

Clara e sucintamente: um estudo em corpus sobre a coordenação de advérbios em *-mente*

Susana Cavadas Afonso

Projecto VISL*, Universidade do Sul da Dinamarca e Linguateca

1. Introdução

As estruturas de coordenação de advérbios em *-mente*, com a redução do sufixo no primeiro elemento coordenado, coloca questões interessantes sobre o ponto de vista da análise automática do português, especialmente no campo da morfossintaxe.

Tendo em conta que um analisador automático não produz sentidos, duas questões se colocam na análise automática desta estrutura: qual a classe morfológica do primeiro elemento coordenado— adjectivo ou advérbio?— e como identificar uma estrutura de coordenação a nível adverbial quando inserida em contexto.

O objectivo deste estudo é a busca de contextos sintáctico-semânticos que permitam 1) a identificação automática de contextos de coordenação a nível adverbial e 2) uma vez identificada a estrutura de coordenação, propor a classe morfológica advérbio para o primeiro elemento coordenado.

A organização do artigo é a seguinte: em primeiro lugar, será apresentado o problema que uma estrutura de coordenação dos advérbios em *-mente* levanta; em segundo lugar apresentar-se-á os contextos sintáctico-semânticos encontrados que permitem a identificação de uma coordenação a nível adverbial por oposição a contextos considerados ambíguos de forma automática. Segue-se a apresentação dos resultados quantitativos do estudo discussão. Como considerações finais, apontam-se futuros desenvolvimentos deste estudo.

2. O problema

A consideração deste problema surge na discussão de o que avaliar e como no contexto de uma avaliação conjunta de analisadores morfológicos para a língua portuguesa (Santos & Rocha, este volume).

É na análise da categoria gramatical e lematização que o problema da coordenação de advérbios em *-mente* é levantado, relativamente ao primeiro elemento coordenado: é este um adjectivo ou um advérbio, em termos de categoria gramatical, e qual o seu lema? A título de exemplo, considere-se a seguinte frase:

* No endereço <http://www.visl.sdu.dk>

(1) Uma situação *clara e sucintamente* abordada.

A primeira observação é que não é garantida a análise de coordenação de advérbios em *-mente* de *clara e sucintamente*. A leitura de não coordenação a nível adverbial mas a nível de dependentes pós-nominais (*clara e abordada*) é possível, como ilustra a paráfrase a (1):

(2) Uma situação clara e abordada sucintamente.

Supondo que adoptaríamos uma leitura de coordenação de advérbios em *-mente*, alguns argumentos para a análise de advérbio em vez de adjetivo do primeiro elemento coordenado devem ser considerados: 1) é invariável (feminino do singular) ao contrário dos adjetivos que apresentam variabilidade em género e número; 2) não partilha da classe de advérbio; e 3) pode (teoricamente) apresentar a forma não reduzida.

Em português um mesmo lexema pode ter duas análises morfológicas distintas como é o caso do advérbio e do adjetivo (advérbios com forma adjectival), como é o caso de 'alto' (Pedro é alto_[adj]; Pedro fala alto_[adv]), casos também verificados em outras línguas românicas como o espanhol (Bosque, 1989). Pensando no léxico electrónico, estas formas teriam duas entradas: adjetivo e advérbio, estando depois a sua desambiguação morfológica dependente do contexto sintáctico. No entanto, estes casos são limitados pelo que a duplicação das formas não constituiriam uma sobrecarga do léxico.

No entanto, no caso da coordenação de advérbios em *-mente*, o primeiro advérbio coordenado, com o sufixo reduzido, pode ser encarado como um adjetivo. Neste caso, a desambiguação automática das formas não é pacífica, como observado em (1). Teoricamente, a relação entre adjetivos e advérbios que deles derivam é fundamental para a análise da forma do primeiro elemento coordenado¹ (Gross, 1990), porque um adjetivo que não possa derivar em advérbio em *-mente* nunca poderia ocupar a posição de advérbio com o sufixo reduzido numa hipotética coordenação de advérbios.

No contexto do processamento de linguagem natural, um estudo exaustivo das relações entre adjetivos e advérbios que deles derivam permitiria, por um lado, a identificação dos adjetivos que não derivam em advérbios em *-mente*, mas, por outro lado, seria uma tarefa de difícil execução- uma vez que não há corpus nenhum que contenha a totalidade dos adjetivos ou advérbios existentes no português- e quase exclusivamente humana/manual², o que não é desejável.

¹ Numa coordenação de advérbios em *-mente*, podem existir mais do que um advérbio com o sufixo reduzido: *Uma situação clara, sucinta e prontamente abordada*.

² Seria difícil a automatização desta tarefa, uma vez que um computador, até à data, não é capaz de produzir sentidos.

3. A busca em corpus: método

A definição de contextos sintáctico-semânticos teve como base a observação do corpus do CETEMPúblico (Rocha e Santos, 2000), usando o sistema de busca do projecto AC/DC (Santos 2000; Santos e Bick 2000).

Extraiu-se, através de expressões regulares, a coordenação (copulativa, adversativa e disjuntiva) de lexemas com os sufixos no feminino do singular, *-eira*, *-osa*, *-ica*, *-ária*, *-esca*, com advérbios em *-mente*³. Este estudo preliminar foi feito apenas sobre a forma das palavras para não ser influenciado por eventuais erros do analisador sintáctico. As ocorrências que não apresentavam, como primeiro elemento coordenado, uma forma adjectival foram eliminadas⁴.

Não se excluíram, porém, os casos em que esta sequência não constituía uma estrutura de coordenação a nível adverbial, isto é, casos em que o primeiro elemento coordenado é na realidade um adjectivo, semântica, funcional e formalmente, o que se revelou extremamente útil na identificação de contextos que não conduzem a uma leitura de coordenação a nível adverbial.

As ocorrências extraídas foram depois analisadas e categorizadas.

4. Descrição dos contextos sintáctico-semânticos

A identificação de contextos sintáctico-semânticos, que permitem a consideração ou rejeição de uma estrutura de coordenação adverbial, focou-se essencialmente 1) na posição da estrutura de coordenação a nível adverbial na oração; 2) a concordância da forma adjectival (com o sufixo *-mente* reduzido) com um núcleo nominal; 3) no tipo de verbos próximos da hipotética coordenação a nível adverbial e na sua estrutura argumental; 4) possíveis expressões fixas⁵; e 5) restrições semânticas.

4.1 Contextos sintácticos

A tabela seguinte refere-se aos contextos de coordenação de advérbios em *-mente*:

Contextos		Exemplos
Posição da estrutura de coordenação na	Início de oração	• Teórica e constitucionalmente , a soberania democrática do reino reside na Câmara dos Comuns.
	Fim de oração	• ...veremos Picasso trabalhar vigorosa e desmultiplicadamente , passear com a mulher,...

³ Este estudo preliminar foi feito apenas sobre a forma das palavras para não ser influenciado por eventuais erros do analisador sintáctico.

⁴ Por exemplo, ...na prática [n] e visivelmente...

⁵ Sobre expressões fixas em geral, ver Hudson 1998.

frase/oração	Isolamento por sinais de pontuação	▪ ...irá necessariamente culminar no Pinhal Novo, de forma a desfazer (precária e temporariamente) o nó da rede ferroviária em Lisboa.
Expressões fixas		▪ única e exclusivamente ▪ única e simplesmente
Não concordância do primeiro elemento coordenado com um núcleo nominal		▪ pareciam-nos os verdadeiros 'outros' _[masc. pl.] , verdadeira _[fem. sg.] e definitivamente diferentes,
Tipos de verbos e estrutura argumental	Verbos transitivos	▪ ...Zimmerman regressa com uma individual nesta galeria especializada que já anteriormente tinha apoiado técnica e comercialmente o seu trabalho
	Verbos regidos de preposição	▪ ...que passam necessária e resumidamente por : <i>Não se recusar a ser submetido a exames físicos e...</i>
	Verbos intransitivos	▪ O Museu Leal da Câmara – que reabriu Irónica e significativamente no passado dia 5 de Outubro
	Verbos copulativos (não concordância sujeito-complemento)	▪ O jogo _[masc. sg.] foi técnica _[fem. sg.] e emotivamente fraco, em particular na primeira metade,...
	Verbos transobjectivos (não concordância objecto-complemento)	▪ sido sempre bem tratado e que se _[masc. sg.] encontra "física _[fem. sg.] e mentalmente" de perfeita saúde"

Tabela 1: Contextos de coordenação de advérbios em *-mente*

O tipo de verbo e a sua estrutura argumental é um contexto de desambiguação das estruturas tratadas neste estudo, na medida em que equaciona a natureza dos objectos e dos complementos, relacionando-a com a forma do elemento coordenado com o sufixo *-mente* reduzido.

Assim, os verbos transitivos directos seleccionam um objecto directo cuja natureza é prototipicamente nominal (não adjectival). Uma vez identificado o objecto directo (no exemplo acima, 'o seu trabalho') o primeiro elemento coordenado não poderá constituir um potencial objecto. O mesmo ocorre com os verbos que regem preposição. O primeiro elemento da coordenação está disponível apenas para ocupar o lugar de elemento coordenado de advérbios em *-mente* se identificada a preposição regida e o seu complemento (no exemplo acima, 'por' regida por 'passar'). Os verbos intransitivos, por outro lado, não seleccionando nenhum objecto ou complemento, só admitem circunstanciais, nos quais se inclui a estrutura de coordenação de advérbios em *-mente*.

Em relação aos verbos copulativos e transobjectivos, complementos são seleccionados. Uma das características do predicativo do sujeito e do objecto é exibirem traços de concordância de género e número relativamente ao sujeito e ao objecto. Não havendo concordância entre aqueles constituintes, estamos perante uma coordenação de advérbios em *-mente*.

Finalmente, quanto às expressões fixas, é importante referir que, neste estudo, não se considera que *única e simplesmente* e *única e exclusivamente* são expressões

fixas no português europeu, mas tão só na amostra observada (cf. resultados quantitativos em baixo).

Os contextos de não coordenação são os seguintes:

Contexto	Exemplo
Paralelismo de construção: <adv><adj:fem.sg.><conj_c><adv-mente><adj>	<ul style="list-style-type: none"> ▪ Watson foi suficientemente ligeira e controladamente ingénua em toda a sua presença em palco. ▪ A promiscuidade fortuita do elevador é mais democrática e potencialmente mais variada do que a de um avião...
Traço [+concordância] entre o modificador e o núcleo nominal	<ul style="list-style-type: none"> ▪ ... terá visto travões excessivos à ocupação_[fem. sg.] turística_[adj: fem. sg.] ou simplesmente uma redução aos poderes absolutos....
Tipo de verbo e estrutura argumental: verbos copulativos	<ul style="list-style-type: none"> ▪ Banu_[fem. sg.] não é religiosa_[fem. sg.] e obviamente o Refah preocupa-a, ...

Tabela 2: Contexto de não coordenação de advérbios em *-mente*

A sequência mínima: advérbio <adv>, adjectivo no feminino do singular <adj:fem. sg.> (potencialmente o advérbio com o sufixo *-mente* reduzido presente na coordenação de advérbios), conjunção coordenativa <conj_c>, advérbio em *-mente* <adv-mente> e um adjectivo <adj>, é indicadora de que não há coordenação a nível adverbial. Algumas razões podem ser apontadas para este facto. Em primeiro lugar, dificilmente um advérbio em *-mente* poderia ser modificador de um outro advérbio em *-mente* ou de coordenação a nível adverbial. Por exemplo, ‘suficientemente’ como modificador de uma hipotética coordenação de ‘ligeira e controladamente’ não parece ser uma análise aceitável. Por outro lado, advérbios quantificadores (ex: ‘mais’) e/ou de intensidade (ex: ‘tão’) podem à partida modificar uma coordenação de advérbios em *-mente* (‘mais democrática e potencialmente variada’). No entanto, se o adjectivo for modificado por mais do que um advérbio, sendo este também de quantidade ou intensidade (‘potencialmente mais variada’), o primeiro advérbio na sequência não pode ser modificador de toda a potencial estrutura de coordenação⁶.

Quanto à não ocorrência dos contextos que auxiliam na identificação de uma coordenação de advérbios em *-mente*, apenas a concordância (traço +conc) e o tipo de verbos/estrutura argumental relativamente aos verbos copulativos foram observados como sendo contextos de não coordenação de advérbios em *-mente*.

⁶ Veja-se a não aceitabilidade de * (*mais(democrática e potencialmente mais variada)*).

4.2. Ambiguidade e restrições semânticas

Como foi acima referido, a não ocorrência dos contextos identificadores de uma coordenação de advérbios em *-mente* poderia supostamente conduzir a uma leitura de não coordenação. Porém, a concordância em género e número entre constituintes ou entre modificadores e núcleos pode igualmente conduzir a uma situação de ambiguidade, como o seguinte exemplo ilustra:

- (3) ...requerem não só um elevado grau de especialização, mas acima de tudo o recurso a informação_[fem. sg.] **rigorosa**_[fem. sg.] e **factualmente correcta**_[fem. sg.].

Em (3), a ambiguidade resulta do facto de ‘rigorosa’ concordar em género e número com o núcleo nominal ‘informação’. Desta forma, as duas análises de coordenação de advérbios em *-mente* (‘rigorosa e factualmente’) e coordenação dos dependentes nominais (‘rigorosa e (...) correcta’) são legítimas.

A ambiguidade presente em (3) é uma ambiguidade real que apenas pode ser resolvida através de traços prosódicos. Porém, há casos em que a ambiguidade pode ser resolvida através da semântica. As restrições à adverbialização de determinados adjectivos, bem como restrições à coordenação de elementos desiguais hierárquica e semanticamente podem auxiliar na desambiguação. As relações lexicais parece ser igualmente importantes. Veja-se a não aceitabilidade dos seguintes exemplos, considerando uma análise de não coordenação:

- (4) * ...paixão por esta namorada **virtual física** e mentalmente dotada.
 (5) * A diáspora pode ser **financeira** e politicamente útil à pátria...

Os adjectivos ‘virtual’ e ‘física’ parecem não poder ser combinados, bem como o conceito de ‘diáspora’ combinado com o adjectivo ‘financeira’.

5. Dados quantitativos: apresentação dos resultados e discussão

No total, extraíram-se 2830 ocorrências, das quais, 2384 foram inspeccionadas e categorizadas⁷.

A tabela 4 representa a distribuição dos contextos extraídos pelas categorias formadas⁸:

⁷ As restantes foram eliminadas por falta de contexto linguístico para a identificação de contextos, extracção cujo primeiro elemento coordenado não tinha forma adjectival, má formação frásica, ou não categorizadas por não identificação de nenhum contexto sintáctico-semântico.

⁸ As abreviaturas significam o seguinte: -conc (traço não concordância), i_or (início de oração), isol. (isolamento por pontuação), f_or (fim de oração), cop. (verbos copulativo), trs (verbos transitivo directo), itr (verbos intransitivos), tO (verbos transobjectivos), trs.prp (verbos que regem preposições), + conc (traço concordância), Tipo vb. (tipo de verbo).

Extr ^{ão}	Coordenação									Não coordenação			
	- conc	Posição			Tipo de verbos e estrutura argumental					Para- lelismo	+conc	Tipo vb. cop.	
		i_or	isol.	f_or	cop.	trs	itr	tO	trs. prp				
-ária	23	5	53	20	9	23	5	3	5	7	7	1	161
-eira	9	2	5	15	3	13	--	1	1	2	9	--	60
-esca	--	--	--	--	--	--	--	--	--	1	--	--	1
-ica	221	60	207	283	97	228	49	20	37	40	60	2	1304
-osa	23	12	33	15	14	29	13	2	5	22	7	1	176
Total	276	79	298	333	123	293	67	26	48	72	83	4	1702

Tabela 3: Distribuição dos contextos extraídos e analisados pelas categorias formadas

Os contextos mais produtivos em termos de identificação de uma estrutura de coordenação de advérbios em *-mente* são o traço de não concordância da forma adjectival com um núcleo nominal (16,2%) e a posição da estrutura na oração, nomeadamente o isolamento por sinais de pontuação (17,5%) e a posição em fim de oração (19,6%).

As expressões *única e exclusivamente* e *única e simplesmente* contabilizaram um total de 207 e 40 ocorrências respectivamente. Não se observou nenhum caso em que estas sequências tivessem outra análise que não a de coordenação de advérbios em *-mente*.

Em termos de rejeição de uma estrutura de coordenação de advérbios em *-mente*, o paralelismo de construção representa 4,23% dos casos categorizados em contextos de não coordenação. O traço +conc é o contexto mais produtivo em termos de não identificação de uma coordenação de advérbios em *-mente* (4,9%). Este facto pode estar relacionado com o contexto pós (hipotética) estrutura de coordenação. Se o que imediatamente segue o advérbio em *-mente* é um adjectivo que não é concordante em género e número com um núcleo nominal ou pertencente a uma qualquer outra classe gramatical, então a forma adjectival (antes da conjugação coordenativa), concordante, não poderá ser um advérbio com o sufixo *-mente* reduzido, mas um adjectivo, morfológica, semântica e funcionalmente:

- (6) ... terá visto travões excessivos à ocupação_[n: fem. sg] turística_[adj: fem. sg] OU simplesmente uma redução_[n: fem. sg] aos poderes absolutos....

Em (6), ‘uma redução’ não pode ser modificador de ‘ocupação’, pelo que se conclui que a concordância em género e número entre ‘ocupação’ e ‘turística’ deve-se a uma relação de dependência.

Como foi referido, o traço +conc pode, porém, conduzir a uma situação de ambiguidade. Os dados quantitativos relativos a ambiguidades são os seguintes:

Ambiguidade real (resolvida por traços prosódicos)	159	414
Ambiguidade sintáctica (resolvida por traços semânticos)	255	

Tabela 4: Resultados quantitativos das ambiguidades observadas

65,6% da ambiguidades observadas podem ser resolvidas com recurso à semântica. Significa isto que uma eficaz desambiguação automática de estruturas de coordenação de advérbios em *-mente* terá de ter em consideração, além de aspectos da morfossintaxe, questões semânticas, como as referidas acima no ponto 4.2.

Conclusão

A identificação automática de uma estrutura de coordenação de advérbios em *-mente* não é tarefa fácil um analisador automático. A não produção de sentidos e, estruturalmente, a forma adjectival que o advérbio com o sufixo *-mente* reduzido exhibe, tornam difícil a identificação de uma coordenação a nível adverbial e, conseqüentemente, a análise morfológica de advérbio para o primeiro elemento coordenado.

O presente estudo pretendeu mostrar que a busca em corpora de padrões sintáctico-semânticos podem auxiliar na identificação das estruturas de coordenação a nível adverbial, permitindo assim uma maior automatização e menor envolvimento de trabalho humano/manual.

Este estudo centrou-se na variante do português europeu e no registo jornalístico. No entanto, uma eficaz e abrangente análise automática de estruturas de coordenação de advérbios em *-mente* deveria abarcar outros registos. Assim, seria interessante observar o comportamento destas estruturas em outros corpora, não jornalísticos, bem como em outra(s) variedade(s) do português.

O prosseguimento natural deste estudo será a implementação dos contextos de desambiguação encontrados no analisador automático PALAVRAS (Bick, 2000) e a avaliação do seu desempenho relativamente às estruturas de coordenação de advérbios em *-mente*.

Agradecimentos

Este trabalho foi desenvolvido no âmbito das actividades da Linguateca, em particular no subprojecto “Floresta Sintá(c)tica”, que é uma colaboração entre o projecto VISL e a Linguateca com o financiamento do POSI, através do projecto da FCCN intitulado “Centro de Recursos Distribuído para o Processamento Computacional da Língua Portuguesa e, desta forma, os meus agradecimentos aos líderes dos projectos em colaboração, Diana Santos (Processamento Computacional do Português) e Eckhard Bick (projecto VISL, Universidade do Sul da Dinamarca). Externamente ao projecto, os meus agradecimentos ao Prof. Jorge Baptista, Universidade do Algarve, pelo tempo gasto nas discussões sobre os advérbios em *-mente*, que em muito me motivaram para este estudo.

Bibliografia

- Bache, Carl e Niels Davidsen-Nielsen (1997). *Mastering English*. Berlin: Mouton de Gruyter.
- Baptista, Jorge e Luís Faísca (2001). “Um filtro para palavras exóticas frequentes do Português”. In *Seminários de Linguística 4*. Faro: CELL, Universidade do Algarve. 65-86.
- Bayer, Samuel (1996). “The coordination of unlike categories”. *Language*, 72/3. 579-616.
- Bick, Eckhard (2000). *The Parsing System “PALAVRAS”. Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Århus University Press.
- Bosque, Ignacio (1989). *Las Categorías Gramaticales*. Madrid: Sintesis.
- Demonte, Violeta (1999). “El adjetivo: clases y usos. La posición del adjetivo en el sintagma nominal”. In Ignacio Bosque e Violeta Demonte (eds.). *Gramática Descriptiva de la Lengua Española. Vol. 1: Sintaxis Básica de las Clases de Palabras*. Madrid: Espasa Calpe. 129-216.
- Dik, Simon C. (1968). *Coordination: its Implications for the Theory of General Linguistics*. Amsterdam: North-Holland.
- Fanost, Claire Hue (1993). *El Adverbio*. Segunda edição (primeira edição, 1987). Madrid: Sociedad General Española de Librería.
- Gross, Maurice (1986). *Grammaire Transformationnelle du Français. Vol. 3: Syntaxe de l'Adverbe*. Paris: Asstril.
- Hudson, Jean (1998). *Perspectives on Fixedness: Applied and Theoretical*. Lund: Lund University Press.
- Mateus, Maria Helena, Ana Maria Brito, Inês Duarte e Isabel Hub Faria (1989). *Gramática da Língua Portuguesa*. 2ª edição. Lisboa: Caminho.
- Merello, Pilar Millán (1996). “Revisión crítica de la clasificación de los adverbios en *-mente* en la tradición gramatical española”. In Gerd Wotjak (ed.). *En Torno al Adverbio Español y los Circunstantes*. Tübingen: Günter Narr Verlag. 127-136.
- Molinier, Christian e Françoise Levrier (2000). *Grammaire des Adverbes: Description des Formes en -ment*. Genève: Librairie Droz.
- Mortega, Soledad Varela (1990). *Fundamentos de Morfología*. Madrid: Sintesis.

- Rocha, Paulo Alexandre e Diana Santos (2000). "CETEMPúblico: um Corpus de grandes dimensões de linguagem jornalística portuguesa". In Maria da Graça Volpe Nunes (ed.). *Actas do V Encontro para o Processamento Computacional da Língua Portuguesa escrita e falada (PROPOR' 2000)*. Atibaia. São Paulo. Brasil, 19 a 22 de Novembro de 2000. 131-140.
- Santos, Diana e Paulo Alexandre Rocha (neste volume). "AvalON: uma iniciativa de avaliação conjunta para o português".
- Santos, Diana (2000). "O projecto Processamento Computacional do Português: Balanço e perspectivas". In Maria da Graça Volpe Nunes (ed.). *Actas do V Encontro para o Processamento Computacional da Língua Portuguesa escrita e falada (PROPOR' 2000)*. Atibaia. São Paulo. Brasil, 19 a 22 de Novembro de 2000. 105-113.
- Santos, Diana e Eckhard Bick. "Providing Internet access to Portuguese corpora: the AC/DC project". In Gavriladou et al. (eds.). *Proceedings of the Second International Conference on Language Resources and Evaluation, LREC 2000*. Athens, 31 May- 2 June 2000. 205-210.