

Paráfrase de advérbios terminados em *-mente* em Português

Paraphrasing Portuguese Adverbs ending in *-mente*

Jorge Baptista

U.Algarve-FCHS, Campus de Gambelas, P-8005-139 Faro, Portugal

INESC-ID Lisboa, L2F-Spoken Language Lab, R. Alves Redol 9, 1000-029 Lisboa, Portugal

jbaptis@ualg.pt

Resumo

Neste artigo, partimos da análise léxico-sintático-semântica das propriedades que foram usadas para classificar advérbios terminados em *-mente* mais frequentes em português e exploramos a geração de diferentes padrões de paráfrase, tanto estruturas regulares ou muito gerais, tais como os advérbios de modo e de ponto de vista, bem como outros, menos produtivos (e às vezes idiomáticas). O objetivo é fornecer um abrangente conjunto de estratégias de paráfrase, que podem ser usadas em várias aplicações de processamento de linguagem natural, como a simplificação de texto ou até mesmo tradução automática.

Palavras chave

advérbios, terminado em *-mente*, paráfrase, desambiguação de sentido, léxico-gramática, português

Abstract

In this paper, we depart from the lexical-syntactic-semantic properties that were used to classify the most frequent adverbs in Portuguese ending in *-mente* ‘-ly’, and explore the generation of different paraphrasing patterns, both regular or very general structures, such as those for manner and view point adverbs, as well as other, less productive (and sometimes idiomatic) structures. The goal is to provide a comprehensive set of paraphrasing strategies, which can be used in several natural language applications, like text simplification or even machine translation.

Keywords

adverbs, ending in *-mente*, paraphrase, word sense disambiguation, lexicon-grammar, Portuguese

1 Introdução

As construções adverbiais são uma parte importante do conteúdo de qualquer texto e mostram uma sintaxe complexa, que pode ser vista como um desafio para muitas aplicações de Processamento de Linguagem Natural (PLN). As suas

propriedades formais (sintáticas) incluem: (i) o escopo dos advérbios (um único constituinte ou uma frase inteira); (ii) posição (básica), (iii) quantificação, (iv) a parte do discurso ou categoria morfossintática (PoS) que modificam; (v) o tipo de paráfrase(s) que podem permitir. É este último tipo de propriedade que será o foco deste artigo.

Além disso, adotamos o conceito de *paráfrase* na perspetiva de Harris (1991) e do Léxico-Gramática (Gross, 1975, 1996b), isto é, sempre que há uma relação de equivalência transformacional entre frases, o que requer que o mesmo material lexical plenamente significativo (mesmo que de uma forma diferente) esteja envolvido, excluindo-se, portanto, situações de mera sinonímia. A paráfrase é, portanto, uma ferramenta teoricamente motivada para a descrição linguística, embora, até onde sabemos, as questões decorrentes da exploração sistemática de mecanismos parafrásticos envolvidos nas construções adverbiais portuguesas em textos reais (concatenados) não tenham sido descritas anteriormente.

Neste artigo exploratório, partimos das propriedades léxico-sintático-semânticas (Molinier & Levrier, 2000) que foram usadas para classificar os advérbios mais frequentes terminados em *-mente* em português (Fernandes, 2011). A nossa hipótese de partida é que a classificação léxico-sintática é a chave para produzir paráfrases adequadas para esses advérbios e testamos a geração de diferentes padrões de paráfrase em exemplos reais selecionados aleatoriamente a partir de um corpus.

Por outro lado, pretendemos obter um conjunto abrangente de estratégias de paráfrase, juntamente com algumas restrições à sua aplicação, que podem ser usadas como diretrizes para a descrição sistemática de construções adverbiais, e podem ser usadas em várias aplicações de PLN, como a simplificação de texto ou até mesmo tradução automática.



A tarefa de desambiguação de sentido de palavra (ing. *word-sense disambiguation*, WSD) é considerada como um passo anterior para a tarefa de parafrasear estas expressões. Uma abordagem anterior baseada de aprendizagem de máquina para WSD dos advérbios derivados mais frequentes, terminados em *-mente*, do Português do Brasil (Fernandes, 2011) relatou uma precisão geral de 81%. Portanto, somente (ou principalmente) as paráfrases de advérbios não ambíguos (monossémicos) serão consideradas aqui.

Este artigo está organizado do seguinte modo: Começamos por uma sucinta revisão da literatura (secção 2) e apresentamos os métodos aqui utilizados (secção 3), para, logo de seguida, enumerar as diferentes paráfrases consideradas neste estudo (secção 4). Apresentamos, então, os resultados obtidos (secção 5), que comentamos em pormenor. O artigo termina (secção 6) com uma síntese das principais conclusões e apontando perspetivas de trabalho futuro.

2 Revisão da literatura

Embora muitos autores tenham produzido descrições perspicazes sobre a sintaxe e a semântica dos advérbios (Costa, 2008; Ernst, 2002; Real Academia Española, 2010; Kovacci, 2000), esses estudos consistem principalmente em observações esparsas e em alguns esclarecimentos quanto às suas propriedades semânticas e (mais raramente) sintáticas (ou seja, formais). Além disso, as taxonomias e esquemas de classificação, quando produzidos, frequentemente mostram critérios que se sobrepõem. É, portanto, seguro dizer que, até onde sabemos, nenhuma descrição sistemática e abrangente desta categoria morfosintática foi produzida para qualquer língua natural, exceto talvez para advérbios compostos franceses (Gross, 1996a) e particularmente para advérbios derivados que terminam em *-ment* (Molinier & Levrier, 2000).

Neste artigo, adotamos a abordagem do Léxico-Gramática (Gross, 1996b) para a descrição da língua e a classificação de advérbios originalmente proposta por (Gross, 1996a), e posteriormente adaptada à descrição dos advérbios portugueses, nomeadamente, aos advérbios que ocorrem com maior frequência e terminam em *-mente* do português brasileiro (Fernandes, 2011); e aos advérbios compostos do português europeu (Palma, 2009). Neste quadro teórico, as construções adverbiais são organizadas em 9 classes: 3 tipos principais de advérbios modificadores de frases (classes Px) e 6 tipos principais de advérbios modificadores internos de pro-

posição (classes Mx). Por falta de espaço, uma descrição detalhada da sintaxe e semântica dessas construções não pode ser apresentada aqui, pelo que remetemos o leitor para os trabalhos acima referidos.

3 Métodos

A partir do corpus CETEMPúblico (Santos & Rocha, 2001), extraímos primeiro todos os lemas de todas as palavras analisadas como advérbios e terminados em *-mente* (4.384). Destes, um número considerável é constituído claramente por formas com erros ortográficos ou de digitação (e.g. *abolutamente* e *aboslutamente*, por *absolutamente*), incluindo nomes terminados em *-mento* com um erro na última vogal (e.g. *adiamente*, por *adiamento*, e *afastamente*, por *afastamento*), erros de acentuação (e.g. *simultâneamente*, por *simultaneamente*), etc.

Praticamente todas as formas válidas dos advérbios encontrados no corpus já estão presentes no léxico do sistema STRING (Mamede et al., 2012)¹. Neste momento, este léxico contém mais de 6.800 advérbios terminados em *-mente*. Note-se que as variantes ortográficas devidas à presença de consoantes surdas etimológicas (e.g. *actualmente*) são consideradas como formas corretas e associadas ao respetivo lema (v.g. *atualmente*), o qual é estabelecido segundo o Acordo Ortográfico e para a variante europeia.

Naturalmente, nem todos estes advérbios dispõem ainda de uma descrição sintática no léxico do sistema. Das 974 construções adverbiais já descritas no léxico apenas 28 (> 3%) não ocorreram no corpus.

Usando a informação codificada nesse léxico, observamos que 110 (11,3%) dos advérbios encontrados no corpus são ambíguos, com 2 ou mais sentidos (e distinto comportamento sintático-semântico correspondente). Por exemplo, *pontualmente* ‘ocasionalmente/atempadamente’ pode funcionar: (i) como um modificador interno de frase, relacionado com tempo/aspecto (frequência) (classe MT), como em *Pontualmente, o Pedro faz isso* (com aproximadamente o mesmo significado que *ocasionalmente*); ou (ii) como um advérbio modificador interno à frase, orientado ao sujeito (classe MS), como em *O Pedro chegou pontualmente à reunião* = *O Pedro foi muito pontual a chegar à reunião*.

Em seguida, para cada classe sintático-semântica, foram descritas as principais estratégias de paráfrase, utilizando transdutores de

¹<https://string.12f.inesc-id.pt> [31/12/2018].

estados finitos com a ferramenta de processamento de corpus Unitex 3.1 (Paumier, 2016). A descrição linguística baseou-se nas concordâncias dos advérbios de cada classe que foram obtidas a partir do corpus, excluindo primeiro aqueles que seriam utilizados na avaliação.

Um conjunto de transdutores de estados finito foi construído em duas etapas, usando as informações mencionadas na Secção 4 e codificadas no léxico para substituir os advérbios pelas paráfrases apropriadas. Estas informações correspondem às propriedades léxico-sintáticas das construções adverbiais e são representadas num formato tabular.

O grafo da Figura 1 ilustra a primeira etapa do processo e constitui um grafo de referência (aqui ligeiramente simplificado) em que essas propriedades são representadas por variáveis (@X), em que X representa o número da coluna correspondente dessa tabela. A variável @B representa o advérbio-alvo; @D e @E as variantes *PrepC* (Secção §4.1); @F a forma adjetival equivalente; as variáveis @H a @H são propriedades binárias, que só dão origem às transduções indicadas a seguir se tiverem o valor '+'. Desse modo, apenas são produzidos nos transdutores finais os caminhos relevantes para cada construção adverbial.

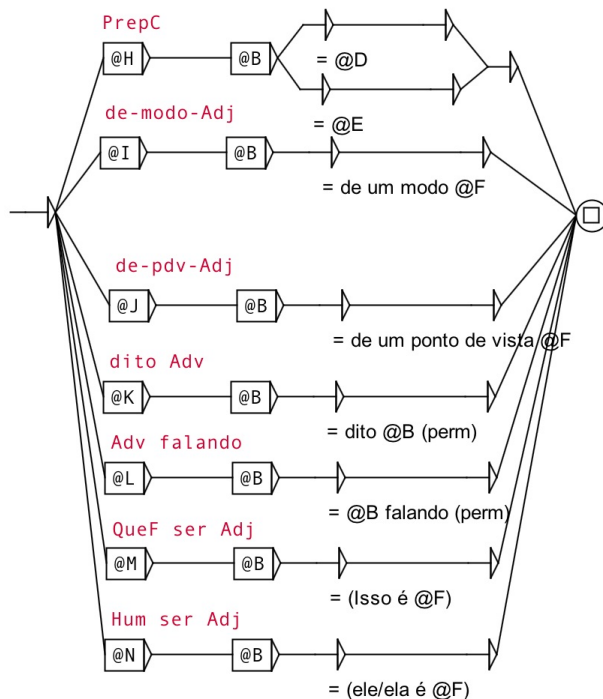


Figura 1: Grafo de referência

Os transdutores finais (aqui apresentados separadamente para cada propriedade, para uma maior clareza) são ilustrados pela Figura 2.

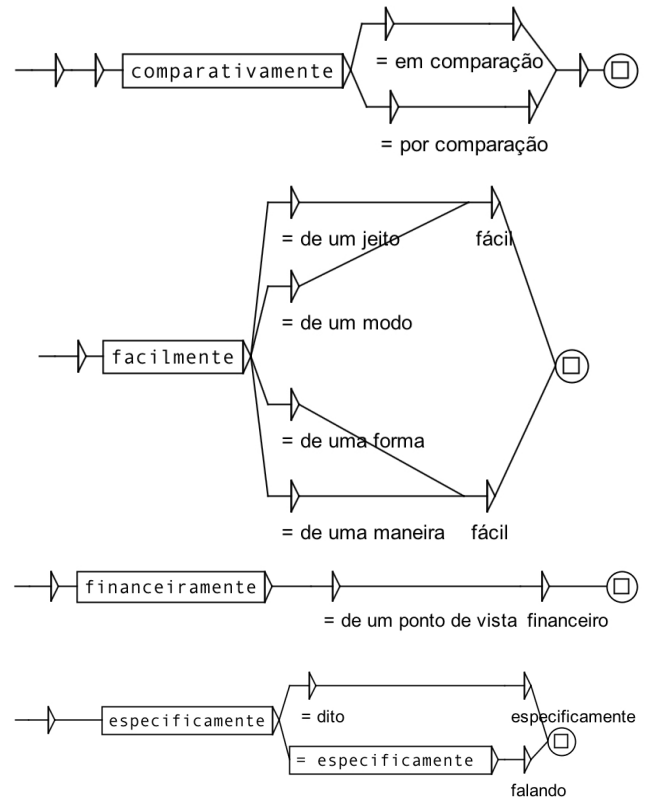


Figura 2: Exemplos de transdutores.

Estes transdutores limitam-se a inserir no texto, imediatamente a seguir ao advérbio, a paráfrase adequada.

Para a avaliação, selecionaram-se os 2 advérbios mais frequentes de cada classe sintático-semântica. No caso de algum desses advérbios ser ambíguo, isto é, apresentar mais do que uma construção sintático-semântica, descartámo-lo e escolhemos o advérbio seguinte, por ordem de frequência, dentro da mesma classe².

De seguida, foram aleatoriamente extraídas do corpus 10 concordâncias para cada um dos advérbios selecionados.

Finalmente, aplicaram-se os transdutores previamente construídos para obter as correspondentes paráfrases. A *qualidade* das frases parafreadas, isto é, sua *equivalência semântica* à frase original extraída do corpus e a sua *aceitabilidade*, foi avaliada independentemente por dois anotadores, ambos falantes nativos de português europeu, e qualquer divergência discutida e resolvida. Como o objetivo do artigo é essencialmente exploratório, nenhuma avaliação do grau de concordância entre anotadores será fornecida aqui.

²Por razões práticas, que apresentaremos mais adiante, nem sempre foi possível fazê-lo de forma sistemática.

4 Paráfrases

Nesta secção, destacamos algumas das estruturas de paráfrase mais comuns encontradas para advérbios terminados em *-mente* em português europeu. Para maior clareza, eles serão organizados por classes, embora algumas destas estratégias possam dizer respeito a várias classes.

Devido à sua sintaxe mais complexa e por não serem propensos a se deixar parafrasear regularmente, os advérbios relacionados com o conceito de *tempo* (classe MT) (Hagège et al., 2009, 2010), os advérbios quantificadores (MQ) e os advérbios de foco (MF) (Baptista & Català, 2011), até porque constituem classes praticamente fechadas, não serão considerados neste artigo. São exemplos destas construções:

Atualmente/ correntemente/ antigamente/ diariamente, o Pedro faz/fazia isso [MT]
O Pedro ficou completamente/ totalmente esgotado [MQ]
O Pedro foi duplamente enganado [MQ]
O Pedro lê essencialmente/ basicamente este tipo de livros [MF]

4.1 PC (advérbio conjuntivo)

Uma das estratégias de paráfrase mais interessantes encontradas para advérbios que terminam em *-mente* é a possibilidade de produzir um grupo preposicional (notado *PrepC*) cuja cabeça é o substantivo morfologicamente associado ao adjetivo base de que o advérbio foi derivado, e.g. *consequentemente* → *em consequência, por consequência*.

Essas paráfrases nominais apresentam frequentemente uma certa fixidez quanto às combinações desse nome com a preposição e o determinante, ou apenas permitem uma variação muito limitada, lexicalmente determinada. Por isso, muitos deles já estão codificados no léxico da cadeia STRING como advérbios compostos (Mamede et al., 2012). Devido às idiosincrasias dessas combinações de palavras, essas paráfrases devem ser diretamente codificadas ou associadas à entrada lexical advérbio terminado em *-mente*.

Os diferentes valores de frequência dessas variantes podem sugerir uma estratégia de paráfrase. Por exemplo, além de *consequentemente* (204 ocorrências), ambas as paráfrases seguintes são encontradas: *em consequência* (767 ocorrências); e *por consequência* (354).

4.2 PS (advérbio disjuntivo de estilo)

Na perspetiva harrissiana (Harris, 1991, p. 91), advérbios disjuntivos de estilo (classe PS) operam como um advérbio de modo que modificam um operador metalinguístico performativo *Eu digo* subjacente a qualquer enunciado efetivamente produzido. Neste sentido, eles poderiam, em princípio, ser parafraseados reconstruindo tal operador (e transformando o enunciado de discurso direto em discurso indireto). Assim, para um enunciado como:

par=ext803257-pol-95b-1: *Especificamente, o principal motivo das consultas é a artrose do joelho, ...*

deve ser possível produzir uma frase como:

Eu digo especificamente que a principal razão para as consultas médicas é a artrite do joelho.

É possível observar variantes destas formas de base, nomeadamente uma oração reduzida de participio (notada *dito Adv*), especialmente quando o advérbio é modificado por *mais*:

... dito mais especificamente, a principal razão para as consultas médicas ...

ou uma oração reduzida gerundiva (notada *Adv falando*), com *falar*:

Especificamente falando, o principal motivo das consultas ...

No entanto, como *especificamente* é ambíguo, sendo classificado tanto como um advérbio de modo (MV) como um advérbio de foco (MF), nos casos em que esta ambiguidade não foi adequadamente resolvida, são esperadas paráfrases incorretas.

4.3 PA (advérbios de atitude disjuntiva)

Esta classe compreende várias subclasses. Começamos pelos advérbios modificadores de frase avaliativos (*PA:eval*):

par=ext4944-eco-97b-2: *Surpreendentemente, o juiz nem leu mais nada ...*

Nestas construções, o advérbio poderia ser considerado uma paráfrase de um verbo opinativo como *Eu acho/penso* tendo como argumento uma frase adjetival:

Eu acho (que é) surpreendente que o juiz nem tenha lido mais nada.

Consideramos ainda os advérbios modais (*PA:modal*), que atribuem uma modalidade específica à frase que modificam:

par=ext436372-des-96a-2: *Provavelmente, [ele] é mais inteligente do que Jean-Jacques.*

Independentemente da classe, estes tipos de construções adverbiais são muitas vezes equivalentes a uma construção adjetival com uma oração completiva sujeito (notado *QueF ser Adj*):

É provável que ele seja mais inteligente do que Jean-Jacques.

No entanto, esta é uma propriedade que sofre fortes restrições lexicais, exigindo descrição explícita, pois alguns advérbios impedem tal transformação:

Aparentemente, ele é mais inteligente do que Jean-Jacques;
cf. **É aparente que ele seja/é mais inteligente que Jean-Jacques.*

Devido à modalidade particular que estes advérbios introduzem, o tempo-modo da oração subordinada na construção adjetival equivalente tem de ser adequado em conformidade, o que representa um grau suplementar de complexidade na formulação da paráfrase:

*Eu acho (que é) surpreendente que o juiz nem tenha/ tivesse/ *tinha/ *tem lido mais nada.*
*É provável que ele seja/*é mais inteligente do que Jean-Jacques.*

Neste momento, ignoramos este tipo de alterações no tempo-modo das orações completivas das construções adjetivais equivalentes às construções adverbiais.

Além das subclasses mencionadas aqui, a classe PA também inclui duas outras subclasses: advérbios disjuntivos de hábito (*PA:habit*), e. g. *habitualmente*, e advérbios de frase orientados para o sujeito (*PA:subj-oriented*), por exemplo *inteligentemente*, e.g. *Inteligentemente, o Pedro não fez isso*. No entanto, como ainda existem poucos advérbios destas subclasses no léxico da STRING e muitas vezes são eles ambíguos com outros empregos, de outras classes, eles não serão abordados neste artigo.

4.4 MV (advérbios de modo)

Quantitativamente, esta é a mais importante classe léxico-sintática de advérbios. A típica estrutura de paráfrase envolve um grupo preposicional com os nomes-operadores de modo (Gross,

1996a) *modo, maneira, forma e jeito* (este último somente para o português do Brasil).

A aceitabilidade da paráfrase está estreitamente dependente da parte do discurso do elemento predicativo que estes advérbios modificam (adjetivo ou verbo), da sua posição relativa, e (muitas vezes) do grau de fixidez da combinação de palavras (colocação) (Vieira et al., 2012).

Considere-se, por exemplo, o advérbio *abertamente*, que pode combinar-se tanto com verbos (1.401), antes (1.256) ou depois (145) deles; como com adjetivos, mas apenas antes destes (131). Apesar do sentido um tanto idiomático do advérbio (não há relação transformacional sincrónica com o verbo *abrir* nem o adjetivo *aberto*), este pode quase sempre submeter-se à paráfrase característica com o nome-operador de modo (e suas variantes) quando combinado com um verbo:

par=ext559653-eco-91b-2: *Os Verdes, o partido que abertamente [= de um modo mais aberto] criticou a cimeira ...*
par=ext21956-pol-95b-1: *Powell criticou abertamente [= de um modo aberto] Robert McNamara ...*

No entanto, como um modificador à esquerda de um adjetivo, e.g.

par=ext391158-opi-97a-2: *Sou abertamente favorável à autonomia regional ...*,

o advérbio parece não poder ser submetido a esta operação, nem antes nem depois do adjetivo:

**Sou de forma aberta favorável à autonomia regional.*
**Sou favorável de forma aberta à autonomia regional.*

Finalmente, deve notar-se que, independentemente da classe sintático-semântica, alguns advérbios que não pertencem à classe MV mas que terminam em *-mente* ainda retêm a possibilidade de serem parafrazeados pelas construções (derivadas analiticamente?) com nomes-operadores de modo. Por exemplo, o advérbio *paradoxalmente* (classe PC):

par=ext1450788-nd -91a-2: *Por esta razão e de modo paradoxal [= paradoxalmente] ... , nos anos 70, a população estrangeira aumentou em vez de diminuir ...*

Este não é frequentemente o caso de advérbios que não são da classe MV, como, por exemplo, o advérbio ambíguo *consequentemente*, que pode ser tanto um advérbio conjuntivo (PC) quanto um advérbio de modo (MV), e cujas construções

se podem distinguir consoante a respetiva posição na frase. No início de frase e destacado por vírgulas:

Consequentemente [→ *De uma maneira consequente*], *o Pedro fez isso*,

apenas encontramos a construção PC, já que a paráfrase com nome-operador só seria interpretada como uma forma de advérbio MV; no final da frase, sem ser separado por vírgulas dos restantes elementos da frase, apenas a análise como MV, ainda que rara, é natural:

O Pedro fez isso consequentemente [= *de uma maneira consequente*].

4.5 MS (advérbios de modo orientados para o sujeito)

Como parte da sua definição de duplo escopo, os advérbios de maneira orientados para o sujeito (MS), como *discretamente* (nos exemplos, nomes próprios foram abreviados):

par=ext203154-pol-92a-1: *O dirigente socialista MS vai apoiar discretamente uma candidatura de NM à liderança*,

além de permitirem a paráfrase com o nome-operador de modo (e suas variantes):

O dirigente socialista MS vai apoiar de modo discreto a candidatura de NM à liderança,

também permitem uma paráfrase com uma construção adjetival que capta a relação entre o advérbio e o sujeito:

O líder socialista MS será discreto ao apoiar a candidatura do NM à liderança.

4.6 MP (advérbios no ponto de vista)

No início de uma frase, os advérbios de ponto de vista (MP) podem ser parafraseados geralmente por um grupo preposicional com o nome-operador composto *ponto de vista* (notado *pdv*):

par=ext1326334-eco-92a-1: *Financeiramente* [= *de_o/um ponto de vista financeiro*], *o mercado de ações foi afetado*, ...

O enquadramento do conteúdo da oração principal pelo advérbio MP apela para o ponto de vista do locutor, o que pode parcialmente explicar a paráfrase com a oração reduzida gerundiva com *falar*:

Financeiramente falando, *o mercado de ações foi afetado*, ...

4.7 Variação posicional

Concluimos esta secção com uma observação sobre variação posicional, outra importante propriedade sintática para caracterizar (e distinguir) construções adverbiais.

Para os advérbios conjuntivos PC como *consequentemente*, verificamos que este se encontra separado por vírgula(s), no começo da frase, em 204 instâncias; 1.199 ocorrências no meio da frase; e, embora teoricamente possível, não se encontraram instâncias deste advérbio em final de frase. O advérbio homógrafo, da classe MV, é bastante raro (11 instâncias) e aparece apenas próximo de (após) um conjunto limitado de verbos, e.g. *agir*.

A variação posicional, embora não envolva exatamente a reformulação do advérbio noutra expressão equivalente, pode ser considerada um tipo especial de paráfrase, pois implica alterações (orto)gráficas no texto (ou seja, alterações de letras de maiúsculas para minúsculas e uso de pontuação). Dadas as dificuldades que essa variação levanta à geração de paráfrases, ela não foi considerada neste trabalho.

5 Resultados

As concordâncias dos 14 advérbios mais frequentes terminados em *-mente* (a maioria não ambíguos), dois de cada uma das classes selecionadas, foram recolhidas do corpus e 10 instâncias de cada advérbio foram selecionadas aleatoriamente para a avaliação.

A Tabela 1 mostra a informação codificada para cada advérbio no conjunto de teste e os resultados (precisão) obtidos pela avaliação tanto da identidade do significado quanto da aceitabilidade das paráfrases produzidas desta maneira. As lacunas lexicais estão marcadas com ‘-’ e as propriedades não relevantes estão marcadas com ‘x’.

Os valores de precisão indicados com ‘*’ representam frases analisadas para as quais não se esperava à partida a possibilidade de estabelecer a paráfrase assinalada. A precisão total por tipo de paráfrase é assim indicada duas vezes: primeiro, considerando apenas os casos julgados relevantes para a classe sintático-semântica considerada (Total 1); e, depois (Total 2), todos os casos analisados, incluindo, portanto, os assinalados com ‘*’.

Globalmente, considerando apenas os casos relevantes (Total 1), obteve-se uma precisão de 0,79; quando se considera todos os casos analisados, esse valor desce para 0,77. Para maior clareza, os resultados serão comentados por tipo

de paráfrase e os exemplos são fornecidos da seguinte forma: primeiro o texto original, depois a paráfrase produzida automaticamente e, finalmente, a forma correta/desejada (sinalizada por ‘→’).

5.1 *PrepC*

O estabelecimento de uma equivalência entre o advérbio e um grupo preposicional encontrou algumas dificuldades. Quando o advérbio é também modificado por um advérbio comparativo (v.g. *mais*, *tão*), o qual precede o advérbio, a ordem inicial das palavras é alterada ou a escolha do elemento comparativo deve sofrer mudanças, e.g.:

adaptar-se mais facilmente
 **adaptar-se mais com facilidade*
 → *adaptar-se com mais facilidade*;

não se entende tão facilmente
 **não se entende tão com facilidade*
 → *não se entende com tanta facilidade*.

Em construções passivas, a paráfrase pelo grupo preposicional é mais natural após o particípio passado, embora o advérbio também possa ocorrer antes dele:

A embaixada da Rússia foi também violentamente atacada pelos manifestantes
 (cp. *A embaixada da Rússia foi atacada violentamente pelos manifestantes*)
 **A embaixada da Rússia foi também com violência atacada pelos manifestantes*
 → *A embaixada da Rússia foi também atacada com violência pelos manifestantes*.

Certas restrições às combinações de palavras parecem estar relacionados com o estatuto de colocação dessas combinatórias lexicais (Vieira et al., 2012):

Cerqueira estava judicialmente impedido de exercer
 **Cerqueira estava na/pela justiça impedido de exercer*
 → *Cerqueira estava impedido *na/?pela justiça de exercer*.

5.2 *de modo Adj*

Como esperado, a maioria dos exemplos de advérbios de modo (classe MV), e menos os advérbios orientados para o sujeito (classe MS), permitem a paráfrase com o nome-operador de modo (*forma*, *maneira* e *modo*; *jeito*, este último

apenas em BP), e os erros são parcialmente devidos à já mencionada combinação com comparativos e a construção passiva. Devido à sua natural ambiguidade com os advérbios de modo (MV), os advérbios disjuntivos de estilo (classe PS) também permitem essa paráfrase. Ainda assim, como um PS, o advérbio *literalmente* parece ser mais usado para expressar uma modalidade (*realis*, ou modalidade real) do que propriamente o modo e a aceitabilidade das frases é geralmente duvidosa ou apenas mantém a interpretação da construção MV. Num uso claro, este valor de modalidade torna a frase inaceitável de uma forma mais evidente:

[*Eu*] *não sei literalmente [= absolutamente] nada*
 *[*Eu*] *não sei de um modo literal nada*.

O resultado obtido com *especificamente* deveu-se ao seu uso predominante como advérbio de foco (MF). Surpreendentemente, o advérbio *curiosamente* (PA:eval) permite esta paráfrase. Como esperado, todos os advérbios de ponto de vista (MP) permitem esta paráfrase com o nome-operador ponto de vista (notado *pdv*).

As próximas propriedades são específicas de advérbios modificadores de frase.

5.3 *Dito Adv-mente e Adv-mente falando*

Embora apenas advérbios PS devessem aceitar estas paráfrases, nenhum dos 2 advérbios selecionados o faz, pelas razões já explicadas acima. Vale ressaltar que *comparativamente* (PC) permite a equivalência com a estrutura gerundiva:

par=ext313886-des-94a-2: *A corrida feminina produziu, comparativamente [falando], melhores resultados [do] que a masculina*.

Surpreendentemente, porém, a maioria dos exemplos de advérbios de ponto de vista (MP) também podem ser parafraseados da mesma maneira.

A avaliação das propriedades seguintes, pela complexidade de que a geração das paráfrases correspondentes se reveste, consistiu basicamente em verificar se seria possível construir uma construção adjetival equivalente com oração completa sujeito (Secção 5.4) ou com sujeito humano (Secção 5.5) e não propriamente uma avaliação da equivalência/aceitabilidade de frases concretas.

Class	Adv-mente	PrepC	Prep-C	de-modo-Adj	de-pdv-Adj	dito Adv	Adv falando	QueF ser Adj	N ^{hum} ser Adj
PC	<i>comparativamente</i>	<i>em/por comparação</i>	0,9	x	x	x	1,0*	x	x
PC	<i>finalmente</i>	<i>por fim</i>	1,0	x	x	x	x	x	x
PS	<i>especificamente</i>	-	x	0,3*	x	0,1	0,2	x	x
PS	<i>literalmente</i>	-	x	0,7*	x	0,0	0,7	x	x
PA:eval	<i>curiosamente</i>	<i>por curiosidade</i>	1,0	1,0*	x	x	x	1,0	x
PA:eval	<i>infelizmente</i>	-	0,9	x	x	x	x	-	x
PA:modal	<i>aparentemente</i>	-	0,9	x	x	x	x	-	x
PA:modal	<i>realmente</i>	<i>na realidade</i>	1,0	x	x	x	x	-	x
MV	<i>facilmente</i>	<i>com facilidade</i>	1,0	0,6	x	x	x	x	x
MV	<i>imediatamente</i>	<i>de/no imediato</i>	0,9	0,9	x	x	x	x	x
MS	<i>cuidadosamente</i>	<i>com cuidado</i>	1,0	0,5	x	x	x	x	1,0
MS	<i>violentamente</i>	<i>com violência</i>	1,0	0,6	x	x	x	x	0,8
MP	<i>financeiramente</i>	-	x	x	1,0	x	0,9	x	x
MP	<i>judicialmente</i>	<i>na justiça</i>	1,0	x	1,0	x	0,8	x	x
Total 1 (casos relevantes)			0,96	0,65	1,00	0,05	0,58	1,00	0,90
Total 2 (todos casos)			0,96	0,60	1,00	0,05	0,66	1,00	0,90

Tabela 1: Propriedades parafrásticas de advérbios selecionados e precisão alcançada.

5.4 (*Eu acho*) *QueF* é *Adj*

Esta propriedade aplica-se especificamente a advérbios avaliativos (subclasse *PA:eval*). É praticamente impossível produzir automaticamente paráfrases, a menos que a frase seja muito simples, ou a oração na qual o advérbio opera tenha sido corretamente extraída. Um exemplo possível é:

par=ext1154964pol-98a-1: [*Ele*] não desmente, curiosamente, a citação ... ,

o que produziria:

→ (*Eu acho que*) é curioso que [*ele*] não desminta a citação.

Observe-se a mudança do modo-tempo do verbo na oração subordinada completiva:

desmente (indicativo-presente)
→ *desminta* (conjuntivo-presente)

e que aqui ignorámos, para efeitos de avaliação.

O advérbio *infelizmente* parece particularmente resiliente a esta transformação, mesmo em frases simples, e.g.:

par=ext1198599-nd-91b-2: *A guerra, infelizmente, criou situações que nos vão dar muito material para escrever*

o que não parece corresponder naturalmente a:

*[*Eu*] *acho que é infeliz que a guerra tenha criado situações que nos vão dar muito material para escrever.*

5.5 *Nhum ser Adj*

Esta propriedade corresponde à elicitación do duplo-escopo dos advérbios de maneira orientados para o sujeito (classe *MS*), portanto, pressupõe não só a correta análise sintática do sujeito do verbo, o que nem sempre é possível de fazer automaticamente, como a correta atribuição de um valor semântico de *humano*, dado o contexto em que esse nome ocorre. Os dois casos assinalados correspondem justamente a construções com sujeito não-humano (v.g. *camioneta* e *excesso de sódio*), em que ambos os advérbios surgem encaixados sob o verbo de uma oração relativa, pelo que seria necessário dispôr de um sistema de resolução de correferência (Marques, 2013).

par=ext1551708-soc-93b-2: ... *despiste de uma camioneta que embateu violentamente numa estação de correios ...*

par=ext1376835-clt-soc-93b-1: ... *não é só o excesso de sódio que ... ataca violentamente cada célula do nosso corpo*

6 Conclusão e trabalho futuro

Em suma e para concluir, parece ser possível produzir as alterações estritamente locais autorizadas pelos modificadores internos às proposições (classes *Mx*) usando as ferramentas aqui usadas, enquanto as propriedades específicas de advérbios modificadores de frases (classes *Px*) são

geralmente mais difíceis de formalizar. A maioria das propriedades parafrásicas é específica da classe ou mesmo lexicalmente dependente, e deve ser dada atenção a certos contextos sintáticos, nomeadamente a presença de outros modificadores adverbiais, a coordenação, integração em construções passivas e, em geral, o fato de que o advérbio está a modificar um verbo ou adjetivo. Está claro, a partir da amostra de advérbios aqui descrita, que a maioria das propriedades é lexicalmente dependente e que a tarefa de paráfrase não pode ser abordada apenas usando a classificação geral das construções adverbiais, aqui resumidamente esboçada.

No futuro, está prevista atingirmos uma maior cobertura do léxico-gramática dos advérbios, com vista ao seu potencial para uso na descrição linguística de parafrases em diferentes aplicações, como simplificação de texto, ferramentas tutoriais de aprendizagem de idiomas e em tradução. Em particular, ainda é necessário elaborar um catálogo de sentidos de advérbios de elevada granularidade, cobrindo uma proporção maior do léxico dos advérbios portugueses terminados em -mente, e associando-lhes as propriedades formais que permitem distinguir diferentes construções de um mesmo vocábulo. Isso só pode ser feito pela paciente descoberta das diferenças entre os usos de palavras, a fim de poder, então, anotar com segurança esses sentidos em textos. Tal permitirá usar técnicas de aprendizagem automática para a desambiguação de sentido destas palavras e, conseqüentemente, a sua utilização em tarefas que envolvam paráfrase.

Agradecimentos

A pesquisa para este trabalho foi parcialmente financiada pelo Governo Português, através da Fundação para a Ciência e a Tecnologia (ref. UID/CEC/50021/2019).

Referências

- Baptista, Jorge & Dolors Català. 2011. Adverbes focalisateurs et analyse syntaxique automatique de groupes nominaux. Em *Passeurs de mots, passeurs d'espoir: lexicologie, terminologie et traduction face au défi de la diversité. Actes des 8èmes Journées scientifiques du LTT-AUF*, 97–110.
- Costa, João. 2008. *O Advérbio em Português Europeu*. Colibri.
- Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge University Press.
- Fernandes, Gaia. 2011. *Automatic Disambiguation of -mente ending Adverbs in Brazilian Portuguese*: Universidade do Algarve e Universitat Autònoma de Barcelona. Tese de Mestrado.
- Gross, Maurice. 1975. *Méthodes en syntaxe*. Hermann.
- Gross, Maurice. 1996a. *Grammaire transformationnelle du français: 3 - syntaxe de l'adverbe*. ASSTRIL.
- Gross, Maurice. 1996b. Lexicon-grammar. Em Keith Brown & J. Miller (eds.), *Concise Encyclopedia of Syntactic Theories*, 244–259. Pergamon.
- Hagège, Caroline, Jorge Baptista & Nuno Mamede. 2009. Portuguese temporal expressions recognition: from TE characterization to an effective TER module implementation. Em *7th Brazilian Symposium in Information and Human Language Technology (STIL)*, 36–43.
- Hagège, Caroline, Jorge Baptista & Nuno Mamede. 2010. Caracterização e processamento de expressões temporais em português. *Linguamática* 2(1). 63–76.
- Harris, Zellig Sabettai. 1991. *A Theory of Language and Information. A Mathematical Approach*. Clarendon Press.
- Kovacci, Ofelia. 2000. El adverbio. Em Ignacion Bosque & Violeta Demonte (eds.), *Gramática Descriptiva de la Lengua Española*, vol. 1, chap. 11, 705–786. Real Academia Española/Espasa.
- Mamede, Nuno, Jorge Baptista, Cláudio Diniz & Vera Cabarrão. 2012. STRING - a hybrid statistical and rule-based natural language processing chain for Portuguese. Em *Computational Processing of the Portuguese Language (PROPOR)*, s/p.
- Marques, João. 2013. *Anaphora resolution*: Instituto Superior Técnico - Universidade de Lisboa. Tese de Mestrado.
- Molinier, Christian & Françoise Levrier. 2000. *Grammaire des adverbes: description des formes en -ment*. Droz.
- Palma, Cristina. 2009. *Estudo contrastivo Português-Espanhol de expressões fixas adverbiais*: Universidade do Algarve. Tese de Mestrado.
- Paumier, Sébastien. 2016. *Unitex 3.1 - user manual*. Université de Paris-Est/Marne-la-Vallée - Institut Gaspard Monge. <http://igm.univ-mlv.fr/~unitex/UnitexManual3.1.pdf>.

- Real Academia Española. 2010. *Nueva gramática de la lengua Española - manual*. Asociación de Academias de la Lengua Española.
- Santos, Diana & Paulo Rocha. 2001. Evaluating CETEMPúblico: A free resource for Portuguese. Em *39th Annual Meeting of the Association for Computational Linguistics*, 442–449.
- Vieira, Lucas Nunes, Cláudio Diniz, Nuno Mamede & Jorge Baptista. 2012. A lexicon of verb and *-mente* adverb collocations in Portuguese: Extraction from corpora and classification. Em *31st International Conference on Lexis and Grammar*, 155–162.