

Orações Relativas em Contextos Apresentativos: dados de um *corpus* oral

Rita Veloso

REPORT, Centro de Linguística da Universidade de Lisboa

Av. Prof. Gama Pinto, 2

1649-003 Lisboa, Portugal

rita.veloso@clul.ul.pt

Resumo*

Na sequência de trabalho anteriormente desenvolvido é nosso objectivo, neste artigo, explorar os tipos de orações relativas que surgem em contextos apresentativos num *corpus* de língua falada do português europeu.

Tomando com o ponto de partida a observação de um pequeno *corpus* de orações relativas de estratégia resumptiva recolhido pelo autor, confrontam-se agora as conclusões obtidas com dados extraídos de um *corpus* de língua falada. Iremos ver que estes nos mostram mais acerca de propriedades discursivas envolvidas na utilização de construções relativas em geral, na língua falada, do que acerca da estratégia resumptiva propriamente dita.

1 Introdução

Em Português Europeu (PE), uma das estratégias disponíveis para construção de orações relativas consiste na utilização de um pronome resumptivo (cf., para o português, Brito 1991, 1995, Peres e Mória 1995, Alexandre, 2000). Em trabalho anterior (Veloso, 2007), descrevemos as características mais regulares manifestadas por orações relativas de estratégia resumptiva, presentes num pequeno *corpus* recolhido por nós.

A descrição desta construção constitui um passo importante para o estudo das principais propriedades sintácticas e discursivas das diferentes estratégias para construção de orações relativas em PE e permitirá comparar padrões de propriedades descritas para outras línguas.

Partindo dos padrões encontrados, será fulcral verificar a sua representatividade num *corpus* de língua falada espontânea, contrastando os seus valores de ocorrência com os valores encontrados para os outros dois grandes tipos de estratégias de relativização, nomeadamente a canónica e a cortadora.

Só assim se poderá perceber realmente se os padrões observados são específicos para a estratégia resumptiva ou se se trata de padrões que, de forma mais generalizada, favorecem a ocorrência de orações relativas.

Deste modo, procedeu-se à extracção de orações relativas dentro de determinados contextos de um *corpus* oral de 320.000 palavras e observaram-se as regularidades aí manifestas.

Na secção 2. iremos resumir o trabalho desenvolvido anteriormente, descrevendo as regularidades encontradas num pequeno *corpus* de orações relativas de estratégia resumptiva e formulando possíveis caminhos de análise

Na secção 3. descreveremos o *corpus* C-ORAL-ROM, que utilizámos agora para verificação dos padrões encontrados anteriormente, e discutiremos as maiores dificuldades que se encontram no processo de extracção e classificação de dados extraídos de um *corpus*.

Na secção 4. apresentar-se-ão as regularidades encontradas nas estruturas-alvo deste trabalho e discutir-se-ão algumas implicações de nível teórico.

Na secção 5. iremos esboçar os caminhos a seguir na continuação da exploração do *corpus* com vista a alcançar-se uma uniformização dos fenómenos descritos bem como da análise a desenvolver.

2 Estratégia Resumptiva – principais propriedades encontradas

2.1 Descrição do *corpus*

O *corpus* utilizado inicialmente contém 139 orações relativas com pronome resumptivo, recolhidas manualmente pela autora entre 2004 e 2005, em situações de formalidade variada: 99 contextos (72%) são produzidos durante conversas espontâneas entre amigos ou colegas; 38 contextos (27%) foram extraídos de *Media*, sobretudo TV; 2 contextos (1%) foram produzidos em contextos formais, nomeadamente aulas universitárias.

* Este trabalho foi realizado no âmbito da bolsa de doutoramento concedida pela FCT (SFRH/BD/13752/2003).

2.1.1 Função de Processamento – contextos de ilhas

Como já descrito para muitas outras línguas por diversos autores, a estratégia de pronome resumptivo ocorre frequentemente quando o contexto de extracção contém uma ilha (McCloskey 1990, Prince 1990, Suñer 1998, Bianchi 2000, entre outros). O PE comporta-se da mesma forma em relação a este tipo de contextos. Das 139 frases do *corpus*, 52 (37.4%) resultam de extracção em contextos de ilhas.

- (1) A um antigo padrão meu / que eu estive quase para lhe dar uma pêra no focinho
- (2) Há coisas fantásticas // que eu nem sei como é que elas me acontecem

Dentro destes 52 contextos de extracção longa, 21 foram produzidos numa coordenação de estruturas relativas. Nestes casos, geralmente é o segundo membro do par coordenado quem exhibe o pronome relativo (cf. (3)), mas o pronome pode estar presente em ambos (cf. (4)):

- (3) porque isso são ordens que eu tenho e que não posso fugir a elas
- (4) Eles tinham umas 200 ou 300 cabeças de gado que ninguém as ia levar nem ninguém as ia buscar ao pasto

2.1.2 Estruturas Apresentativas

Verbo principal – dos 139 contextos em análise, 29 não continham a forma verbal que introduzia o DP relativizado, fosse por a estrutura

não o ter de facto, fosse por ter sido impossível reconstruir o contexto total. Dos 110 remanescentes, é interessante notar que 77% contêm um verbo apresentativo a seleccionar o DP relativizado (*ser* – 28 ocorrências; *ter* – 27; *haver* – 24; *estar* 5; e 2 ocorrências de verbos de aparição). Este tipo de estruturas aparece em exemplos de diversos artigos sobre diferentes línguas (Deulofeu 1981, Prince 1990, Suñer 1998, Bianchi 2000), apesar de não ser identificado como tal e de não dispormos, de momento, de valores estatísticos de ocorrências.

Determinante do antecedente – excluindo 6 contextos para os quais também não foi possível reconstruir o antecedente, nos 113 restantes o antecedente é introduzido por um determinante indefinido em 86% dos casos. Esta propriedade já tinha sido observada para outras línguas por Prince (1990), Suñer (1998) e Bianchi (2000), entre outros.

Estas duas características fazem do contexto preferencial para a estratégia resumptiva observado neste pequeno *corpus* um contexto de estrutura apresentativa (cf. Lambrecht 2000).

Na tabela abaixo apresenta-se um pequeno conjunto ilustrativo, apenas para os casos introduzidos por *haver* e *ser*.

Será agora imprescindível verificar num *corpus* de língua falada espontânea se este contexto se confirma de facto como um contexto específico para a estratégia resumptiva ou se se trata de um contexto que, de forma mais generalizada, favorece a ocorrência de orações relativas.

Há aqui		coisas	que	eu não sei exactamente o que é que eles querem dizer	com isto	
(na Índia) há	um	sítio	que /	as pessoas quando sabem que vão morrer / vão	para lá	
Há		problemas	que	às vezes as pessoas não sabem como lidar	com eles	
Há		coisas	que	eu acho que ninguém tem que passar	por elas	
Há		coisas	que	eu acho / que não vale a pena / a gente preocupar-se	com elas	
Há	imensas	crianças	que	os pais	lhes	dão tudo
Nestas coisas assim há		coisas	que	tu não sabes o que fazer	com isto	
Há	uns	amigos meus lá de baixo	que		eles	viam muito
(o Frota) e sobretudo é	um	actor	que	eu gosto de trabalhar	com ele	
É	uma	peessoa	que	eu não queria mesmo ter que votar	nele	
(a propósito do fôlego e do tabaco) É	uma	coisa	que	eu dei logo	por isso /	porra //
Ela é assim	uma	menina	que		ela	trabalha numa mercearia
É	uma	hipótese	que	já tínhamos falado	nela	
Isso são	uns	dados adicionais	que	eu também ainda não olhei	para eles	
E este livro / é	um	livro	que	eu andava com o olho	nele	
Este é	um	problema	que	eu já me debatia	com ele	
Ele era	um	rapaz	que	tanta rapariga	o	queria

2.1.3 Propriedades Discursivas

Com base nestas características, gostaríamos de poder relacionar diferentes análises, oriundas de diversos campos teóricos.

Lambrecht (2000) relaciona orações relativas predicativas, que ocorrem maioritariamente em contextos apresentativos, com as propriedades discursivas envolvidas. De acordo com este autor, nestas estruturas o antecedente é o Foco da oração principal, mas, simultaneamente, é o Tópico da oração relativa, a qual expressa informação nova acerca desse tópico. Apesar de as orações em análise aqui não caberem na classe das relativas predicativas, interessa-nos preservar esta ideia da duplicidade de funções discursivas exibida pelo antecedente da oração relativa, pelo menos nestes contextos apresentativos.

Já Prince (1990), no seguimento de Heim (1983), apresenta uma análise semelhante para orações relativas indefinidas com estratégia resumptiva: “The NP is indefinite, it represents a ‘Brand-new entity (...) and the hearer must add that entity, or construct a new file-card (...) the relative clause serving simply to predicate some property of that entity, that is, where the appropriate file card has already been independently constructed”.

Por outro lado, Kato (1993) relaciona a estratégia resumptiva com propriedades de Tópico, defendendo que, neste tipo de orações relativas, o local de extracção para relativização é uma posição de Tópico, semelhante à da construção de Clitic Left Dislocation.

Através da interacção das suas propriedades discursivas e sintácticas, é possível estabelecer o paralelo entre a análise das orações relativas em contextos apresentativos e a análise das relativas de estratégia resumptiva, uma vez que estas aparecem sobretudo no contexto daquelas.

3 O corpus C-ORAL-ROM

3.1 Descrição do corpus

O corpus português C-ORAL-ROM foi construído recentemente no âmbito do projecto homónimo² e tem cerca de 318 000 palavras (correspondendo a cerca de 30 horas de gravação). A utilização deste corpus oferece diversas vantagens: trata-se de um corpus contemporâneo, uma vez que cerca de 93,3% é posterior a 1995, e é constituído por

² C-ORAL-ROM é o acrónimo do projecto intitulado *Integrated reference corpora for spoken romance language; tools of analysis; standard linguistic measures for validation in HLT*, financiado pelo *Information Society Technologies Programme (IST)*, com o contrato n.º IST-2000-26228.

exemplos de diferentes registos do Português falado, como se pode ver na tabela seguinte:

INFORMAL		
Family /	Conversations	24.449
/Private	Dialogs	62.738
	Monologs	46.005
		133.192
Public	Conversations	1.817
	Dialogs	23.119
	Monologs	7.710
		32.646
TOTAL		165.838

FORMAL		
Natural	Business	10.215
Context	Conferences	9.750
	Law	6.315
	Political Debate	8.923
	Prof. Explanation	6.473
	Preaching	6.127
	Political Speech	8.649
	Teaching	9.822
		66.274
Media	Interviews	14.570
	Meteo	1.930
	Reportages	10.762
	Scientific Press	9.923
	Sport	5.676
	Talk Show	17.396
	News	1.859
		62.116
Telephone	Private	24.365
TOTAL		152.755

3.2 O processo de extracção

A existência de anotação morfossintáctica ao nível das categorias principais³ permite extrair com alguma celeridade as orações relativas aí existentes (deixando de lado a margem de erro inevitável num processo de anotação automática, se bem que essa margem de erro tenha sido bastante reduzida graças à revisão manual de algumas categorias).

No entanto, o facto de a etiqueta atribuída aos pronomes relativos ser a mesma para os pronomes exclamativos e interrogativos, exige um trabalho bastante moroso na selecção dos dados relevantes. Por outro lado, são necessárias várias tentativas de extracção até se encontrarem as delimitações de contexto necessárias para observação de dados interessantes. Por fim, as recorrentes ambiguidades de estruturas (por exemplo, entre interrogativas indirectas e relativas sem antecedente expresso – cf. Móia 1992 –, entre relativas e explicativas ou ainda entre relativas e consecutivas) tornam o processo de classificação e organização de dados numa tarefa extremamente difícil e longa. É necessário encontrar testes fiáveis que permitam, com relativa segurança, atribuir a etiqueta certa à estrutura em questão.

³ Para uma apresentação dos critérios de anotação morfossintáctica ver Bacelar do Nascimento et al. (2005).

Deste modo, optou-se para este trabalho, por extrair apenas contextos de pronomes-*wh*⁴ quando antecidos de verbos apresentativos. Os verbos seleccionados foram *aparecer*, *estar*, *haver* e *ter* (limitações de tempo impediram-nos de analisar para este trabalho os contextos extraídos para o verbo *ser*).

Os contextos foram seguidamente classificados e organizados, isolando-se as estruturas relativas presentes. Nalguns casos foi necessário ir ao ficheiro de texto para completar o contexto; noutros foi necessário ouvir o ficheiro de som para esclarecer alguma dúvida. As estruturas potencialmente ambíguas foram postas de lado, para análise posterior. Também não se tiveram em conta estruturas em que o elemento relativo tinha como antecedente um constituinte não relacionado com o verbo em questão.

4 Descrição de resultados

Os resultados obtidos, não sendo particularmente reveladores para a análise das orações relativas enquanto objecto sintáctico puro, são, no entanto, interessantes para o estudo da interacção entre a sintaxe e o discurso.

Na pesquisa de pronomes-*wh* invariáveis, obtiveram-se 882 contextos: 13 para o verbo *aparecer*, 165 para o verbo *estar*, 321 para *ter* e 383 para *haver*. Depois de seleccionadas as estruturas relativas em contextos de verbos apresentativos, ficámos com um *corpus* de 454 contextos (*aparecer* – 7; *estar* – 28; *ter* – 122; *haver* – 297). Na pesquisa de pronomes-*wh* variáveis obtiveram-se 25 contextos, dos quais só um correspondia a uma oração relativa em contexto apresentativo. Esta surge com o verbo *haver* e o pronome relativo em causa é *cujo*. Curiosamente, resulta de uma reformulação:

(5) ora / há indivíduos que têm / cujo gene / é deficitário // portanto é um gene mutante // ⁵

Há mais 3 casos em que esta reformulação se dá, nomeadamente para a locução pronominal *o qual*.⁶

⁴ Independentemente do estatuto de complementador de *que*, nalgumas construções relativas – cf. para o Português, Tarallo (1985), Faria e Duarte (1989), Brito (1991), Brito (1995), a equipa de anotação do projecto optou por manter uma classificação mais tradicional).

⁵ Limpámos os exemplos de lemas e etiquetas de anotação para facilitar a leitura.

⁶ Estes exemplos ocorrem no *corpus* de 454 contextos, com pronomes invariáveis, e não no de 25 contextos, com pronomes variáveis, dado o pronome relativo *o qual* ter sido etiquetado como locução pronominal.

(6) e depois / há um / que ela / &pe / pelo qual ela se apaixonou / e a meio / chega à conclusão

(7) e eu tive colegas / &ah / na minha turma eu tive colegas mestiços / que ainda hoje / dos quais sou amiga //

(8) e / e nós / temos direito / esses direitos que nós / pelos quais / &eh lutámos e / e / e

Nestes 4 exemplos é notório que a reformulação impediu o aparecimento de uma relativa de estratégia resumptiva ou cortadora. Este facto permite-nos pensar que, apesar da espontaneidade dos textos orais existentes no *corpus* C-ORAL-ROM, poderá haver ainda algum auto-controlo por parte dos falantes.

4.1 Representatividade

Considerando que o *corpus* contém 8.215 pronomes *wh* anotados como invariáveis e que destes já se procedeu à análise de 3.128 contextos (análise fora do âmbito deste artigo), tendo-se identificado 68.7% de estruturas relativas, prevê-se, informalmente, que, a manter-se esta proporção, haja um total de aproximadamente 5.640 orações relativas no *corpus*. Os 454 contextos extraídos para este trabalho corresponderão, nesse caso, a 8% das estruturas relativas do *corpus*.

4.2 Estratégia utilizada

Em relação ao tipo de construção relativa envolvida nos 454 contextos, podemos confessar que os resultados foram decepcionantes, pois esperávamos uma mais expressiva frequência de orações com estratégia resumptiva. No entanto, as construções relativas estudadas apresentam a seguinte distribuição: 91.6% de estratégia canónica, 5% de estratégia cortadora e 2% de estratégia resumptiva⁷.

Uma vez que os verbos pertenciam à classe identificada anteriormente como favorecendo o aparecimento de relativas de estratégia resumptiva e que 86.5% das estruturas apresentam um determinante indefinido a introduzir o antecedente da relativa, fomos tentar perceber que outros factores poderiam estar a interferir com os resultados.

4.2.1 Constituinte relativo

84.6% das estruturas apresentam o morfema *que* isoladamente no constituinte relativo; 6.6% apresentam *que* precedido de preposição; 3.7% apresentam *quem* e 3.3% apresentam *onde*.

Cruzando agora estes dados com os anteriores, vemos que 83.6% das construções canónicas apresentam *que*.

⁷ Resultados de alguma forma semelhantes aos encontrados em Alexandre (2000).

4.3 Função do constituinte relativo

Para 75% das estruturas, o constituinte relativo tem a função sintáctica de sujeito; em 12%, o constituinte relativo tem a função (semântica) de Locativo (espacial ou temporal) e em 8% tem a função de complemento directo.

Continuando o cruzamento de valores, verifica-se que as construções relativas canónicas apresentam um constituinte relativo com função de sujeito em 80% dos contextos e com a função de complemento directo em 7.9%.

Ou seja, em 87.9% das estruturas, a estratégia canónica não envolveu um constituinte relativo complexo (são, por assim dizer, relativas-*that* e não relativas-*wh*).

4.4 Verbo da oração relativa

Os verbos que surgem na oração relativa pertencem aos tipos mais variados. No entanto, não podemos deixar de observar que *ser* ocorre em 21.5% dos casos, seja numa estrutura identificadora seja numa estrutura caracterizadora, e em 2.4% em estruturas passivas.

É nas estruturas canónicas que este facto se torna mais interessante: *ser* está presente em 22.5% destas estruturas; *ser* passivo ocorre em 2.4%. Se aliarmos estes valores aos obtidos para os casos em que se recorre ao *-se* passivo (6.7%), encontramos 9.1% de estruturas em que a construção passiva permite mapear o constituinte relativo com o sujeito.

5 Conclusões

Os dados extraídos do *corpus* não permitiram confirmar a hipótese anteriormente formulada de que o contexto de verbos apresentativos favorecia o aparecimento de orações relativas de estratégia resumptiva. Esperamos, no entanto, que, quando se alargar o número de verbos em análise, se obtenham melhores resultados. De qualquer modo, apenas quando dispusermos dos valores finais para as diferentes estratégias, poderemos avaliar a representatividade deste contexto na utilização daquela estratégia.

Também a exploração de contextos de coordenação de orações relativas em estruturas apresentativas permitirá alcançar resultados mais significativos. Note-se que nos 454 contextos analisados foram detectados 6 casos de estruturas marginais no segundo elemento da coordenação (4 resumptivas e 2 cortadoras, que não foram contabilizadas acima, dado que não tínhamos feito o levantamento sistemático de estruturas de coordenação neste contexto). Este valor aumentará consideravelmente quando se alargar o contexto de forma a que contenha as restantes orações relativas

relacionadas (o que não foi possível fazer para este trabalho).

Em relação à utilização da estratégia canónica na língua falada, observaram-se alguns aspectos interessantes, nomeadamente no que diz respeito à interacção entre a sintaxe e o discurso: se acreditarmos que uma construção relativa envolve realmente uma estrutura de topicalização, o recurso a estruturas em que o constituinte relativo corresponde ao tópico não marcado – o sujeito – permite evitar complexidade estrutural. A utilização de construções passivas vem também contribuir para essa simplificação estrutural.

Parece-nos possível afirmar que, na língua falada, os falantes recorrem à estrutura informacional não só para valorizar determinados constituintes mas também para evitar complexidades estruturais.

A exploração da interacção entre sintaxe e discurso parece-nos ser um caminho interessante a prosseguir no estudo das estruturas relativas e das propriedades que condicionarão a utilização de diferentes tipos de estratégias na sua construção.

6 Agradecimentos

Ao CLUL, por estes 12 anos de convivência saudável, pela amizade e pelo estímulo.

Referências

- N. Alexandre. 2000. *A Estratégia Resumptiva em Relativas Restritivas do Português Europeu*, Diss. de Mestrado, Lisboa, FLUL.
- M.F. Bacelar do Nascimento, J. Bettencourt Gonçalves, R. Veloso, S. Antunes, F. Barreto e R. Amaro. 1995. «The Portuguese Corpus», in *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*, E. Cresti e Massimo Moneglia (eds.), John Benjamins Publishing Company, Amsterdam, Philadelphia, cap. V, pp. 163-207.
- V. Bianchi. 2000. «On resumptive relatives and the theory of LF chains», ms., Scuola Normale Superiore, Pisa.
- C. Blanche-Benveniste. 1990. «Usages normatifs et non normatifs dans les relatives en français, en espagnol et en português» J. Bechert, G. Bernini e C. Buridant. (eds.) *Toward a Typology of European Languages*, Berlin, New York, de Gruyter, pp. 317-335.
- A.M. Brito. 1991. *A Sintaxe das Orações Relativas em Português: Estrutura, mecanismos interpretativos e condições sobre a distribuição dos morfemas relativos*, Diss. de Doutoramento defendida em 1988, Porto, INIC.

- A.M. Brito. 1995. «As Orações Relativas Restritivas nas Variantes Culta e Oral em Quatro Línguas Românicas, com Incidência Especial em Português», *Lusorama*, 27, Frankfurt, Instituto Camões pp. 70-81.
- J. Deulofeu. 1981. «Perspective linguistique et sociolinguistique dans l'étude des relatives en français», *Recherches sur le Français Parlé*, 3, GARS, Provence, pp. 135-193.
- I. Faria e I. Duarte. 1989. «O Paradoxo da Variação: Aspectos do Português Europeu», *Revista Internacional de Língua Portuguesa*, 1, Lisboa, Associação das Universidades de Língua Portuguesa, pp. 21-27.
- M. Kato. 1993. «Recontando a História das Relativas em uma Perspectiva Paramétrica», I. Roberts and M. Kato (eds.) *Portugues Brasileiro – Uma viagem diacrônica*, Campinas, Unicamp.
- K. Lambrecht. 2000. «Prédication seconde et structure informationelle. La relative de perception comme construction présentative», *Langue Française* 127, 49-66.
- J. McCloskey. 1990. «Resumptive Pronouns, A'-binding, and levels of representation in Irish», *Syntax and Semantics*, vol. 23: *The syntax of modern Celtic languages*, R. Hendrick (ed.), 199-248, New York: Academic Press.
- J. Peres e T. Móia. 1995. *Áreas Críticas da Língua Portuguesa*, Lisboa, Editorial Caminho.
- E. Prince. 1990. «Syntax and Discourse: a Look at Resumptive Pronouns», *Berkeley Linguistic Society* 16, pp. 482-497.
- M. Suñer. 1998. «Resumptive Restrictive Relatives: A Crosslinguistic Perspective», *Language* 74, pp. 335-364.
- F. Tarallo. 1985. «The Filling of the Gap: *pro*-drop Rules in Brazilian Portuguese», L. King e C. Maley (eds.), *Selected Papers from the XIIIth Linguistic Symposium on Romance Languages*, Amsterdam, Philadelphia, John Benjamins Pub., pp. 355-375.
- R. Veloso. 2007. «European Portuguese Relative Clauses in a Spoken Corpus», poster apresentado ao *Corpus Linguistics 2007*, realizado em Birmingham, de 28 a 30 de Julho.