe

- Atirar → morrer (relação de causa), comprar pipoca → assistir um filme (relação de sequência)
- [I don't want a truck;] [I'd prefer a convertible.]
  - Indicações da relação de contraste: paralelismo sintático, negação na primeira oração, relação entre truck e convertible

82

## PÁGINA OFICIAL DA RST

The screenshot shows the official website for Rhetorical Structure Theory (RST). The page features a navigation menu on the left with the following items: HOME, INTRO TO RST, ANALYSES, RESEARCH TOPICS, TEXT GENERATION, BIBLIOGRAPHIES, TOOLS, PAGES IN FRENCH, PAGES IN PORTUGUESE, PAGES IN SPANISH, E-MAIL LIST, and SITE MAP and SEARCH. The main content area is titled "INTRO TO RST / RHETORICAL STRUCTURE THEORY/" and includes a welcome message and several paragraphs of introductory text. A text size selector is visible in the top right corner, with options for "LARGE TEXT SIZE" and "SMALL TEXT SIZE".

## CONCORRENTES DA RST

### o Vários outros “modelos relacionais”

- Hobbs (1979): entre os trabalhos pioneiros
  - *Result, Explanation, Parallel, Elaboration, Occasion*
- Jordan (1992): relações mais semânticas, com tipos e subtipos
  - Diferencial da semântica: desprovida de força argumentativa
- Marcu (1997): mais de 70 relações
- Kehler (2002): 3 relações (mais semânticas) são suficientes
  - *Resemblance, cause-effect* e *contiguity*
  - Mapeamento possível com RST
    - ELABORATION, LIST e CONTRAST são relações *resemblance*
    - CAUSE e RESULT, EXPLANATION e JUSTIFY são relações *cause-effect*
    - SEQUENCE é uma relação *contiguity*
- E muitos outros

## RELAÇÕES DE JORDAN (1992)

Relações	Tipo das Relações			
Identification	Detail	Authority	Modal (cont.)	
Classification		Determination		
Specification		Permission		
Appearance		Obligation		
Characteristics		Willingness		
Function		Desire		
Material		Time		Time
Parts		Before		
Active		After		
Passive		Simultaneous		
Agent	Inverted time			
Source	General	Elaboration	Text manipulation	
Assessment		Summary		
Basis		Repetition		
Cause		Paraphrase		
Effect		Forecast		
Emotive Effect		Transition		
Purpose		Collateral inversion		Special
Means		Concession		
Problem		Compatibility		
Solution		Contrast		
Possibility	Comparison			
Capability	Conditionals			
Correctness	Document structures	Other		
Propriety	Hypothetical-Real			
Necessity	Transition couplets			
Need	Accompaniment			
Completion	Circumstance			
Achievement	Inverted circumstance			
Future	Connection			
Intention	Enablement			
Mandate	Example			
	Extent			
	Location			
	Inverted Location			
	Manner			
	True			

85

## RST E OUTROS MODELOS: ALGUMAS POSSIBILIDADES

- Intenções e RST: sugestão de Moore e Paris (1993)
  - Mapeamento n-para-n entre intenções e relações RST

Intenções	Relações Retóricas
<ul style="list-style-type: none"> <li>▪ persuadir o leitor sobre uma proposição</li> <li>▪ persuadir o leitor a realizar uma ação</li> <li>▪ tomar o leitor competente para compreender algo</li> <li>▪ tomar o leitor competente para realizar uma ação</li> </ul>	EVIDENCE MOTIVATION BACKGROUND ENABLEMENT
<ul style="list-style-type: none"> <li>▪ capacitar o leitor a identificar algo</li> </ul>	CIRCUMSTANCE CONDITION CONTRAST ELABORATION PURPOSE SEQUENCE
<ul style="list-style-type: none"> <li>▪ fazer com que o leitor acredite em uma proposição</li> </ul>	CONTRAST ELABORATION

86

## RST E OUTROS MODELOS: ALGUMAS POSSIBILIDADES

- RST e semântica: método de Korelsky e Kittredge (1993)

Se a relação retórica EVIDENCE é observada entre duas proposições P1 e P2, em que P1 é o núcleo e P2 é o satélite, então:

- 1) se há um agente consciente de tal forma que P1 e P2 fazem referência a suas ações, então a relação semântica VOLITIONAL CAUSE se estabelece entre as proposições;
- 2) se não há um agente consciente, então a relação semântica NON-VOLITIONAL CAUSE se estabelece entre as proposições;
- 3) se P2 é uma proposição genérica, então a relação semântica ELABORATION se estabelece entre as proposições.

87

## RST E OUTROS MODELOS: ALGUMAS POSSIBILIDADES

- Outras visões

- A retórica já traz semântica (Hovy, 1991)
  - Deve-se explicitar isso?
    - Pode ocorrer uma grande proliferação de definições
    - Estruturas retóricas e semânticas podem não ser isomórficas
- A retórica já traz intenções (no campo “efeito” das definições, principalmente)
- Semântica é “mais livre”, podendo formar grafos (Dale, 1993)
  - Moser e Moore (1996): tornar as relações semânticas “parasitas” das retóricas, restringindo as possibilidades

88

## RST E DCG

- **Questão**

- **DCG**

- É possível lidar com RST?
  - Se sim, como?
  - Se não, por quê?