

Perguntas-respostas em textos escritos: uma análise no âmbito das relações discursivas

Amália Mendes¹, Pierre Lejeune^{1,2} & Carolina Nunes¹

¹ Centro de Linguística da Universidade de Lisboa, Universidade de Lisboa

² Faculdade de Letras da Universidade de Lisboa, Universidade de Lisboa

Abstract:

Question-answer pairs are typically associated to spoken discourse and to directive speech acts, although they are also found in written texts. We analyse contexts extracted from the CRPC-DB, a written subcorpus annotated with discourse relations in the PDTB-style. We focus on the nature of the question and of the answer in interactional contexts, but also in contexts where a single locutor poses the question and answers it. Contexts of question-answer pairs with a single locutor have a textual function related to the topic-comment structure of the text and involve a virtual locutor and a modalized assertion. We discuss the treatment of question-answer pairs in discourse banks following different theoretical frameworks and make a proposal that integrates these contexts in our annotation scheme. Our results contribute to theoretical proposals that focus on the pragmatic and textual functions of question-answer pairs, and their application to a new resource for the study of discourse in Portuguese.

Keywords: Discourse analysis, question-answer pair, discourse relation, PDTB

Palavras-chave: Análise do discurso, par pergunta-resposta, relação discursiva, PDTB

1. Introdução

Os pares pergunta-resposta (PR) estão tipicamente associados a contextos de oralidade e a atos de fala diretivos, em que o falante procura obter informação junto do seu interlocutor, podendo ainda ter outras funções ilocutórias (Barbosa et al., 2020). Os contextos de PR também ocorrem em textos escritos, com diversas funções¹. O nosso objetivo é analisar contextos de PR em português escrito, extraídos do Corpus de Referência do Português Contemporâneo (CRPC) *online* (Généreux et al., 2012)². Para uma análise desses contextos, teremos em consideração:

(i) A natureza da pergunta. Partiremos da definição habitual e restritiva do conceito de pergunta, como sendo um pedido ou a manifestação de um desejo de informação dirigido por um dos locutores ao seu interlocutor. Considerando a natureza dos dados de escrita, esta definição tem sido alargada para incluir perguntas retóricas, isto é, frases interrogativas com função assertiva e perguntas postas e respondidas por um único locutor, que não são portanto um pedido de informação. A natureza global vs. parcial da pergunta é ainda outro aspeto a ter em conta.

(ii) A natureza da resposta. À semelhança do ponto (i), a resposta consistiria em providenciar a informação solicitada. No entanto, os segmentos linguísticos que seguem a pergunta podem consistir em dizer que a resposta não é conhecida, ou é impossível de obter, ou ainda serem comentários à pergunta, ou comentários que, embora não constituam uma resposta direta, estão relacionados com o tópico da pergunta. Veja-se o exemplo (1).

¹ Usamos o termo *par pergunta-resposta* para todas as estruturas compostas por uma frase interrogativa seguida de resposta; discutimos no artigo os valores ilocutórios de cada segmento do par.

² <http://alfclul.clul.ul.pt/CQPweb/crpcf16>



- (1) PÚBLICO - Em sua opinião, quais são os principais obstáculos à melhoria das previsões meteorológicas? (...)

GODWIN OBASI - Antes de mais, deixe-me dizer-lhe que as previsões meteorológicas de curto e médio prazo (respectivamente de 1-3 dias e 4-10 dias) melhoraram enormemente a sua qualidade nos últimos anos. (pu92214)

(iii) a existência de um ou dois locutores, e ainda o conceito de enunciador. Na nossa análise, iremos organizar os dados distinguindo entre os contextos verdadeiramente interacionais, que envolvem duas entidades, e contextos com um único locutor, embora os dados mostrem que a distinção nem sempre é óbvia. Um único locutor pode ainda convocar vários enunciadores (no sentido de Ducrot (1972)) enquanto diferentes pontos de vista, que podem ser partilhados ou opostos (Culioli, 1990; Ducrot, 1972; Kerbrat-Orecchioni, 1999).

A análise dos contextos de PR insere-se num projeto de constituição de um corpus anotado com relações discursivas (também chamadas relações retóricas), isto é, relações conceptuais que se estabelecem, a nível intrafrásico ou interfrásico, entre elementos proposicionais, como a relação de CAUSA, EXPANSÃO, CONTRASTE. Assim, pretende-se analisar de que forma os pares PR são tratados nas teorias de relações discursivas e a sua aplicação em propostas de anotação em bancos do discurso (*discourse bank*), isto é, em corpora anotados com relações retóricas entre as orações e frases dos textos. O banco do discurso para o português, CRPC Discourse Bank (CRPC-DB) segue o modelo de anotação do Penn Discourse Treebank - PDTB (Prasad et al., 2008) com inclusão de alguns elementos da versão PDTB 3.0 (Webber et al., 2018). Assim, o nosso trabalho terá igualmente em consideração:

(iv) a função do par pergunta-resposta em termos de relação discursiva, isto é, de que forma o par-pergunta resposta contribui para a coesão e coerência e que funções desempenha.

Desta forma, a partir de um trabalho de anotação de relações retóricas em textos escritos, analisamos a complexidade dos contextos de pares PR em textos escritos, do ponto de vista dos locutores envolvidos, da identificação dos enunciadores, da natureza da frase interrogativa e do contexto seguinte, das funções pragmáticas e discursivas das interrogativas e das estratégias utilizadas pelos autores. Aplicamos esta análise à identificação das relações discursivas que compõem um texto e propomos um tratamento no âmbito da anotação dessas relações discursivas.

Começamos por rever algumas propostas de anotação de PR em bancos do discurso existentes na secção 2, apresentamos o CRPC-DB na secção 3, analisamos os contextos de pares de PR em 4 e propomos uma anotação destes casos em 5, antes de concluirmos na secção 6.

2. Os contextos de pares PR em bancos do discurso

Nesta secção, apresentamos uma revisão das opções que foram tomadas em vários projetos e vários quadros teóricos para a anotação de pares de PR. Começamos pelos corpora anotados de acordo com o PDTB. O Penn Discourse Treebank – PDTB contém textos escritos do *Wall Street Journal* e está anotado com relações discursivas (também designadas de relações de coerência ou relações retóricas) a nível intra e inter-frásico (Prasad et al., 2007). As relações discursivas marcam a relação entre dois argumentos: “discourse relations hold between two and only two (possibly discontinuous) spans of text that can be interpreted as propositions, eventualities, beliefs, etc. (what Asher [1993] has called **abstract objects**)” (Prasad et al., 2014:921). Uma relação discursiva é explícita quando existe um conector a relacionar os dois argumentos (tipicamente uma



conjunção, um advérbio ou um conector), implícita quando não ocorre um conector (sendo então sugerido um conector na anotação) ou AltLex, quando a relação entre os dois argumentos é assegurada por uma lexicalização alternativa que desempenha uma função idêntica à do conector. Para além destes três tipos de relação, o PDTB anota relações entre entidades (EntRel), isto é, casos em que o segundo argumento fornece mais informação sobre uma entidade introduzida no primeiro argumento, e existe ainda a possibilidade de considerar que não existe uma relação discursiva entre os dois argumentos (NoRel). As relações discursivas do tipo explícito, implícito e lexicalização alternativa são associadas a um sentido, p.e. condição, razão, temporal, especificação (a lista hierarquizada dos sentidos da versão PDTB 3.0 é apresentada em anexo). O PDTB adota uma “theory-neutral approach to the annotation, making no commitments to what kinds of high-level structures may be created from the low-level annotations of relations and their arguments” (Prasad et al., 2007:2961).

No PDTB 2.0, os pares de PR não são tratados de forma específica. Assim, se existir um conector a introduzir o segundo argumento (a resposta), a relação será anotada como uma relação de tipo explícito e será atribuído um sentido. No exemplo (2), a frase interrogativa é o argumento 1 e o conector *because* introduz a resposta, que constitui o segundo argumento. Entre o primeiro e o segundo argumento existe uma relação com o sentido de Reason (Razão). No exemplo (3), a relação entre os dois argumentos é de tipo implícito, uma vez que não há nenhum conector explícito, e a anotação propõe um conector que expressa o sentido de Expansão inferido a partir do contexto. Os exemplos extraídos de corpora anotados apresentam o conector sublinhado, o argumento 1 em itálico e o argumento 2 a negrito. Após o exemplo, indica-se a fonte, o tipo de relação e o sentido atribuído.

- (2) *Why constructive?* Because despite all the media prattle about comedy and politics not mixing, they are similar in one respect: Both can serve as mechanisms for easing tensions and facilitating the co-existence of groups in conflict. (wsj-2369) [Explicit;Contingency:Cause:reason]
- (3) *How does a nice new tax, say 5% on any financial transaction sound?* (Implicit=*indeed*) **That ought to make sure we’re all thinking for the long term.** (wsj-0118) [Implicit; Expansion]

Num corpus de notícias como o PDTB, é possível encontrar pares de PR produzidos no âmbito de entrevistas transcritas, mas há igualmente contextos de pares de PR em que o autor das notícias produz a frase interrogativa e também a resposta. No corpus multilingue TED-MDB, um corpus paralelo anotado de acordo com o estilo do PDTB em 6 línguas (Zeyrek et al., 2018, 2019), estes últimos são aliás os únicos tipos encontrados, uma vez que as TED Talks são apresentações orais, preparadas, perante um público ao vivo e sem intervenção da audiência. Nelas, o falante frequentemente usa interrogativas para captar a atenção da audiência. Estes pares de PR são portanto encontrados quer em dados escritos quer em dados de oralidade. No TED-MDB, estas estruturas são anotadas com um novo sentido de topo na hierarquia do PDTB, chamado Hypophora, figura pragmática da oralidade com uma função apelativa (Lanham, 1991; Mayoral, 1994). Se necessário, os contextos são etiquetados com um segundo sentido, como Razão. A anotação destes pares de PR centra-se, no TED-MDB, na sua natureza interrogativa. Assim, a palavra interrogativa é etiquetada como uma lexicalização alternativa que introduz a pergunta. Seguindo as regras de anotação do PDTB, a pergunta é portanto o segundo argumento, enquanto a resposta é o primeiro argumento. Apresentam-se em (4) e (5) exemplos extraídos de Zeyrek et al. (2018). Em interrogativas globais (ou polares), o auxiliar invertido ou o elemento interrogativo são anotados como uma lexicalização alternativa em inglês. Nas línguas românicas, as interrogativas globais podem ser expressas apenas pela entoação e estes casos são anotados no corpus como uma relação implícita com o sentido Hypophora, como em (5).



Uma secção do Penn Treebank foi anotada no quadro da Rhetorical Structure Theory - RST (Carlson & Marcu, 2001; Mann & Thompson, 1988). No quadro da RST, considera-se que “the text structure conveys propositions, and propositions conveyed in this way are called ‘relational propositions’ (Mann & Thompson, 1988: 259), sendo que estas proposições relacionais não são necessariamente expressas por orações e podem não ser assinaladas por nenhum elemento. Neste corpus com anotação discursiva no quadro da RST, os pares PR são anotados com sentidos específicos (neste quadro teórico uma pergunta pode não ser realizada como frase interrogativa), em combinação com os conceitos de núcleo e satélite. Os contextos são anotados como Question-Answer-N quando a resposta é o núcleo (i.e, quando a resposta é entendida como mais importante do que a pergunta), como Question-Answer-S quando a resposta é o segmento satélite, e simplesmente como Question-Answer quando a relação é multinuclear (não é identificado um núcleo e um satélite). Existe uma categoria adicional para as perguntas retóricas, consideradas mononucleares. As categorias para pares de PR são incluídas num dos 16 tipos de relações retóricas, o tipo “Tópico-Comentário”, que inclui: “problema-solution, question-answer, statement-response, topic-comment, comment-topic, rhetorical-question” (Carlson & Marcu, 2001:33). A lista completa é apresentada no Anexo 2.

A natureza dos dados e das opções teóricas nos vários projetos é muito diversa e determina até certo ponto diferentes abordagens na fase de anotação de pares PR. O corpus STAC é verdadeiramente interacional e não contém perguntas com função fática nem perguntas retóricas. O corpus TED-MDB é oral mas não existe interação e as únicas interrogativas encontradas são produzidas e respondidas pelo mesmo locutor. Nos corpora anotados no quadro do PDTB e da RST, os dados são notícias de jornal, pelo que podem ser interacionais ou não. Finalmente, enquanto, no TED-MDB, no corpus STAC e no Penn Treebank anotado no quadro da RST, os pares de PR são anotados com sentidos retóricos, no PDTB 3.0 foi adicionado um novo tipo de relação especificamente para os casos de Hypophora.

3. O corpus

A nossa discussão das relações de coerência nos pares PR é baseada nos dados encontrados no CRPC-DB, um corpus do português anotado com relações discursivas de acordo com o quadro do PDTB. O CRPC-DB está atualmente em desenvolvimento e contém textos de diferentes géneros (textos jornalísticos, literários, didáticos e científicos) do corpus PAROLE³, por sua vez um subcorpus escrito do Corpus de Referência do Português Contemporâneo (Généreux et al., 2012).

Ao contrário do PDTB, em que cada conector é anotado por sua vez na totalidade dos dados, no CRPC-DB o anotador lê todo o texto e anota as relações discursivas identificadas, sem haver pré-anotação de elementos lexicais. Este método garante que o anotador não está condicionado para a identificação de certas relações explícitas. Uma avaliação de 3 estratégias diferentes de processo de anotação (Sharma et al., 2013) mostra que a abordagem em que o anotador procede com a anotação de um texto de cada vez, quer anotando as relações sequencialmente, quer anotando primeiro as explícitas e depois as implícitas, tem melhores resultados do que a abordagem do PDTB. Os textos do CRPC-DB são anotados sequencialmente: o anotador identifica e anota todas as relações que encontra em cada texto. A anotação é realizada por um anotador treinado e revista por um anotador com experiência. A anotação segue os tipos de relação do PDTB 2.0 (Explicit, Implicit, AltLex, EntRel and NoRel), como exemplificado nos exemplos (7)-(11). Cada relação dos tipos Explicit, Implicit e AltLex) é etiquetada com um dos sentidos da hierarquia do PDTB 3.0 (ver lista no anexo 1).

³ O corpus PAROLE está disponível no catálogo ELRA: http://catalogue.elra.info/en-us/repository/browse/ELRA-W0024_01/



- (7) As previsões meteorológicas de curto e médio prazo registaram grandes melhorias nos últimos anos, **mas é necessário continuar a investir na rede de estações de recolha à escala mundial** (pu92214.txt) [Explicit; Comparison:Concession:arg2-as-denier]
- (8) *As empresas atingidas agora pela operação policial faziam ampla concorrência às lojas francas.* (Implicit=para tal) **Vendiam mais barato e eram mais rápidas na satisfação dos pedidos dos clientes** (dn85705.txt) [Implicit; Contingency:Manner:arg2-as-manner]
- (9) *A Federação de Pais Portugueses em França solicitou uma audiência ao primeiro-ministro com o objetivo de lhe expor a situação da comunidade portuguesa naquele país e «exigir a concessão do passaporte europeu aos emigrantes que aí residem»* (dn86307.txt) [AltLex; Contingency:Purpose:arg2-as-goal]
- (10) *O Banco Comercial Português na sua qualidade de co-líder do conjunto de instituições financeiras que tomaram a seu cargo a colocação da operação de privatização da Portugal Telecom, S.A., que vai agora decorrer, não poderia deixar de o convidar a investir na aquisição de ações do operador de serviço básico de telecomunicações portuguesas. **A Portugal Telecom é a terceira maior empresa nacional e uma das mais sólidas e rentáveis empresas de serviços de Portugal.*** (banco001.txt) [EntRel]
- (11) *“Mas estava escrito que o Paços de Ferreira não poderia vencer este encontro, que bem o merecia, visto a sorte não estar com a equipa. **Arbitragem sem quaisquer problemas, apesar do três «amarelos» mostrados.**”* (bo86102) [NoRel]

Indica-se na Tabela 1 a dimensão atual do CRPC-DB, por tipo de relação, e na Tabela 2 a frequência dos 4 sentidos principais, por tipo de relação. O CRPC-DB contém neste momento 200 textos anotados e um total de 7042 relações anotadas. Os dados da Tabela 1 permitem desde já apontar que as relações implícitas são mais frequentes do que as explícitas, o que distingue o português dos dados do inglês no PDTB). Os dados confirmam os resultados da anotação do corpus multilingue TED-MDB, em que o subcorpus português apresentava frequência elevada de relações implícitas, comparado com o inglês ou com o turco, por exemplo (Zeyrek et al., 2019). A Tabela 2, por sua vez, permite observar que o sentido mais frequente (considerando apenas os 4 sentidos do topo da hierarquia) é o de Expansão, com uma frequência muito superior aos outros 3. Permite igualmente perceber que as relações implícitas são mais frequentes do que as explícitas no caso único do sentido Expansão. O CRPC-DB continua em desenvolvimento e em revisão.

Tipo de relação	Frequência
Explicit	3059
Implicit	3213
NoRel	349
Altlex	242
EntRel	179
Total	7042

Tabela 1: Total de relações no CRPC-DB, por tipo de relação



	Explicit	Implicit	AltLex	Total
Expansion	1456	2238	73	3767
Contingency	597	441	94	1132
Comparison	583	375	18	976
Temporal	423	159	57	639
Total	3059	3213	242	6514

Tabela 2: Número de relações discursivas no CRPC-DB, por tipo de relação e sentido

4. Análise

O CRPC-DB não contém, por enquanto, dados de oralidade, pelo que não há casos de transcrição de interações orais em tempo real. O corpus tem dados comparáveis com os do PDTB (notícias do Wall Street Journal) e com os do banco do discurso turco (notícias e novelas) (Zeyrek & Webber, 2008). No entanto, e como já referimos, as notícias podem incluir transcrições de entrevistas e também frases interrogativas com função fática de captar a atenção do leitor. Nesta secção, vamos observar e analisar o nosso corpus para identificar os tipos de interrogativas encontrados nos nossos dados. Iremos explorar algumas questões já identificadas na introdução, que são relevantes quando se consideram contextos com interrogativas, como: quem faz a pergunta e quem responde; a identificação mais ou menos óbvia do enunciador; de que forma a pergunta se relaciona com o co-texto; se existe uma resposta válida à pergunta.

No caso das interrogativas prototípicas, o locutor faz um pedido de informação ao seu interlocutor. Usa-se, nestes contextos, o termo de “interrogativas puras”. As interrogativas podem ser globais (ou polares): o falante ignora o valor de verdade da proposição e a resposta fornece esse valor de verdade; ou interrogativas parciais: existe, na proposição, uma variável por identificar e a resposta fornece o valor da variável dessa proposição. Em ambos os casos estamos perante uma proposição única: isto é, a resposta não acrescenta uma nova proposição.

A literatura identifica outras funções para as construções interrogativas, que não envolvem um pedido de informação. Assim, as interrogativas podem ser atos diretivos de pedidos, que mitigam a força ilocutória, por delicadeza (Lopes, 2018); podem ser retóricas quando constituem um ato de fala indireto em que uma asserção é apresentada sob a forma de pergunta; podem ainda ter valor avaliativo e serem atos de fala expressivos (interrogativas de surpresa) (Barbosa et al., 2020). Interessa-nos observar os valores das interrogativas no nosso corpus, do ponto de vista das suas funções pragmáticas e retóricas. Quanto à natureza da pergunta, tomaremos uma definição alargada, que abarca não só as interrogativas que formulam um pedido de informação, mas também outros valores. Analisaremos não só contextos em que a resposta fornece a informação pedida, mas ainda outros casos, à semelhança de autores que têm questionado uma visão restrita do conceito de pergunta (veja-se, por exemplo, o caso das perguntas retóricas (Borillo, 1981)) e ainda outras propostas mais alargadas, em que a resposta constitui de alguma forma uma “reação” à pergunta (Grésillon & Lebrave, 1984).

Na nossa análise, iremos organizar os dados distinguindo entre os contextos verdadeiramente interacionais, que envolvem dois locutores, e contextos com um único locutor. Os dados mostram que esta distinção está longe de ser uma categorização clara. Um único locutor pode ainda convocar vários enunciadores (Ducrot, 1972), com pontos de vista partilhados ou opostos. Teremos sempre em atenção a função do par de PR em termos de relação discursiva, isto é, de que forma o par de PR contribui para a coesão e coerência e que



funções desempenha. Para entender de que forma a pergunta se relaciona com o co-texto, apresentamos exemplos do corpus com alguma extensão.

4.1 Contextos com dois locutores

No CRPC-DB encontramos contextos de entrevistas, que são transcrições revistas de interações orais (por vezes as entrevistas são feitas por escrito, mas não temos qualquer indicação desse tipo). Estes são contextos com interação entre dois locutores, em que um faz a pergunta e o outro responde. Estes casos são ilustrados pelo exemplo (12): o entrevistador pede uma informação e o entrevistado fornece uma resposta apropriada (damos uma versão encurtada da resposta). Como os exemplos da secção 4 não pretendem discutir as opções de anotação, marca-se apenas em itálico a pergunta e o segmento linguístico seguinte.

- (12) A. *E como se vão defender os agricultores do quase certo desastre que se avizinha?*
B. *“Em Trás-os-Montes, os seguros de colheita quase não existem, porque as companhias fogem a isso ou só os fazem se forem acompanhados por seguros de outro tipo”.* (pu92205)

No entanto, na maioria dos casos, a resposta não é uma resposta direta, mas tem relação com a pergunta. Por exemplo, no exemplo (1), aqui repetido em (13), o entrevistado comenta a pergunta, dando informação adicional antes de apresentar a resposta propriamente dita:

- (13) *PÚBLICO - Em sua opinião, quais são os principais obstáculos à melhoria das previsões meteorológicas ? (...)*
GODWIN OBASI - Antes de mais, deixe-me dizer-lhe que as previsões meteorológicas de curto e médio prazo (respectivamente de 1-3 dias e 4-10 dias) melhoraram enormemente a sua qualidade nos últimos anos. (pu92214)

Nos dois exemplos anteriores, a identidade do locutor que faz a pergunta e do locutor que responde é conhecida e clara a partir do contexto. Mas nem sempre é o caso. Enquanto (13) é um caso de entrevista e segue o formato deste género textual ao identificar o autor de cada turno de fala, o exemplo (14) é extraído de uma notícia que usa várias estratégias para relatar a informação providenciada pelo entrevistado: quer como discurso direto, como acima, quer como discurso reportado que é elusivo quanto à identidade do enunciador. Presume-se que, em (14), o autor faz a pergunta, mas não é claro se a pergunta foi de facto parte da interação real com o entrevistado ou se a pergunta é usada na notícia para introduzir parte da informação, como uma interação construída. A pergunta marca uma rutura em relação ao contexto anterior (*mas*) e introduz um novo desenvolvimento do tópico. A primeira frase após a pergunta (*a água não pode cair de qualquer maneira*) não é marcada como uma citação, mas antes como um relato indireto da fala do entrevistado e é seguida de uma citação direta com informação mais detalhada. O contexto ainda pode ser considerado como um caso de interação, embora os termos exatos da interação sejam modelados pelo autor.

- (14) Isto é, espera-se da natureza o milagre de fazer cair dos céus, nas próximas dez semanas, sensivelmente o dobro da água dos últimos três meses. *Mas ... chover como? A água não pode cair de qualquer maneira. “É bom que chova bem, forte, durante três semanas, por exemplo, mas não em grandes batedas” (...), explicou Nuno Moreira.*” (pu92205)

No exemplo (15), a pergunta é feita por jornalistas (embora o autor, sendo jornalista, possa estar incluído no coletivo de locutores). A pergunta contém uma asserção implícita: se Theunisse é penalizado, Delgado



também o deveria ser, e a oração gerundiva fornece a justificação para a asserção. O nível de quase certeza da pergunta faz com que partilhe propriedades com as perguntas retóricas. Não é claro se a pergunta é dirigida a alguma entidade específica e pode ser apenas a expressão de uma falta de compreensão em relação às regras. Por conseguinte, não é esperada nenhuma resposta e não há uma resposta válida no contexto.

- (15) O corredor holandês (...) foi penalizado com dez minutos e passou ao 11º lugar por a segunda análise de urinas ter confirmado a presença de Codeína - uma substância (...) proibida pelo regulamento anti-"doping" da União Ciclista Internacional (UCI).

Então, porquê Theunisse e não Delgado? - perguntavam alguns jornalistas, sublinhando que o Probenicidol deverá ser incluído na lista dos produtos proibidos da UCI no princípio de Agosto. Este medicamento, de venda livre nas farmácias, é normalmente utilizado com anti-inflamatório nos casos de gota (...) (ex88420)

4.2 Contextos com um só locutor

Noutros contextos do corpus, o autor faz uma pergunta para a qual fornece ele próprio uma resposta, correspondendo portanto aos casos de hypophora discutidos na secção 2, e definidos como tendo uma função fática. Estamos perante textos escritos que não reproduzem uma interação entre dois locutores (como as entrevistas dos exemplos anteriores). Estes contextos são frequentes em textos didáticos, como os manuais escolares. Assim, no excerto de um manual escolar de geografia em (16), o autor apresenta uma pergunta que não constitui, naturalmente, um verdadeiro pedido de informação a um interlocutor, e que constrói antes o que poderia ser a pergunta de um leitor virtual. Nesse sentido, é útil o conceito de **dupla locução**, uma vez que o par de PR não remete para dois locutores reais (objeto da análise da pragmática), mas sim para um “segundo locutor *virtual*” (Grésillon & Lebrave, 1984: 123). Quanto ao segmento que segue a pergunta, não é estritamente uma resposta, mas sim a introdução de uma nova secção da obra. A mesma informação poderia aliás ter sido apresentada como uma asserção, por exemplo “*vamos procurar as razões deste facto através de um estudo pormenorizado de cada continente*”. A pergunta introduz assim um novo tópico ou o desenvolvimento de um tópico precedente e consiste numa forma específica de estruturar informação conhecida e informação nova: “la scission de S en un couple question-réponse aurait l’avantage de faire apparaître plus clairement le rapport topic/comment.” (Grésillon & Lebrave, 1984: 119).

- (16) desta observação poderás deduzir que existem áreas favoráveis e áreas desfavoráveis à presença do Homem. - *Quais as razões deste facto? Vamos procurá-las através de um estudo pormenorizado de cada continente.* 2.1 Europa (geo9_1)

Outro exemplo em que um único locutor faz a pergunta e responde é apresentado em (17). A pergunta é uma rutura com o co-texto anterior no sentido em que **modaliza** a asserção anterior atribuída a Koestler (‘a sociedade humana é uma realidade cujas manifestações típicas são inequivocamente patológicas’) e assegura o desenvolvimento textual do segmento seguinte, onde várias perspetivas teóricas são referidas. Essa função de desenvolvimento textual é designada por **função textualizante** da resposta, que desenvolve a pergunta e assegura a progressão do texto (Grésillon & Lebrave, 1984:125).

- (17) Koestler, em 1969, levanta a hipótese de a sociedade humana, desde a sua origem, ser uma realidade cujas manifestações típicas são inequivocamente patológicas (...). *O homem social terá sido desde sempre (...) um ente enfermo e condenado à degenerescência vital, o que, a longo prazo, poderia significar a sua condenação inapelável como forma de vida? Esta pergunta, posta explícita ou implicitamente, levou a uma intensa, diríamos frenética, busca de explicações e justificações.* Saía-



se assim da posição equívoca de que o homem não devia obedecer à natureza, mas sim vencê-la, para outra muito mais humilde (...) (dn81201)

Em (18), a pergunta é igualmente uma rutura com a frase precedente (veja-se o uso de *afinal*), havendo uma sequência de duas perguntas. Contrariamente a (17), em (18) as perguntas desenvolvem o trecho anterior a partir de uma asserção implícita (não existe muita diferença entre os primatas e os humanos) e estão numa relação de Resultado com valor pragmático (Contingency: Cause:Result+Speech Act) com o co-texto: 'é devido a esta informação que estou em condições de fazer a pergunta'. O segmento seguinte não oferece resposta mas abre a apresentação de uma nova abordagem teórica.

- (18) De igual modo, os livros de Desmond Morris (...) estabelecem uma aproximação um tanto desmoralizadora entre os animais simpáticos que visitamos nas jaulas dos jardins zoológicos e os seres humanos vencedores da gravidade e protagonistas da espantosa aventura espacial. *Que distância separa afinal o homem médio do chimpanzé Nim, do chimpanzé Washoe ou da sua congénere Sara? Que fluido subtil distingue o eleitor médio canadiano ou luxemburguês de Koko, a gorila a quem M. Patterson conseguiu, com paciência evangélica, ensinar mais de 400 sinais e símbolos? Serge Moscovici, ao pôr frontalmente a questão de que o homem civilizado, habitante de uma sociedade antinatural, é um animal doméstico que perdeu ao longo do processo de domesticidade a sua qualidade de ser livre, coloca um problema pleno de sugestões e até de confrontações de natureza ética.* (dn81201)

Também em (19) existe um bloco de perguntas e uma rutura com a sequência anterior, neste caso assinalada pela conjunção *mas*. Neste exemplo, a pergunta incide sobre uma asserção (o stress está presente em todas as sociedades humanas) e apresenta informação nova (a condenação do homem como ser vivo). O segmento seguinte não fornece resposta direta, mas tem uma função textualizante ao focar essa condenação através da insuficiência dos tratamentos.

- (19) O estudo das sociedades primitivas (das quais não se encontram ausentes nem a violência, nem a crueldade, nem os outros pecados mortais da espécie, que Koestler assinalou) traz-nos unicamente a revelação de que nelas não tinham lugar os psiquiatras, mas em todas se encontra o feiticeiro (o psiquiatra colectivo ou o Moreno de então) e, pior que isso, o soma pacificante de que Aldous Huxley falara no seu Admirável Mundo Novo.
- Este mal do século, como lhe chamou Hans Selve, o stress, é doença que já existia nas sociedades primitivas, que existe hoje e que Huxley previa existir no seu mundo mecanizado e artificial, dominado por uma tecnologia genética que afeiçoava o organismo humano às tarefas sociais.
- Mas esse stress que é afinal o resultado de uma incapacidade de adaptação do homem ao meio-ambiente civilizacional não significará a sua condenação como ser vivo?*
- O seu carácter epidémico e o facto de ter atingido os centros nervosos e a própria vida psicológica de base, não pode significar que a natureza decidiu que ele começa a não ter condições de continuar a viver, tal como um carro que começa a diminuir a velocidade por falta de combustível?*
- As tentativas de resposta modernas na luta contra o stress: a picanálise, o yoga, a meditação transcendental (ou o recurso à religião, imprudentemente desacreditada pelos livres-pensadores do século XVIII) parecem não ser suficientes, nem mesmo apoiadas pela tecnologia química, com os barbitúricos, relaxantes ou inibidores beta.* (dn81201)



O exemplo (20) ilustra um caso particular, uma vez que a pergunta e o segmento seguinte são produzidos por um locutor que o jornalista cita e não pelo próprio jornalista autor da notícia. Mais uma vez, a pergunta não é um pedido de informação, pelo contrário serve para estabelecer o tópico (informação partilhada) e fornece informação nova (comentário).

- (20) Manuel Baptista, pastor da povoação de Suções, perto de Mirandela, confessou-nos que nunca viu ano como este: "Não há água nem há que comer para elas" - e apontou para as 130 ovelhas do rebanho à sua guarda - "nem para a gente. Olhe que até comem os ramos das oliveiras, senhor!" Os bovinos não terão melhor sorte.
As forragens não duram sempre e pasto quase não há. *O senhor viu as vacas na televisão, que não se podiam levantar com a fraqueza? É o que vai acontecer aqui*", concluiu Manuel Batista. (pu92205)

Em suma, realçamos as seguintes propriedades dos pares de PR nos géneros textuais escritos (excluindo as transcrições de interações reais):

- (a) os pares de PR na escrita envolvem frequentemente um locutor virtual, criando o que pode ser designado como "dupla locução";
- (b) a pergunta estabelece uma rutura com o segmento anterior, frequentemente expresso por um marcador discursivo (*mas, afinal*), que introduz um tópico (embora em casos mais raros a pergunta desenvolva e feche o segmento anterior);
- (c) os pares de PR assumem uma função de estruturação da informação, ao separarem tópico e comentário;
- (d) a resposta desempenha uma função textualizante, ao desenvolver o tópico introduzido pela pergunta;
- (e) embora os exemplos aqui apresentados não sejam verdadeiras perguntas retóricas, neles a pergunta está frequentemente associada a uma asserção implícita fortemente modalizada que as aproxima das perguntas retóricas: "Si ce potentiel assertif a pu être posé jusqu'à present dans le cas des questions rhétoriques, nous le suggérons pour toutes les formes de l'interrogation écrite." (Grésillon & Lebrave, 1984:126).

Assim, uma análise de estruturas interrogativas em contextos escritos aponta, para além das já conhecidas funções diretiva, expressiva, e assertiva (esta última no caso das perguntas retóricas), para uma função estruturadora da informação, que frequentemente opera ao nível tópico-comentário e apresenta um potencial assertivo.

5. Anotação de pares PR

O esquema de anotação do CRPC-DB é baseado na anotação do PDTB, isto é, segue uma abordagem de base lexicalista baseada na presença de conetores que asseguram relações de coesão e coerência no texto. Quando não ocorre um conector explícito, é proposto um conector que expressa a relação de coerência existente entre as duas orações ou frases adjacentes. A anotação contempla o conjunto de relações apresentado na secção 2 sobre o PDTB: relação explícita; relação implícita; relação com Lexicalização Alternativa (AltLex), por exemplo "a razão era que", "um exemplo é"; relação entre entidades referidas em frases adjacentes (EntRel); e ausência de relação (NoRel). Para cada relação do tipo explícito, implícito ou AltLex, o anotador deve indicar o sentido da relação, usando para tal a hierarquia do PDTB 3.0, organizada em 3 níveis (ver anexo). O PDTB prevê a possibilidade de anotar a atribuição de cada relação (atribuição de um locutor identificado por uma sequência textual, como por exemplo, "disse o presidente"), mas o CRPC-DB não inclui por enquanto essa informação.



As propriedades dos pares de PR que identificámos em textos escritos levantam algumas considerações do ponto de vista da sua anotação. A função de muitos dos pares de PR com um locutor diz respeito à estrutura da informação e à organização da estrutura tópico-comentário. As diferentes teorias e abordagens às relações retóricas apresentam conjuntos mais ou menos abrangentes de relações. No quadro da teoria RST, um dos 16 tipos principais de relações retóricas diz respeito à estrutura tópico-comentário e inclui as perguntas-respostas. A RST toma a anotação discursiva numa perspetiva lata, em que as relações discursivas envolvem relações semânticas (causa, resultado, condição), avaliativas, relações textuais como a expansão e relações tópico-comentário ou de mudança de tópico. No caso do PDTB o conjunto de relações é mais restrito e não envolve estas duas últimas relações. A mudança de tópico é anotada como ausência de relação discursiva. No entanto, todos os pares de frases adjacentes devem ser anotados, e o PDTB utiliza, no caso dos pares PR que não sejam hypophora, os sentidos disponíveis na hierarquia. Assim, as propriedades relacionadas com a marcação de tópico-comentário nos pares de PR que foram assinaladas na secção 3 não têm um sentido específico neste esquema de anotação: por exemplo, se o par de PR incide sobre o modo de realização do predicado, a relação entre pergunta e resposta é etiquetada como Modo (Expansion:Manner).

Uma vez que seguimos a anotação do PDTB, mantemos uma anotação próxima do original, sem incluir relações de tópico. Uma das razões é manter uma comparabilidade com outros recursos que usaram o esquema do PDTB. Outra é o facto de encarmos as relações ao nível da estrutura da informação como outro nível de anotação, distinto do nível das relações discursivas, da mesma forma que a avaliação e a modalidade constituem, quanto a nós, outros níveis complementares para uma abordagem geral do texto e do discurso. Sendo assim, seguimos a generalidade das opções do PDTB 2.0, acrescentando no entanto uma informação que permite recuperar os casos de pares PR. Antes de avançarmos com essa apresentação convém ainda referir que teria sido possível não anotar os pares de PR como relação discursiva, e optar antes por NoRel, casos que seriam posteriormente tratados no nível da estrutura da informação. No entanto, existe naturalmente uma relação entre a pergunta e a resposta, contrariamente ao que acontece com os contextos de NoRel.

Por outro lado, a opção de anotação do PDTB envolve uma inconsistência de que temos consciência. Com efeito, este esquema requer que os argumentos da relação sejam objetos abstratos, como referido na secção 2. Ora, nos pares de PR existe, como vimos, uma única proposição, sendo que a resposta a uma interrogativa global fornece o valor de verdade da proposição e a resposta a uma interrogativa parcial fornece o valor da variável identificada na pergunta. Não há, por conseguinte, duas proposições mas apenas uma. Apesar dessa inconsistência com os critérios que definem de raiz a noção de argumento, optamos por manter o esquema do PDTB, em prol da comparabilidade com outros corpora anotados.

A nossa proposta tem como objetivo permitir recuperar os casos de pares de PR no futuro de forma rápida. Para tal, importa que não sejam anotados com tipos de relação diferentes, que dificultem qualquer recuperação futura. Assim, a nossa proposta é que os pares de PR sejam anotados de forma idêntica a outras sequências de frases, não tratando de forma específica estes contextos na escolha do tipo de relação. Tal como noutras sequências frásicas, assume-se uma relação de tipo Explícito ou Implícito, consoante ocorra ou não um conector a ligar as duas frases. Por exemplo, a relação nos casos dos exemplos (21) e (22) é de tipo explícito, uma vez que a resposta é iniciada pelo conector *na verdade* e *porque*, respetivamente. No caso de (23) não há nenhum conector a ligar o par de frases e anota-se como relação implícita, sugerindo o conector implícito *na verdade*. No caso de (21) e (23), entende-se que a resposta providencia informação adicional e permite desenvolver um tópico, pelo que se atribui o sentido de Especificação (Expansion:Level-of-detail:arg2-as-detail) (ver Asher et al. (2017)). O sentido de Especificação é inferido através do conector e do contexto linguístico em (21) e apenas através do contexto em (23). O exemplo (22) é anotado com o sentido de Contingency:Cause:Reason. Além disso, todos os pares de PR recebem um segundo sentido, que identifica a sequência como um par de PR produzido em situação de interação (QAP) ou um par de PR em que o locutor coloca a pergunta e responde ele próprio (QAP-Hypophora). Em casos de dúvida quanto a haver um ou dois locutores, a anotação é conservadora



e opta-se pela etiqueta QAP. Quando não existe uma relação entre a pergunta e o segmento seguinte (isto é, quando o segmento seguinte não se refere diretamente à pergunta), é anotado como NoRel.

- (21) *Cinema soviético, o que é ? Na verdade, deste lugar - que é o lugar específico do espectador português dentro do mercado cinematográfico português - é difícil, senão impossível, responder* (dn81323.txt) [Explicit, Expansion:Level-of-detail:arg2-as-detail; QAP:Hypophora]
- (22) *A que propósito poderá vir tudo isto , se se trata de escrever sobre uma noite de música? Porque, simplesmente, agora que parecem estar resolvidos definitivamente os problemas das ideologias (pelo menos em Portugal), quando o candidato de direita nos quer fazer querer que não é dessas bandas, e que só por acaso topográfico lá passou (neste jogo de acasos que é a política), e que o candidato de esquerda é tido igualmente por direita, cantar debaixo daquela «soleníssima» cúpula que os burgueses são como os porcos e quanto mais velhos mais estúpidos são cheira ainda a sacrilégio!* (de86114.txt) [Explicit, Contingency:Cause:Reason; QAP:Hypophora]
- (23) *Subsídios?* (implicit = na verdade) **“Para o ano esperamos tê-los”**, diz-nos Leandro Vale. (dn8114.txt) [Implicit, Expansion:Level-of-detail:arg2-as-detail; QAP]

A anotação é idêntica à do PDTB 2.0, exceto no facto de aplicar um segundo sentido que identifica os pares de PR; difere do PDTB 3.0 porque os casos de PR com um único locutor (hypophora) não são tratados no CRPC-DB como um tipo de relação diferente, mas sim como um sentido específico, e porque se aplica o sentido QAP a todos os pares PR. Difere, ainda, do TED-MDB porque nesse corpus os casos de hypophora são tratados com o sentido primeiro de hypophora, e a anotação incide sobre a interrogação ao selecionar um elemento lexical interrogativo como conector (auxiliar invertido ou palavra interrogativa). Aproxima-se do corpus STAC por anotar todos os pares PR com o sentido QAP, embora este seja sempre um segundo sentido, para além do primeiro sentido identificado. Assim, mantém-se uma anotação consistente com o PDTB 2.0, no que diz respeito ao tipo de relação e ao primeiro sentido anotado. As perguntas retóricas não têm tipicamente uma resposta, mas podem ter relação com o segmento seguinte e são anotadas de acordo com as normas gerais do PDTB. Estes são casos não recuperáveis com a etiqueta QAP. No entanto, se necessário é possível extrair todos os casos de NoRel em que o primeiro argumento termina com ponto de interrogação.

Fundamental para a discussão dos pares de PR na escrita é o conceito de Atribuição, usado no PDTB. A riqueza de contextos enunciativos diversos e de diferentes funções desempenhadas ao nível textual pelos pares de PR constitui uma área natural para explorar o conceito de Atribuição e a sua futura aplicação no âmbito do CRPC-DB.

6. Conclusão

Os textos do CRPC-DB, um banco do discurso para o português, são uma fonte de reflexão sobre a natureza dos pares de PR em textos escritos. Os pares de PR mostram que estas estruturas envolvem na escrita outras funções, além das funções diretiva (pergunta como pedido de informação) ou expressiva (interrogativas de surpresa). As questões de enunciação são fundamentais para esta análise, pois têm efeito direto nas funções desempenhadas pelos pares PR. A presença de um único locutor nos pares PR invoca um locutor virtual, criando o que pode ser designado como “dupla locução”. Nos contextos que analisámos, a pergunta caracteriza-se por estabelecer uma rutura com o contexto anterior, introduzindo um tópico que é de seguida desenvolvido na resposta. Esta raramente constitui, nestes casos, uma resposta a um pedido de informação e levanta-se a questão de estabelecer o que constitui uma resposta válida. Consideramos que uma não-resposta não deixa de constituir



uma “reação” à pergunta. A rutura que a pergunta estabelece com o contexto anterior e que a resposta desenvolve corresponde assim a uma estruturação do tópico-comentário. A função dos pares PR ao nível da estrutura da informação aparece como fundamental para entender a sua utilização na escrita. A natureza assertiva e modalizada destas perguntas mostra que estamos perante um *continuum* entre as interrogativas ditas “puras”, como pedido de informação de um locutor ao seu interlocutor, os pares de PR com um único locutor, com forte potencial assertivo e modalização, e as perguntas retóricas, que constituem verdadeiras asserções.

Diferentes abordagens foram usadas na anotação discursivas de perguntas, muitas vezes diretamente relacionadas com a modalidade e género dos textos, mas também com os objetivos e conceitos operativos dos quadros teóricos subjacentes à anotação. Enquanto algumas teorias discursivas integram relações ao nível da estrutura tópico-comentário nas categorias do sistema de anotação, outras limitam as relações retóricas a um conjunto mais restrito. Naturalmente, a fronteira entre relações retóricas que tratam da progressão textual, como a Expansão, e relações de tópico-comentário pode ser difícil de delimitar. Seguimos, na nossa proposta de anotação, a versão do PDTB 2.0, embora adicionando etiquetas que identificam PR interacionais e PR de tipo hypophora (no sentido do TED-MDB).

Um desenvolvimento natural deste trabalho será a utilização dos resultados para a discussão da categoria “Atribuição”, que foi proposta no PDTB para associar os argumentos ao seu enunciador. Como o CRPC-DB ainda não contempla a anotação da Atribuição, os casos de PR serão certamente um ponto de partida rico e produtivo para o estudo deste conceito.

Referências:

- Asher, N., Muller, P., Bras, M., Ho-Dac, L. M., Benamara, F., Afantenos, S., & Vieu, L. (2017). ANNODIS and Related Projects: Case Studies on the Annotation of Discourse Structure. In *Handbook of Linguistic Annotation* (pp. 1241–1264). Springer Netherlands. https://doi.org/10.1007/978-94-024-0881-2_47
- Barbosa, P., Santos, P., & Veloso, R. (2020). Tipos de frase e força ilocutória. In *Gramática do Português* (pp. 2517–2586). Fundação Calouste Gulbenkian.
- Borillo, A. (1981). Quelques aspects de la question rhétorique en français. *DRLAV. Documentation et Recherche En Linguistique Allemande Vincennes*, 25(1), 1–33. <https://doi.org/10.3406/drlav.1981.969>
- Carlson, L., & Marcu, D. (2001). Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 2, 1–87. <ftp://128.9.176.20/isi-pubs/tr-545.pdf>
- Culioli, A. (1990). *Pour une linguistique de l'énonciation, Opérations et représentations Tome 1*. Ophrys.
- Ducrot, O. (1972). *Dire et ne pas dire. Principes de sémantique linguistique*. Hermann.
- Généreux, M., Hendrickx, I., & Mendes, A. (2012). Introducing the Reference Corpus of Contemporary Portuguese On-Line. In N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (Eds.), *LREC'2012 -- Eighth International Conference on Language Resources and Evaluation* (pp. 2237–2244). European Language Resources Association (ELRA).
- Grésillon, A., & Lebrave, J.-L. (1984). Qui interroge qui et pourquoi? In *La langue au ras du texte*. Presses Universitaires de Lille.
- Kerbrat-Orecchioni, C. (1999). *L'énonciation*. Armand Colin.
- Lanham, R. (1991). *A Handlist of Rhetorical Terms*. University of California Press.
- Lopes, A. C. M. (2018). *Pragmática: Uma introdução*. Imprensa da Universidade de Coimbra.
- Mann, W. C., & Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3), 243–281.
- Mayoral, J. A. (1994). *Figuras Retóricas*. Editorial Sintesis.
- Prasad, R. ;, Webber, B., & Joshi, A. (2014). Reflections on the Penn Discourse Treebank, Comparable Corpora, and Complementary Annotation. *Dissertation Abstracts International, B: Sciences and*



- Engineering*, 40(4), 921–950. <https://doi.org/10.1162/COLI>
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., & Webber, B. (2008). The Penn Discourse Treebank 2.0. *Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2008*, 2961–2968.
- Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., Joshi, A., Robaldo, L., & Webber, B. (2007). The Penn Discourse Treebank 2.0 Annotation Manual. *IRCS Technical Reports Series*. <https://doi.org/10.1136/bmj.331.7518.689>
- Sharma, H., Dakwale, P., Sharma, D. M., Prasad, R., & Joshi, A. (2013). Assessment of different workflow strategies for annotating discourse relations: A case study with HDRB. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7816 LNCS(PART 1), 523–532. https://doi.org/10.1007/978-3-642-37247-6_42
- Webber, B., Prasad, R., Lee, A., & Joshi, A. (2018). *The Penn Discourse Treebank 3.0 Annotation Manual* (Issue December).
- Zeyrek, D., Mendes, A., Grishina, Y., Kurfalı, M., Gibbon, S., & Ogrodniczuk, M. (2019). TED Multilingual Discourse Bank (TED-MDB): a. *Language Resources and Evaluation*, 587–613. <https://doi.org/10.1007/s10579-019-09445-9>
- Zeyrek, D., Mendes, A., & Kurfalı, M. (2018). Multilingual Extension of PDTB-Style Annotation: The Case of TED Multilingual Discourse Bank. *Proceedings of LREC 2018*. <http://textlink.ii.metu.edu.tr>
- Zeyrek, D., & Webber, B. (2008). A Discourse Resource for Turkish : Annotating Discourse Connectives in the METU Corpus. *Middle East*, 65–72.

Anexos

Anexo 1: Lista das relações retóricas do PDTB 3.0

(Webber et al., 2018)

Temporal	Synchronous	
	Asynchronous	Precedence
		Succession
Contingency	Cause	Reason
		Result
	Cause+Belief	Reason
		Result
	Cause+SpeechAct	Reason
		Result
	Purpose	Arg1 as goal
		Arg2 as goal
	Condition	Arg1 as cond
		Arg2 as cond
	Condition+SpeechAct	
	Negative Condition	Arg1 as negcond



		Arg2 as negcond
	Negative Condition+SpeechAct	
Comparison	Contrast	
	Smilarity	
	Concession	Arg1 as denier
		Arg2 as denier
	Concession+Speech act	Arg2 as denier+speech act
Expansion	Conjunction	
	Disjunction	
	Specification	Arg2 as detail
		Arg1 as detail
	Equivalence	
	Instantiation	
	Exception	Arg1 as exception
		Arg2 as exception
	Substitution	Arg1 as subst
		Arg2 as subst
	Manner	Arg1 as manner
		Arg2 as manner

Anexo 2: Lista das relações retóricas usadas na anotação do Penn Treebank segundo a RST
(Carlson & Marcu, 2001:32-33)

São 78 relações divididas em 16 classes que partilham algum tipo de significado retórico:

- Attribution: attribution, attribution-negative
- Background: background, circumstance
- Cause: cause, result, consequence
- Comparison: comparison, preference, analogy, proportion
- Condition: condition, hypothetical, contingency, otherwise
- Contrast: contrast, concession, antithesis
- Elaboration: elaboration-additional, elaboration-general-specific, elaboration-part-whole, elaboration-process-step, elaboration-object-attribute, elaboration-set-member, example, definition
- Enablement: purpose, enablement
- Evaluation: evaluation, interpretation, conclusion, comment
- Explanation: evidence, explanation-argumentative, reason
- Joint: list, disjunction
- Manner-Means: manner, means
- Topic-Comment: problem-solution, question-answer, statement-response, topic-comment, comment-topic, rhetorical-question
- Summary: summary, restatement
- Temporal: temporal-before, temporal-after, temporal-same-time, sequence, inverted-sequence
- Topic Change: topic-shift, topic-drift

