

UNIVERSIDADE DE LISBOA
FACULDADE DE LETRAS
DEPARTAMENTO DE LINGÜÍSTICA GERAL E ROMÂNICA

**VARIATION AND CHANGE IN THE
SYNTAX OF RELATIVE CLAUSES**

New evidence from Portuguese

Maria Adriana Príncipe Cardoso

Tese orientada por:
Ana Maria Martins
Mark de Vries

DOUTORAMENTO EM LINGÜÍSTICA
(Linguística Histórica)

2010

This dissertation was supported by national and European funding
– grant SFRH/BD/22475/2005 –

FCT Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR Portugal



UNIÃO EUROPEIA
Fundo Social Europeu

If there are no data, there cannot be any theory. If there is no theory, there can hardly be any understanding.

A. Devine and L. Stephens,
Latin Word Order: Structured Meaning and Information

Acknowledgements

When I decided to do my PhD in historical linguistics, I asked Ana Maria Martins if she wanted to be my supervisor. She said 'Yes', but on one condition: I had to work on generative syntax. I accepted without hesitation. I was not fully convinced by the generative grammar, but I was pretty sure that I wanted her to be my supervisor, so I took the risk.

So the adventure began. From the first moment, the amount of papers, theses, and conferences developed within the generative framework overwhelmed me. But as I began to emerge from the chaos, I began to develop a critical perspective and to select my favourite papers, books, and linguists. Mark de Vries was at the top of my list, with his Syntax of Relativization. I started to talk to Ana Maria Martins about his work, showing my enthusiasm for the way he combined linguistic theory with typological research. As a result, I was challenged again. First, why don't you go to Groningen and work with him? Then, why don't you write the thesis in English and ask him to be your co-supervisor?

This thesis is the result of all these challenges. I would like to express my gratitude to Ana Maria Martins, for challenging me, for believing in my work, for teaching me how to question the data correctly and how to combine the philological dimension with the theoretical research. My gratitude also goes to Mark de Vries, for giving me the opportunity to discuss my work with him, for commenting on it, and for showing me how generative syntax works in its different dimensions.

To the staff of the Centro de Linguística da Universidade de Lisboa, my thanks for being my host institution and for giving me all the necessary conditions for the development of my work. In particular, I would also like to thank my dearest 'cordiais' Ana Maria Martins, Ernestina Carrilho, Maria Sousa Lobo, Catarina Magro, Sandra Pereira, and Cristina Morgado for the friendship, good company, complicity, and support. Special thanks also go to João Saramago, Gabriela Vitorino, Luísa Segura, and Amália Andrade, for their hospitality and for always making me feel a part of the team.

To the linguistics department in Groningen, my thanks for having accepted me as a visiting scholar and for making my stay at Groningen a great experience. Special thanks are due to Mark de Vries, Jan Koster, Jan-Wouter Zwart, Marlies Kluck, Herman Heringa, Radek Šimik, Jordi Fortuny, and Janneke ter Beek, for allowing me to participate in the weekly advanced syntax seminar, for the discussions and comments. I would also like to thank them for the unforgettable dinners (namely the one with special piano performances) and for the 'zwerftocht' in the Waddenzee.

To Maria Francisca Xavier, André Eliseu, Ana Maria Martins and Ernestina Carrilho, I am extremely grateful for their help in allowing me to become a researcher at the Faculdade de Ciências Sociais e Humana (Universidade Nova de Lisboa), Instituto de Linguística Teórica e Computacional, and Centro de Linguística da Universidade de Lisboa.

To my degree and master's professors at the Faculdade de Ciências Sociais e Humana (Universidade Nova de Lisboa), for having given me the tools to survive in the different dimensions of my activities as a researcher and teacher.

To the best teachers I ever had: Maria José, my 6th grade Portuguese teacher; Rosalina, my 10th and 11th grade Latin teacher; and Eduardo Paiva Raposo, my Syntax teacher on the Master's.

To my colleagues at Escola de Educação de Lisboa, thank you for recognizing my work and for waiting for my return. In particular, I would like to thank Otilia Sousa, Susana Pereira, and Encarnação Silva, for the confidence they showed in my work and for the unconditional support. Special thanks should also go to my former colleagues, who despite the physical distance have always been a phone call away: Dulce Sarroeira, Ana Silva, Luísa Amado, Luísa Matos, Manuel Luís Costa, and Conceição Aleixo.

To all those linguists with whom I had linguistic and non-linguistic discussions during conferences, presentations, and seminars and to all people who I asked for opinions on grammaticality and other linguistic-related things, I offer my thanks. I am not able to recall all the people that deserve to be mentioned here, but I would like to mention the members of the G4 (Gramática Gerativa às Quartas), Ana Maria Brito, Rita Marquilhas, Nélia Alexandre, Beatrice Santorini, Theresa Biberauer, João Costa, Volker Struckmeier, and Jairo Nunes.

Closer to me, I deeply thank Susana Pereira for sincerely wishing the best for me and for making my life easier when I needed it most.

And, better later than never, I would really like to thank Catarina Magro, for being my friend, for making my life more interesting, and for having shared so many adventures with me on this long journey. It's true, without you, things would have been much less fun! I am also particular indebted to André Eliseu, for having taught me so many things since we met (in 1997!) and for having helped me to solve some of the difficulties I faced along the way. Having finished this thesis, I promise to be a better 'godmother'!

I also would like to thank my friends who are not involved in linguistics: Cláudia, Patrícia, Rita, Helena, Joana, Alexandra, Sara, Ângelas, Sílvia, Mário, Nuno, Sérgio, Zénon, and André. Thank you for the continuous support.

Now the moment has come to thank my closest family. I'm grateful to my parents, Maria João and Carlos, my sister Cristiana (and Nuno), my aunt Josete and my uncle Hermínio, my cousins Paula, Nicolau and Mariana, who support me unconditionally. They were even more anxious than me to see this dissertation completed.

Finally, I want to thank Carlos for encouraging me throughout the whole course of this work and for filling my life with so many memorable moments.

Abstract

This dissertation sheds light on language variation and change from a generative syntactic perspective, based on a case study of relative clause constructions in Portuguese and other languages. The research methodology adopted involves comparative syntax (see, e.g., Cinque and Kayne 2005), both in the diachronic and the synchronic dimensions: contemporary European Portuguese is systematically compared with earlier stages of Portuguese; moreover, Portuguese is compared with other languages, in particular Latin, English, Dutch, and Italian.

Furthermore, the present research contributes to the theoretical debate on the structural analysis of three different phenomena: restrictive relatives, appositive relatives, and extraposition. Two important findings are (i) that competing theoretical analyses need not be either false or true universally, but could be instrumental in explaining language variation (both diachronically and synchronically); and (ii) simple lexical changes concerning the availability of (abstract) functional items can have dramatic consequences in the behavior of certain 'construction types' in a particular language. As for the structural analyses of relative clauses and extraposition, both Kayne's (1994) raising analysis and De Vries' (2006b) specifying coordination analysis proved to be central to the understanding of the issues dealt with in this dissertation.

Ultimately, this dissertation demonstrates that cross-linguistic and diachronic contrasts offer invaluable means of enhancing our understanding of various linguistic phenomena, thereby contributing also to our understanding of the interaction between principles and parameters.

Keywords: *relative clauses, remnant-internal relativization, extraposition, appositive relative clauses, comparative syntax, diachronic generative syntax*

Resumo em Português (Summary in Portuguese)

A presente dissertação visa contribuir para a compreensão dos fenómenos de variação e mudança linguísticas, centrando-se no estudo de construções de relativização em português e em outras línguas.

A perspectiva adoptada neste estudo insere-se na área de investigação que tem sido designada de *sintaxe comparada* (cf., a título de exemplo, Cinque e Kayne 2005). Esta opção metodológica manifesta-se em dois planos distintos. Por um lado, é estabelecida uma comparação entre os dados de estádios anteriores do português e do português europeu contemporâneo. Por outro lado, estes dados são confrontados com dados de outras línguas, nomeadamente do latim, inglês, holandês e italiano.

Esta metodologia foi determinante para o desenvolvimento do presente estudo. Permitiu, por exemplo, mostrar que as construções de relativização não exibem um comportamento uniforme no decurso da história do português. Permitiu ainda concluir que, no que diz respeito a algumas das construções investigadas, há uma clara coincidência entre os dados diacrónicos do português e os dados de línguas germânicas contemporâneas (com particular ênfase para o inglês e holandês). O enquadramento conceptual e os instrumentos associados à investigação comparativa possibilitaram assim ultrapassar algumas das limitações inerentes à investigação em linguística histórica, nomeadamente no que diz respeito à natureza limitada das fontes escritas e à impossibilidade de manipular dados.

A análise apresentada toma como referência o quadro da Teoria de Princípios e Parâmetros (Chomsky 1981 e trabalhos posteriores), na sua versão minimalista (Chomsky 1993, 1995 e trabalhos subsequentes). A variação e mudança linguísticas são interpretadas com base no modelo proposto por Lightfoot (1991, 1999 e trabalhos posteriores), que estabelece uma relação clara entre a mudança linguística e a aquisição da linguagem. A explicação dos fenómenos de variação linguística beneficia ainda dos contributos de Kroch (1989, 1994, 2001) e da investigação em sintaxe comparada paramétrica (cf., e.g., Holmberg and Roberts 2010). A este respeito, cumpre referir que a adopção de um modelo teórico foi fundamental para os resultados alcançados neste trabalho. Para mencionar apenas algumas das mais valias obtidas, o enquadramento teórico foi decisivo na organização, descrição e explicação dos dados. Permitiu ainda fazer predições importantes, que orientaram a pesquisa de

corpora numa fase adiantada dos trabalhos. Parafrazeando a epígrafe desta dissertação, *não pode haver teoria sem dados e, sem teoria, dificilmente há compreensão.*

A investigação que se apresenta assenta em fundamentação empírica proveniente de diversas fontes, em função dos diferentes objectivos delineados.

No que concerne aos dados da história do português, o período compreendido entre os séculos XIII e a primeira metade do século XVI é estudado com base nos textos editados em Martins (2001), em *Documentos Portugueses do Noroeste e da Região de Lisboa*. Os dados recolhidos são complementados por outras fontes, em particular por textos disponíveis no *Corpus Informatizado do português Medieval* (Xavier, Coord.) e no *Corpus Histórico do Português Tycho Brahe* (Galves, Coord.). São também consideradas algumas edições de documentos da época, nomeadamente: *Crónica do Conde D. Pedro de Meneses* (ed. Brocardo 1997); *Livro de Linhagens do Conde D. Pedro* (ed. Brocardo 2006); *Demanda do Santo Graal* (ed. Piel and Nunes 1988); *Livro dos Ofícios* (ed. Piel 1948); *Crónica de D. Fernando* (ed. Macchi 1975); *Gil Vicente: todas as obras* (coord. Camões 1999). Para o período compreendido entre a segunda metade do século XVI e o século XIX, a fundamentação empírica é proveniente do *Corpus Histórico do Português Tycho Brahe* (Galves, Coord.), bem como dos textos disponibilizados em CARDS - *Cartas Desconhecidas* (Marquilhas, Coord.) e *Corpus do Português – CdP* - (Davies and Ferreira 2006-). Tal como para o período anterior, são igualmente consideradas edições de outros documentos da época: *Os Autos do Processo de Vieira na Inquisição* (ed. Muhana 1995); *Documentos para a História da Inquisição em Portugal* (ed. Pereira 1987); *Inquisição de Évora: dos Primórdios a 1668* (ed. Coelho 1987).

Para o português europeu contemporâneo, para além das minhas próprias intuições e dos juízos de outros falantes nativos, são tidos em conta textos provenientes de diversas fontes. Os corpora consultados são: *Corpus de Extractos de Textos Electrónicos MCT/Público* (Rocha e Santos 2000), *Corpus de Referência do Português Contemporâneo* (Bacelar do Nascimento, Coord.), *Corpus do Português Oral* (Bacelar do Nascimento, Portuguese Coord. 2005), *Corpus Dialectal para o Estudo da Sintaxe* (Martins, Coord.).

No que diz respeito à evidência empírica relativa a outras línguas, para além de dados recolhidos em textos da especialidade, são considerados dados dos seguintes corpora: *Corpus of Contemporary American English* (see Davies 2008-), *Penn Parsed*

Corpus of Modern British English (Kroch et al. 2010), *Penn-Helsinki Parsed Corpus of Early Modern English* (Kroch et al. 2004). Os dados do holandês, italiano e latim são maioritariamente provenientes de estudos sobre construções de relativização (com excepção do holandês, língua para a qual se apresentam também produções de vários falantes).

Do ponto de vista teórico, a presente dissertação contribui para o debate que se tem gerado em torno da análise sintáctica das orações relativas restritivas, orações relativas apositivas e extraposição. Duas conclusões gerais podem ser destacadas: (i) as diferentes análises apresentadas na literatura para dar conta destes fenómenos linguísticos não têm necessariamente de ser consideradas verdadeiras ou falsas (ou correctas/incorrectas), podendo constituir instrumentos fundamentais para a compreensão da variação linguística (sincrónica e diacrónica); (ii) simples mudanças no léxico relativas à presença de itens funcionais abstractos (i.e., sem realização fonética) podem ter um impacto bastante significativo nas propriedades e estrutura sintáctica das tradicionais ‘construções linguísticas’.

A dissertação é constituída por três estudos principais, que se descrevem de seguida.

O primeiro estudo apresenta e discute o fenómeno de *remnant-internal relativization* (RIR). Este termo é usado para referir os contextos de não adjacência entre o núcleo de uma oração relativa restritiva e o seu modificador/complemento, como se ilustra esquematicamente em (1):

(1) [_{S-matrix} ... [_N [_{RC} ... modificador/complemento ...]]] ...]

Dois padrões de ordem de palavras são identificados na história do português:

- RIR com o modificador/complemento em posição final (no interior da oração relativa):

(2) os livros que eu compus da philosophia

(D. Pedro, *Livro dos Officios*, séc. XV, in Pádua 1960: 70, cit. em Martins 2004)

(3) e pasará huũ rrio que perhy core dagoa doce
(*Carta de Pero Vaz de Caminha*, fol. 3v., séc. XV, cit. em Martins 2004)

(4) eram sobrinhas da molher que faleseo de lamsarote rodrigues
(Coelho 1987; *Inquisição de Évora: dos primórdios a 1668*; 1612)

(5) como eu me encontro num estado miseravel pella falta que há do vinho
(CARDS 4002; 1890)

- RIR com o modificador/complemento na periferia esquerda da oração relativa:

(6) e qualquer que de nos primeiro morer
(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; 1266)

(7) e esto por prool e verdade de hũa Licença que do dito senhor pera ello tenho
(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; 1540)

(8) e o deradeiro que delas fiquar posa amte de sua morte nomear a terceira
(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; 1528)

(9) Nas bombas que de fogo estão queimando
(Ramos 1982; L. de Camões, *Os Lusíadas*, V, 90; séc. XVI)

Estes padrões coexistem com o padrão mais frequente, que envolve adjacência entre o núcleo e o seu modificador/complemento, como em (10):

(10) [_{S-matrix} ... [N modificador/complemento [_{RC} ... t ...]] ...]

O objectivo central do presente estudo é demonstrar que a construção RIR constitui evidência para a elevação para as orações relativas restritivas por elevação (*raising analysis*), proposta originalmente por Schachter (1973) e Vergnaud (1974, 1985) e recentemente adoptada (e actualizada) por Kayne (1994), Bianchi (1999) e De Vries (2002).

Numa primeira etapa do estudo, mostra-se que a análise clássica por *adjunção* (cf. Chomsky 1977, Brito 1991, entre outros) não permite explicar a presença de modificadores/complementos no interior da oração relativa. Dois argumentos são apresentados para justificar esta impossibilidade: (i) assumindo que os

modificadores/complementos do nome são compostos com o nome numa posição externa à oração relativa, os contextos de RIR teriam de envolver a descida (*lowering*) dos modificadores/complementos para uma posição que não c-comanda o seu vestígio; (ii) se o núcleo e o seu modificador/complemento fossem gerados em posições distintas (o núcleo numa posição externa à oração relativa e o modificador/complemento numa posição interna), ficaria por explicar a existência de uma dependência semântica e de uma relação de selecção entre estes elementos.

Numa segunda etapa, mostra-se que a construção RIR pode ser derivada pela combinação da análise de elevação das orações relativas com o apagamento de cópias em PF (Bošković e Nunes 2007). De forma esquemática, propõe-se que os casos de RIR e os casos de adjacência entre o núcleo e o seu modificador/complemento são derivados por uma estrutura como em (11), implementada em termos da teoria do movimento como cópia:

(11) [os [livros da philosophia]ⁱ [que eu compus [livros da philosophia]ⁱ]]

Em função da estrutura informacional e de requisitos relacionados com a satisfação dos traços EPP de algumas categorias funcionais, mais do que uma cadeia de movimento pode ter lugar e podem ocorrer diversos tipos de apagamento de cópia (incluindo apagamento de cópias mais baixas e apagamento de partes distintas de diferentes cópias - *scattered deletion*), de forma a derivar os diferentes padrões de ordem de palavras identificados nestas construções.

Os dados relativos à construção RIR são ainda integrados no contexto mais geral de ocorrência de sintagmas nominais descontínuos na história do português.

O segundo estudo apresentado nesta dissertação discute a extraposição de orações relativas restritivas. Neste âmbito, o termo extraposição é usado para referir os contextos em que não existe adjacência entre a oração relativa e o seu antecedente, ocorrendo entre ambos material lexical da matriz, como se observa no esquema em (12). De forma a facilitar a leitura, o material da matriz que ocorre entre o antecedente e a oração relativa é destacado graficamente (sublinhado).

(12) [... [antecedente] OR...]

Em português europeu contemporâneo, as orações relativas restritivas podem ocorrer extrapostas, como se ilustra em (13)-(16):

(13) Ainda por cima, dá-se conta de que as obras não têm licença camarária e faz diligências na Câmara das Caldas da Rainha que levam ao seu embargo

(CETEMPúblico -Primeiro milhão)

(14) Não havia ontem nenhum aluno na minha aula que estivesse concentrado.

(15) Encontrei uma pessoa ontem que não via há muito tempo.

(http://coisasbelasesujas.blogspot.com/2004_09_01_archive.html)

(16) Houve alguém no meio da noite que decidiu agarrar uma cana que supostamente seria do Aranha (...)

(exemplo recolhido em: <http://www.pescador.com.pt/livre/viewtopic.php?f=9&t=1772>)

Tal possibilidade também se regista na história do português. Contudo, em fases anteriores do português, a extraposição de orações relativas restritivas é mais permissiva do que em português europeu contemporâneo. O mesmo tipo de contraste se observa em diferentes línguas contemporâneas. A este respeito, é interessante notar que as línguas germânicas (em particular o inglês e o holandês) exibem possibilidade de extraposição generalizada, opondo-se desta forma ao português europeu contemporâneo (mas aproximando-se claramente da situação observada na diacronia do português).

Os contrastes observados, quer na dimensão sincrónica quer na diacrónica, podem ser descritos de acordo com os seguintes parâmetros: (i) efeito de definitude; (ii) extraposição a partir de posições encaixadas; (iii) extraposição a partir de constituintes pré-verbais.

Considerando, a título de exemplo, a situação do português contemporâneo, verifica-se que: (i) o antecedente de uma oração restritiva extraposta pode ser um sintagma nominal ‘fraco’, mas não um sintagma nominal ‘forte’ (Milsark 1974) (cf. (17)-(18)); (ii) o antecedente da oração restritiva extraposta não pode ser um constituinte encaixado, como é o caso do complemento de uma preposição (cf. exemplo (19) e (20); note-se que no último caso há um nível adicional de encaixe); (iii) o antecedente da relativa extraposta pode ser um sujeito pré-verbal, mas não um sujeito pós-verbal (cf. (21)-(22)).

- (17) a. Encontrei um rapaz no cinema que perguntou por ti.
 b. *Encontrei o rapaz no cinema que perguntou por ti.
- (18) a. Foram publicados dois livros recentemente que vale a pena ler.
 b. *Foram publicados aqueles livros recentemente que vale a pena ler.
- (19) a. Agradei no jantar a alguns amigos que me ajudaram nesta fase difícil.
 b. *Agradei a alguns amigos no jantar que me ajudaram nesta fase difícil.
- (20) a. Vi ontem a filha de um rapaz_i que_i joga no Benfica.
 b. *Vi a filha de um rapaz_i ontem que_i joga no Benfica.
- (21) a. Ontem explodiu uma bomba em Israel que causou 5 mortos.
 b. *Ontem uma bomba explodiu em Israel que causou 5 mortos.
- (22) a. Chegou um senhor ontem que fez muitas perguntas sobre ti.
 b. *Um senhor chegou ontem que fez muitas perguntas sobre ti.

Ainda no que diz respeito à extraposição a partir de outros antecedentes pré-verbais, verifica-se que o antecedente de uma oração restritiva extraposta pode ser um constituinte *wh-*, um foco contrastivo ou um constituinte enfático/avaliativo, mas não um tópico.

Quando analisados à luz de uma perspectiva comparativa, estes resultados tornam-se particularmente interessantes. Por um lado, as restrições acima apresentadas não se observam em fases anteriores do português. Por outro lado, línguas germânicas como o inglês e o holandês não exibem as restrições identificadas em português contemporâneo, apresentando antes a extraposição generalizada documentada na diacronia do português.

Do ponto de vista teórico, estes contrastes são explicados pelo facto de a extraposição de orações restritivas poder envolver diferentes estruturas: uma estrutura que envolve coordenação especificante (e elipse), como em (23), (cf. De Vries 2002) e uma estrutura que envolve abandono (*stranding*) da oração relativa na sua posição básica, como em (24) (cf. Kayne 1994).

- (23) ... [CoP [XP₁ antecedente YP] [Co [XP₂ [~~antecedente~~ OR] YP]]] (coord. especificante)

(24) ... [antecedente_i YP [t_i OR]] (abandono da OR)

As diferentes propriedades observadas resultam assim da estrutura sintáctica que gera as orações relativas restritivas extrapostas. Do ponto de vista diacrónico, propõe-se que a extraposição de restritivas é gerada por coordenação especificante em fases anteriores do português, sendo gerada por abandono da oração relativa em português europeu contemporâneo. Do ponto de vista inter-linguístico, coloca-se a hipótese de línguas como o inglês e o holandês gerarem a extraposição por coordenação especificante, em contraste com o português europeu contemporâneo (e possivelmente o italiano, o castelhano e o francês), que derivam a extraposição por abandono da oração relativa.

Sugere-se, por fim, que a variação encontrada neste domínio decorre da presença de um núcleo coordenativo restritivo abstracto no léxico de diferentes línguas (ou de diferentes estádios de uma mesma língua). Salienta-se, ainda, que parece haver uma correlação entre a perda de extraposição por coordenação especificante e a possibilidade de extrapor membros coordenados em contextos tradicionais de coordenação.

O terceiro e último estudo apresentado investiga um caso de (micro-)variação que envolve relativas apositivas introduzidas pelo morfema *o qual*. Do ponto de vista descritivo, existe um contraste marcado entre as propriedades das relativas apositivas com *o qual* na sincronia e diacronia do português. Este contraste pode ser descrito tendo em conta os seguintes parâmetros: (i) núcleo interno; (ii) extraposição; (iii) pied-piping; (iv) antecedentes oracionais; (v) antecedentes descontínuos; (vi) coordenação do morfema *wh* com outro sintagma nominal; (vii) força ilocutória; (viii) conjunção coordenativa.

Considere-se, a título de exemplo, as propriedades referidas em (i), (iv) e (vi). Em português europeu contemporâneo, as relativas apositivas introduzidas por *o qual* não podem ocorrer com um núcleo interno (cf. (25)), não permitem antecedentes oracionais (cf. (26)), nem a coordenação do morfema-*wh* com outro sintagma nominal (cf. (27)). Estas estruturas encontram-se, porém, documentadas na história do português, como se ilustra em (25)-(28).

- (25) *Existem argumentos fortes a favor dessa análise, os quais argumentos apresentarei de seguida.
- (26) *O João chegou a horas, o qual muito me surpreendeu.
- (27) *O presidente elogiou o João, o qual e a sua mulher têm desenvolvido um óptimo trabalho naquela instituição.
- (28) entrego e outorgo. ao Mosteiro de san Saluador de Moreyra. hũu casal que e en Rial de Pereyra. o qual casal a dita dona Mayor uëegas [...] mandou ao dito Mosteiro.
(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; 1282)
- (29) e se obrygou de pagar os dytos duzemos Reaes e dous fframguãos e a dyta galinha de fforo despoys do ffaleçimemto da dyta molher do dito alluaro fernandez em cada hũu Ano pelo dito dia de natall pera o qual loguo obrygou seus bães
(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; 1540)
- (30) filho de hum seu filho chamado per nome dom Henrrique, o qual era lidimo e, segundo conta a cronica, era o primeiro filho que o dito rei de Ungria ouve. O qual dom Henrrique e hum seu tio, irmão de sua madre,[...] se vierão a Castela aa corte, donde o dito rei dom Affonso estava
(CdP; *Cronica de Portugal*; 1419)

Tomando em consideração estes e outros contrastes, propõe-se que as orações apositivas introduzidas por *o qual* não envolvem apenas uma estrutura sintáctica, podendo ser geradas por coordenação especificante (cf. De Vries 2006b) ou por elevação do núcleo (cf. Kayne 1994). Assim, defende-se que em estádios anteriores do português as orações apositivas introduzidas por *o qual* são geradas por coordenação especificante, enquanto em português europeu contemporâneo são geradas por elevação do núcleo. Tal hipótese deriva as propriedades contrastivas acima mencionadas, bem como os contrastes observados entre diversas línguas contemporâneas (e.g., entre o português europeu contemporâneo e o inglês contemporâneo).

Por fim, é de salientar o facto de esta proposta de uma análise não unitária das orações apositivas assentar em evidência empírica sólida e bem controlada. Tomando como objecto uma construção muito específica – orações relativas apositivas introduzidas pelo morfema relativo *o qual* –, analisada em diferentes fases de uma mesma língua, este estudo envolveu um elevado grau de controlo de variáveis, muitas vezes inviável em estudos que comparam línguas histórica e tipologicamente distantes.

Com base nos dois últimos estudos, é possível concluir que a organização tradicional dos fenómenos linguísticos em torno dos tradicionais tipos de ‘construção’ (e.g., orações relativas apositivas e extraposição) pode facilmente induzir em erro. De facto, apesar de a classificação tradicional apontar para um comportamento uniforme das construções abrangidas sob determinada designação, nem sempre tal se verifica. Estes estudos demonstram que uma análise unitária das orações relativas apositivas e da extraposição de relativas restritivas não permite explicar a variação observada nos planos inter e intra-linguísticos.

A presente dissertação ilustra também, de forma bastante evidente, que os dados de variação linguística podem contribuir para a compreensão de diversos factos linguísticos e para a clarificação das relações que se estabelecem entre princípios e parâmetros.

Palavras-chave: *orações relativas, grupos nominais descontínuos, extraposição, orações relativas apositivas, sintaxe comparada e sintaxe diacrónica generativa*

Contents

Chapter 1. Introduction

1. Preamble	3
2. Goals of the dissertation	4
3. Methodology	5
3.1. Empirical evidence sources	5
3.1.1. Historical Portuguese	5
3.1.2. CEP	8
3.1.3. Other languages	9
3.2. Methodological options	10
4. Theoretical framework	11
4.1. Syntactic structure	11
4.2. Language change	13
4.3. Information structure	14
4.3.1. Focus	14
4.3.2. Topic	20
4.3.3. Topicalization/focalization	23
5. Overview of the dissertation	25

Chapter 2. Remnant-internal relativization: new evidence for the raising analysis

1. Introduction	29
1.1. Goals	29
1.2. Chapter outline	32
2. The syntax of RRCs	33
2.1. The definition of <i>RRCs</i>	33
2.2. Competing analyses of RRCs	33
2.2.1. The adjunction analysis	34
2.2.2. The raising analysis	34
2.3. Arguments for the raising analysis	35
2.4. Summary	42
3. The syntax of split noun phrases	42

3.1. The definition of <i>split noun phrases</i>	42
3.2. Core properties of split noun phrases	42
3.3. Competing analyses	45
3.3.1. Movement analyses.....	45
3.3.2. Base-generation analyses	54
3.4. Summary	56
4. The syntax of remnant-internal relativization: first take	57
4.1. The definition of <i>remnant-internal relativization</i>	57
4.2. Empirical evidence	59
4.2.1. Split noun phrases in earlier stages of Portuguese.....	59
4.2.2. Remnant-internal relativization in earlier stages of Portuguese	63
4.3. Competing analyses	66
4.4. Problems with the <i>raising analysis of RRCs + movement analysis of split noun phrases</i> combination	71
4.5. Summary	73
5. The proposal	73
5.1. Preliminaries	74
5.1.1. The interpretation of remnant-internal relativization.....	74
5.1.2. The copy theory of movement in the PF side	77
5.1.2.1 Bošković and Nunes (2007).....	77
5.1.2.2 Stjepanović (2007).....	78
5.2. The derivation of remnant-internal relativization	79
5.2.1. Pattern I.....	80
5.2.2. Pattern II	82
5.2.3. Pattern III	88
5.3. Remnant-internal relativization in CEP?	90
5.4. Summary	91
6. Conclusion	93

Chapter 3. Variation and Change in the Syntax of Restrictive Relative Clause

Extrapolation

1. Introduction.....	99
1.1. Goals	100
1.2. Chapter outline.....	101
2. The syntax of RRC extrapolation.....	102
2.1. Competing analyses	102
2.1.1. Rightward adjunction analyses	102
2.1.2. Stranding analyses	104
2.1.3. Coordination analyses.....	105
2.2. Unitary vs. non-unitary approaches to extrapolation	106
3. RRC extrapolation in Portuguese: previous scholarship.....	108
4. Properties of RRC extrapolation in contrast	112
4.1. Properties of RRC extrapolation in CEP.....	112
4.1.1. Summary.....	122
4.2. Properties of RRC extrapolation from a cross-linguistic perspective.....	123
4.2.1. Summary.....	128
4.3. Properties of RRC extrapolation in earlier stages of Portuguese.....	129
4.3.1. Summary.....	137
5. The analysis of RRC extrapolation in CEP.....	138
5.1. The stranding analysis	138
5.2. RRC extrapolation and movement to the left periphery	140
5.3. RRC extrapolation and short scrambling.....	142
5.3.1. Subject and object scrambling in CEP.....	142
5.3.2. The involvement of short scrambling in RRC extrapolation	146
5.3.3. The trigger for short scrambling in RRC extrapolation	148
5.3.3.1 Short scrambling and information structure in CEP.....	148
5.3.3.2 Short scrambling and information structure in RRC extrapolation.....	152
5.3.4. Deriving the intervening material	156
5.3.4.1 Deriving the occurrence of adverbs and PPs in the intervening position.....	157
5.3.4.2 Blocking the occurrence of other constituents in the intervening position	169
5.3.5. Summary	174
5.4. Deriving the properties of RRC extrapolation in CEP.....	175

5.4.1. Summary	181
5.5. Problems with the stranding analysis.....	182
5.5.1. Summary	192
5.6. Conclusion	193
6. The analysis of RRC extraposition in earlier stages of Portuguese.....	193
6.1. The specifying coordination analysis	194
6.1.1. Koster (2000)	194
6.1.2. De Vries (2002)	200
6.2. Deriving the properties of RRC extraposition in earlier stages of Port.	204
6.2.1. Summary	208
6.3. Problems with the specifying coordination plus ellipsis analysis.....	209
6.4. Conclusion	214
7. Comparative remarks	215
7.1. Diachronic change in RRC extraposition	216
7.2. Cross-linguistic contrasts	222
8. Conclusion	224

Chapter 4. Variation and Change in the Syntax of Appositive Relative Clauses

1. Introduction.....	229
1.1. Goals	229
1.2. Chapter outline.....	230
2. The syntax of ARCs.....	231
2.1. The competing analyses.....	231
2.1.1. Orphanage analyses	231
2.1.2. Constituency analyses.....	233
2.2. Unitary vs. non-unitary analyses	235
2.2.1. Cinque (1982)	236
2.2.2. Smits (1988).....	240
2.2.3. Bianchi (1999)	241
2.2.4. Cinque (2008)	243
3. Towards a dual approach to ARCs: evidence from Portuguese	246
3.1. <i>O qual</i> -ARC: an overview of its basic use and properties.....	247
3.2. Two types of <i>o qual</i> -ARCs: diachronic evidence	254

3.2.1. Some properties in contrast.....	255
3.2.1.1 Internal head.....	255
3.2.1.2 Extraposition.....	267
3.2.1.3 Pied-piping.....	286
3.2.2. Additional properties in contrast.....	303
3.2.2.1 Clausal antecedent.....	303
3.2.2.2 Split antecedents.....	307
3.2.2.3 Coordination of the <i>wh</i> -pronoun with another DP.....	310
3.2.2.4 Illocutionary force.....	312
3.2.2.5 Coordinator.....	315
3.3. Summary.....	317
4. An analysis of the two types of ARCs.....	319
4.1. <i>o qual</i> -ARCs in earlier stages of Portuguese.....	319
4.1.1. The specifying coordination analysis.....	320
4.1.2. The expansion of the specifying coordination analysis.....	325
4.1.3. Summary.....	334
4.2. <i>O qual</i> -ARC in CEP.....	335
4.2.1. The raising analysis.....	335
5. Deriving the contrasting properties of the two types of <i>o qual</i> -ARCs.....	338
5.1. Internal head.....	338
5.2. Extraposition.....	340
5.3. Pied-piping.....	353
5.4. Clausal antecedent.....	361
5.5. Split antecedents.....	361
5.6. Coordination of the <i>wh</i> -pronoun with another DP.....	364
5.7. Illocutionary force.....	365
5.8. Coordinator.....	369
5.9. Summary.....	369
6. Some comparative remarks.....	370
6.1. Diachronic change in the syntax of <i>o qual</i> -ARCs.....	370
6.2. Synchronic variation in the syntax of <i>o qual</i> -ARCs.....	377
6.3. Cross-linguistic contrasts.....	383
7. Conclusion.....	385

Chapter 5. Conclusion 389

References 399

Chapter 1

Introduction

Contents

1. Preamble	3
2. Goals of the dissertation	4
3. Methodology	5
3.1. Empirical evidence sources	5
3.1.1. Historical Portuguese	5
3.1.2. CEP	8
3.1.3. Other languages	9
3.2. Methodological options	10
4. Theoretical framework.....	11
4.1. Syntactic structure.....	11
4.2. Language change	13
4.3. Information structure	14
4.3.1. Focus.....	14
4.3.2. Topic	20
4.3.3. Topicalization/focalization	23
5. Overview of the dissertation	25

1. Preamble

The concept of *relative clause* is difficult to characterize, given the diversity of constructions traditionally grouped under this label. As a working definition, I will adopt De Vries's (2002: 14) proposal, which defines relative clauses as having the properties in (1) (Note: the term *pivot* refers to a constituent that is semantically shared by the matrix clause and the relative clause).

(1) *Defining properties of relative constructions:*

- a. A relative clause is subordinated.
- b. A relative clause is connected to surrounding material by a pivot constituent.

Relative clauses pose many challenges to syntactic theory. Chomsky (1995) explicitly acknowledges this fact by noting that “(...) we still have no good phrase structure theory for such simple matters as attributive adjectives, relative clauses, and adjuncts of many different types” (Chomsky 1995: 382, n. 22). Bianchi (2002: 242), in turn, regards relative clauses as an intriguing empirical domain “(...) both because of the complexity of the data and of the theoretical relevance of the construction (especially with respect to the syntax-semantic interface).”

To further complicate the matter, relative clauses allow for other syntactic structures that, despite all efforts, remain poorly understood (e.g., pied-piping and extraposition). Other difficulties can be identified that are not exclusive to analyses of relative clauses but are frequently encountered in generative syntax. Such difficulties include an overemphasis on the study of English, the predominance of theoretically oriented studies (as opposed to corpus-based and corpus-driven studies), and incomplete coverage of relevant properties. The last two limitations are perfectly expressed by Bianchi (2002: 242):

Each analysis is designed to account for certain aspects of a domain, and leaves others unaccounted for. But the choice of the “core” data to be analysed is to some extent arbitrary, for we cannot know a priori which set of data is fully representative of the empirical domain under investigation; and, as I have tried to illustrate above, the way an analysis is designed is crucially affected by more general theoretical concerns.

From these considerations it follows that many gaps remain in our understanding of relative clauses. This is not to say that little has been written about this topic; on contrary, relative clauses have become a subject of extensive debate in the generative field (e.g., the Chomsky/Bresnan debate on unbounded dependencies and the Bianchi/Borsley debate on the antisymmetric/symmetric analyses of relative clauses) (see Bianchi 2002). However, most studies focus primarily on a specific contemporary language, neglecting the cross-linguistic and diachronic variation. This dissertation aims to fill some of these gaps by incorporating new empirical data that shed light on the syntax of relative clauses. The data considered are mainly drawn from earlier stages of Portuguese and Contemporary European Portuguese (CEP), but other languages are considered, including Latin, French, Italian, Spanish, English, Dutch, German, and Russian.

2. Goals of the dissertation

This dissertation sheds light on language variation and change from a generative syntactic perspective, based on a case study of relative clause constructions in Portuguese and other languages. It demonstrates that inter and intra-linguistic contrasts offer precious means of enhancing our understanding of various linguistic phenomena, thereby contributing also to our understanding of the interaction between principles and parameters.

This dissertation has the following specific aims.

- Provide a comprehensive description of three distinct phenomena documented in the synchrony and diachrony of Portuguese: remnant-internal relativization, restrictive relative clause (RRC) extraposition, and appositive relative clauses (ARCs).
- Compare and contrast the properties of several languages (with respect to the three topics under analysis) with earlier stages of Portuguese and CEP.
- Explain the variation facts in the light of generative syntax and the theories that combine language change with language acquisition.

- Contribute to a better understanding of the syntax of earlier stages of Portuguese and CEP regarding not only relative constructions but also other topics such as word order, information structure, split noun phrases, and pied-piping.

3. Methodology

The present research adopts a comparative perspective, contrasting the behavior of different languages and different stages of the same language with respect to three syntactic phenomena: remnant-internal relativization, RRC extraposition and ARCs.

Such an approach requires the collection and analysis of data from different languages and periods, namely, from historical Portuguese, CEP, and other languages. The empirical evidence sources and the methodology used are presented in Sections 3.1 and 3.2.

3.1. Empirical evidence sources

3.1.1. Historical Portuguese

At present, three important digital corpora are available for the study of historical Portuguese:

- *Tycho Brahe Parsed Corpus of Historical Portuguese* (TYC), available at <http://www.tycho.iel.unicamp.br/~tycho/corpus/en/index.html>. It contains 52 texts written by authors born between 1380 and 1845, amounting to a corpus size of 2,407,422 words¹ (see Galves, Coord.).

¹ Information available online on October 25, 2010.

- *Digital Corpus of Medieval Portuguese* (CIPM – *Corpus Informatizado do Português Medieval*), available at: <http://cipm.fsh.unl.pt/>. It consists of texts dating from the 12th to the 16th century, amounting to a corpus size of 2,053,582 words² (see Xavier, Coord.).
- *Unknown Letters* (CARDS - *Cartas Desconhecidas*), available at <http://alfclul.clul.ul.pt/cards-fly>. By the end of 2010, 2,000 private letters from the 15th to the 19th century will be available online (at the time of this writing, only 400 were available)³ (see Marquilhas, Coord.).

Although these corpora have made important contributions to the study of earlier stages of Portuguese, they have non-trivial limitations for the development of syntactic studies. The most serious limitation concerns the lack of syntactically annotated texts. With the exception of TYC (which contains eleven syntactically annotated texts),⁴ no other texts are available for syntactic searches. This means that these algorithms permit users to search for words or sequences of words but not for syntactic structures.

This lack of syntactically annotated texts is a serious obstacle to the development of studies on the syntax of historical Portuguese because the time limit imposed by a master's or PhD research project is hardly compatible with a manual search for specific syntactic structures.

Given this scenario, the following corpora were selected for this study. For the period between the 13th century and the second half of the 16th century, the majority of the data came from the texts edited by Martins (2001) in *Documentos Portugueses do Noroeste e da Região de Lisboa*.⁵ This publication contains 218 notarial documents, produced mostly between the second half of the 13th century and the second half of the 16th century and amounting to a corpus size of approximately

² Information available online on October 25, 2010.

³ These figures correspond to the texts available online on October 25, 2010.

⁴ Information available online on October 25, 2010.

⁵ This corpus is available online in the *Digital Corpus of Medieval Portuguese* (CIPM; *Corpus Informatizado do Português Medieval*) at <http://cipm.fsh.unl.pt/>.

140,000 words. The empirical evidence for this period is complemented by other sources:

- Texts available in CIPM.
- Texts available in TYC.
- Individual editions: *Crónica do Conde D. Pedro de Meneses* (ed. Brocardo 1997); *Livro de Linhagens do Conde D. Pedro* (ed. Brocardo 2006); *Demanda do Santo Graal* (ed. Piel and Nunes 1988); *Livro dos Ofícios* (ed. Piel 1948); *Crónica de D. Fernando* (ed. Macchi 1975), *Gil Vicente: todas as obras* (coord. Camões 1999).

From the second half of the 16th century to the 19th century, data are drawn from:

- TYC.
- CARDS.
- *Corpus do Português* (CdP), available at: <http://www.corpusdoportugues.org>. This corpus currently contains almost 57,000 texts written between the 14th and the 20th centuries, amounting to a corpus size of almost 45 million words⁶ (see Davies and Ferreira 2006-).
- Individual text editions: *Os Autos do Processo de Vieira na Inquisição* (ed. Muhana 1995); *Documentos para a História da Inquisição em Portugal* (ed. Pereira 1987); *Inquisição de Évora: dos Primórdios a 1668* (ed. Coelho 1987).
- Grammars and studies in the history of Portuguese language.

Examples from historical Portuguese are identified by the corpus (e.g., TYC), edition (e.g., Martins 2001) or study/grammar (e.g., Said Ali 1931/1971) and by basic information concerning the author, title, and date (year or century, according to the information available).

For ease of reading, the editorial conventions adopted by the different editors in the transcription of texts have been simplified, i.e.: (i) parentheses and italics that

⁶ Information available online on October 25, 2010.

indicate the expansion of abbreviations have been eliminated; (ii) the indication of line breaks and the *hyphen* sign (which indicates the division of a word at the end of a line) have been removed; and (iii) the tildes and the superscript marks (similar to an acute accent) that editors transcribe after a letter appear here above the letter.

Examples from historical (and contemporary) Portuguese are followed by glosses and the translation into English. All glosses and translations are mine. The same holds for glosses and translations of examples drawn from grammars and studies written in Portuguese.

3.1.2. CEP

The empirical evidence from CEP was based on my own linguistic intuitions and on intuitions from other speakers. As a complement, the following sources were used:

- Grammars and studies on the syntax of CEP.
- Written texts available in the internet.
- Written and spoken texts, available in the following corpora:
 - *Corpus de Extractos de Textos Electrónicos MCT/Público* (CETEMPúblico), available at <http://www.linguateca.pt/CETEMPUBLICO/>. This corpus contains some 190 million words⁷ from a daily Portuguese newspaper (see Rocha and Santos 2000).
 - *Corpus de Referência do Português Contemporâneo* (CRPC), available at http://www.clul.ul.pt/index.php?option=com_content&view=article&id=183%3Areferencia-corpus-of-contemporary-portuguese-crpc&catid=91%3Aanagrama&lang=pt. This corpus contains 334 million words taken from samples of several types of written and spoken texts (see Bacelar do Nascimento, Coord.).
 - *Corpus of Spoken Portuguese* available in *C-ORAL-ROM - Integrated Reference Corpora for Spoken Romance Languages* (C-ORAL-ROM). The Portuguese corpus contains 320,452 words (see Bacelar do Nascimento, Portuguese Coord. 2005).

⁷ Information available online on November 4, 2010.

- *Syntax-oriented Corpus of Portuguese Dialects* (CORDIAL-SIN), available at <http://www.clul.ul.pt/en/research-teams/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects>. This corpus contains 600,000 words; it is a geographically representative body of selected excerpts of spontaneous and semi-directed speech (see Martins, Coord.).

Examples drawn from corpora are identified by the corpus (e.g., CRPC) and by basic information regarding, for instance, the version/file (e.g., CETEMPúblico 1.7 v. 4.0). Examples from the internet are identified by the worldwide web addresses; grammars and studies are referenced as usual (e.g., Brito and Duarte 2003).

3.1.3. Other languages

Whenever possible, a cross-linguistic perspective was adopted to show how different languages (or different stages of the same language) behave with respect to the phenomena under analysis. In this case, contemporary data were mainly drawn from:

- Grammars and several studies on the syntax of relativization.
- *Corpus of Contemporary American English* (COCA), available at: <http://www.americancorpus.org>. This corpus currently contains 410 million written and spoken words⁸ (see Davies 2008-).

Some data from historical English are also presented, taken from:

- *Penn Parsed Corpus of Modern British English* (PPCMBE), available at: <http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html>. It consists of just under one million words and spans roughly 210 years (1700-1914)⁹ (see Kroch et al. 2010).

⁸ Information available online on November 4, 2010.

⁹ Information available online on November 4, 2010.

- *Penn-Helsinki Parsed Corpus of Early Modern English* (PPCEME), available at <http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html>. It consists of over 1.7 million words and spans 210 years (1500-1710)¹⁰ (see Kroch et al. 2004).

Examples drawn from corpora are identified by the corpus (e.g., PPCEME) and by basic information regarding, for instance, the version/file. Examples from grammars and studies are referenced as usual (e.g., De Vries 2002).

Non-English examples are always followed by the translation into English. Glosses are given when the author provides them. Exceptions are Latin, Dutch and German, for which I provide the glosses and translation in cases for which the authors do not provide them.

3.2. Methodological options

Having presented the empirical basis of this dissertation, some remarks about the methodology are in order.

In the study of earlier stages of Portuguese, given the limitations of the resources available, a small corpus of studies was selected for systematic syntactic analysis. In this case, a corpus-driven methodology was conducted that, having no pre-defined search structures, allows all types of constructions to emerge from the corpora (see Tognini-Bonelli 2001).

Following this procedure, the texts edited by Martins (2001) were submitted to a systematic study. The relative constructions were extracted manually and stored in a database. This database has approximately 4000 tokens and contains information about the syntactic and semantic properties of relative constructions. The adoption of this methodology was rewarding; a variety of constructions (or properties of the constructions) were found that have not been reported in the grammars and studies of the history of Portuguese.

Based on the collected data, the phenomena to be studied were selected according to the following principles: (i) the contrasting properties of the relevant

¹⁰ Information available online on November 4, 2010.

structures in earlier stages of Portuguese with respect to CEP; (ii) the novelty of the constructions (i.e., constructions/properties not yet reported in the literature), and (iii) the theoretical relevance of the construction.

Once the study topics were selected, other sources of historical Portuguese were inspected to: (i) document phenomena unattested in Martins (2001); (ii) broaden the variety of text-types documenting a specific phenomenon; and (iii) cover the period from the second half of the 16th century to the 19th century. In this phase, a corpus-based methodology was followed, which involves the selection of particular examples for specific and pre-determined purposes (see Tognini-Bonelli 2001). The same procedure was adopted for the corpora data from CEP and other languages.

The research is framed within a qualitative approach (as opposed to quantitative approach) that involved the search for particular syntactic constructions, independent of their frequency in the particular corpus being analyzed. Frequencies were only available for Martins' (2001) corpus.

4. Theoretical framework

The theoretical approaches used throughout this dissertation are briefly presented below. Emphasis is given only to the aspects that directly concern the subject of this dissertation.

4.1. Syntactic structure

The syntactic analysis is developed within the Principles-and-Parameters (P&P) framework (Chomsky 1981) under its minimalist version (Chomsky 1993, 1995 and subsequent work).

The P&P model constitutes a “(...) radical break from the rich tradition of thousands of years of linguistic inquiry (...)” (Chomsky 1995: 5). Perhaps one its most relevant contributions has been the rejection of the view that a language consists of rules for forming grammatical constructions (e.g., relative clauses and passives). As Chomsky (1995: 5-6) states:

The P&P held that languages have no rules in anything like the familiar sense, and no theoretically significant grammatical constructions except as taxonomic artifacts. There are universal principles and finite array of options as to how they apply (parameters), but no language-particular rules and no grammatical constructions of the traditional sort within or across languages.

This move was crucial for the development of generative syntax. It asserts that the notion of *construction*,¹¹ which was used in the traditional grammar (and in earlier periods of the generative grammar) to refer to clause types (among other syntactic patterns), can be used non-technically to refer to a variety of apparently related constructions but has no theoretical relevance.

In view of the new paradigm, the concept of *relative clause* is used in this dissertation as a mere descriptive label with no explanatory force. Similarly, the view that relative clauses (and other constructions, such as extraposition) involve uniform underlying structure and movement is rejected. In actuality, this ideal, which is still pursued in many generative studies, can somehow be taken as a revival of the traditional concept of construction.

Given its relevance to the present dissertation, The Antisymmetry Theory, proposed by Kayne (1994), should also be addressed. This theory claims that word order reflects structural hierarchy by means of the linear correspondence axiom (LCA), which states that asymmetric c-command imposes a linear ordering of terminal elements. The LCA imposes severe restrictions on the syntactic structure, in particular, (i) the impossibility of right-hand adjunction; (ii) the impossibility of rightward movement; (iii) strict binary branching; and (iv) the Specifier-head-complement universal order.

Many studies have been devoted to arguments for and against this theory. As Takano (2003: 516) observes:

¹¹ See Schönefeld (2006), where the notion of construction is examined from a number of different theoretical perspectives. In this respect, it is worth mentioning that there are many theories that diverge radically from the generative view presented above. This is, for instance, the case of the Construction Grammar model, pursued by a growing number of researchers (e.g., Ivan Sag, Charles Fillmore, and William Croft).

(...) for a number of phenomena involving rightward positioning, such as rightward adjuncts, heavy NP shift, extraposition, postverbal subjects, and postverbal constituents in OV languages, there now exist both an approach consistent with Kayne's theory (the antisymmetric approach) and another not consistent with it (the symmetric approach).

Kayne's theory has also energized the debate over relative constructions. As a consequence, and following the pattern referenced by Takano (2003), there now exists both an approach consistent with Kayne's theory (the raising analysis of relative clauses, originally proposed by Schachter 1973 and Vergnaud 1974, 1985, and more recently revived by Kayne 1994, Bianchi 1999, and De Vries 2002) and another that is not consistent with it (the right-hand adjunction analysis, originally proposed by Ross 1967, Chomsky 1977, and Jackendoff 1977).

4.2. Language change

The interpretation and explanation of grammatical changes is developed within the model proposed by Lightfoot (see Lightfoot 1991, 1999 and subsequent work), which associates diachronic change with language acquisition.

Within this model, grammars are regarded as mental organs (represented in the mind of the speaker) and not as social entities (codifying the data presented in a particular period). Following the Chomskian view of language acquisition, it is assumed that children are born with a universal grammar (UG) (a set of linguistic principles common to all languages) and that, when exposed to primary linguistic data (crucial experiences, what the children hear), they develop a specific grammar.

Grammatical changes occur when learners converge on a grammatical system different from the one internalized by the speakers of the previous generation. Lightfoot's model can therefore be seen as a synchronic approach to language change, according to which changes have local causes, and are not driven by diachronic generalizations about language change.

The analysis proposed here also builds on the insights of Kroch (1989, 1994, 2001) and on the contributions of parametric syntax (cf., e.g., Holmberg and Roberts 2010).

4.3. Information structure

Information structure can be regarded as a phenomenon of information packaging that responds to the immediate communicative needs of interlocutors (see Chafe 1976, cited in Krifka 2007).

There is a broad agreement that word order is, at least to some extent, determined by information structure. It is also generally assumed that information structure interacts not only with syntax but also with other grammatical domains, such as interpretation, intonation, and morphology (Erteschik-Shir 2007).

Because the present dissertation deals with several word order patterns in CEP and in earlier stages of Portuguese, some clarifications of the terminology used in information structure approaches are necessary.

According to Lambrecht (1994: 5-6), information structure involves the analysis of four major categories: propositional information, identifiability and activation, topic and focus. In the present research, only two of these categories are explicitly addressed: topic and focus. Since the terminology associated with these two concepts is notoriously varied, Sections 4.3.1 and 4.3.2 clarify the use of these terms, introducing some theoretical details that are crucial for the development of this dissertation. At the end, Section 4.3.3 clarifies the use of the terms *topicalization* and *focalization*.

4.3.1. Focus

Focus is a term used to refer to a linguistic expression that conveys new or textually and situationally non-derivable information (Halliday 1967). Although it can be subclassified in different ways, in this dissertation only two main distinctions are used: (i) broad focus/narrow focus and (ii) information focus/emphatic-contrastive focus.

The distinction between broad and narrow focus is based on the scope of focus. Broad information focus (also known as sentential focus or unmarked focus) is used to refer to contexts in which the focus is assigned to the whole sentence. Narrow information focus refers to contexts in which only part of the sentence is assigned focus.

Information focus is commonly distinguished from contrastive focus. Information focus (also known as semantic focus or presentational focus) signals distinctions between shared and new information (Enkvist 1980). It represents new information related to what has been called a topic, presupposition, background, or common ground. A typical test used to identify information focus is a question-answer pair, where the focused constituent of the answer replaces the *wh*-word in the question (see, among others, Costa 1998, 2004a). For illustrations, see examples (2)-(3), from CEP:¹²

(2) *Sentence-focus (broad focus)*

A: a. **O que** é que aconteceu?
‘What happened?’

B: b. **O João partiu a janela.**
the J. broke a window
‘João broke a window.’

(Costa 2004a: 79)

(3) *Object focused (narrow focus)*

A: a. **O que** é que o Paulo partiu?
‘What did Paulo break?’

B: b. O Paulo partiu **a janela.**
the P. broke the window
‘Paulo broke the window.’

(Costa 2004a: 79)

Emphatic/contrastive focus (also known as identificational focus) is commonly defined as evoking a suitable set of alternatives from which a subset is chosen (see, e.g., Rooth 1992). Some authors also define it on the basis of semantic features, such as exhaustiveness (see, e.g., Kiss 1998). In this study and in Zimmermann (2007), emphatic/contrastive focus is taken as a discourse-pragmatic phenomenon related to “(...) the speaker’s assumptions about what the hearer considers to be likely or

¹² For the sake of clarity, in examples (2)-(3), the *wh*-word in the question (*o que* ‘what’) and the focused constituent of the answer are marked in bold.

unlikely, introducing a certain degree of subjectivity.” (Zimmermann 2007: 148). This definition has a broader scope because it includes not only the concepts of contrast and exhaustivity but also the more general concept of emphasis.

Languages may resort to different strategies of emphatic/contrastive focus marking, namely, intonation contour, syntactic movement, particular syntactic structures (e.g., clefts), focus-sensitive particles and morphological markers. The different ways of expressing emphatic/contrastive focus in CEP are displayed in I. to IV. below:¹³

I. Prosodic prominence alone

- (4) Partimos DIA VINTE de Abril.
leave:1PL day twenty of April
 ‘We leave on April 20 (and not on April 21).’

II. Focus-movement to the sentential left periphery (i.e., focalization)¹⁴

- (5) COM ESTAS PALAVRAS me despeço.
with these words me:CL say.goodbye:1sg
 ‘I say goodbye WITH THESE WORDS.’

III. Specific syntactic constructions (e.g., clefts)

- (6) É COM ESTAS PALAVRAS que me despeço.
is with these words that me:CL say.goodbye:1SG
 ‘I say goodbye WITH THESE WORDS.’

¹³ The emphatic/contrastive focus (identified by the E/CSR rule) is indicated by small caps.

¹⁴ In line with Zubizarreta (1999) and Martins (in prep.), the preposed constituent in this construction (i.e., *com estas palavras* ‘with these words’ in (5)) is referred to as *preposed focus*. For the use of the term *focalization*, see Section 4.3.3 below.

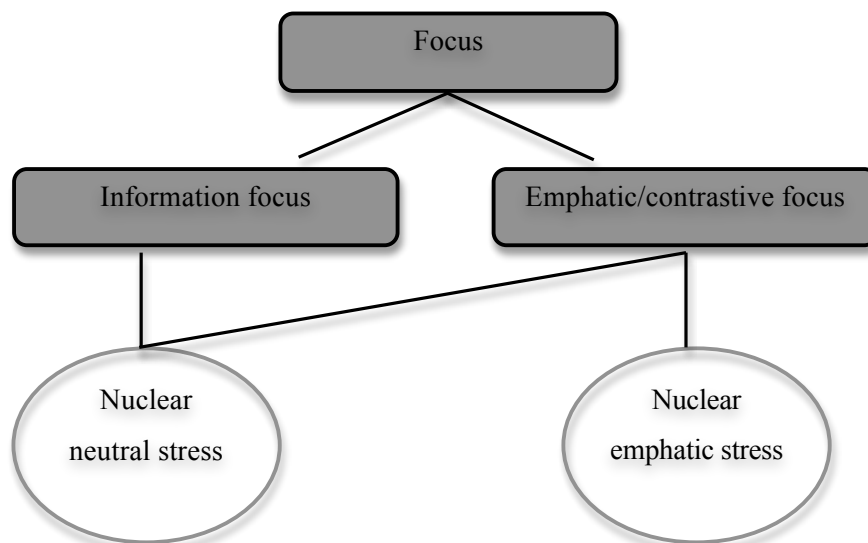
IV. Focus-sensitive particles (e.g., *só* ‘only’)

- (7) *SÓ OS MEUS AMIGOS percebem o que quero dizer.*
only the my friends understand the that want:1SG say:INF
 ‘Only my friends understand what I mean.’

From the above patterns, it follows that the emphatic/contrastive focus in CEP can be marked in the following ways: only prosodically (I.); syntactically (II.-III.); or lexically (IV.). It is also worth noting that in CEP, both syntactic and lexical strategies co-occur with prosodic marking.

Several authors (among others, Costa 1998, 2004a, and Frota 1998, 2002) have proposed that prosody plays an important role in the identification of focus in CEP. This dissertation adopts Zubizarreta’s (1998, 1999) view of the relationship between prosodic prominence and focus, which can be schematically represented as follows:

Figure 1. The relationship between focus and prosody (Zubizarreta 1998, 1999)

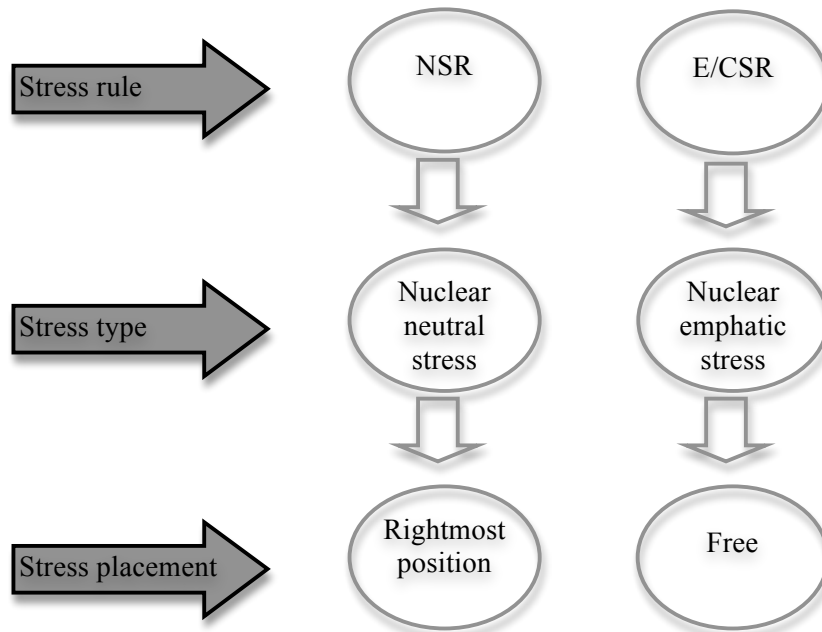


Starting with a slightly impressionistic generalization, the crucial factor seems to be that a focused constituent always carries the nuclear stress. However, there is no one-to-one correlation between the subtypes of focus and the subtypes of stress: information focus is identified by neutral stress, whereas emphatic/contrastive focus

may be identified by either emphatic or neutral stress (see Zubizarreta 1999: 4229, fn. 16 and 4242: fn. 27).¹⁵

Neutral and emphatic stresses are, in turn, assigned by different rules. The effects that these rules have on stress placement are outlined in Figure 2:

Figure 2. Effects of stress rules on stress placement



The Nuclear Stress Rule (NSR) (see (8)) assigns prominence to the rightmost/lowest sentential constituent, whereas the Emphatic/Contrastive Rule (E/CSR) (see (9)) assigns prominence to any position.

¹⁵ As will become clear below, in CEP the ambiguity between information focus and emphatic/contrastive focus can only arise if the focused constituent is rightmost. Following Frota (1998, 2002 and other related work), I propose that the aforementioned ambiguity is eliminated by differences in peak alignment (or choice of pitch accent) (see Chapter 2, Section 5.2 for further details).

(8) *Nuclear Stress Rule (CEP)*¹⁶

Given two sisters C_i and C_j , the one lower in the asymmetric c-command ordering is more prominent.

(Zubizarreta 1998: 19)

(9) *Emphatic/Contrastive Rule*

A word with contrastive stress must be dominated by every F[ocus]-marked constituent in the phrase.

(Zubizarreta 1998: 45)

In addition to these rules, Zubizarreta (1998) proposes the Focus Prominence Rule (FPR) (see (10)), which aims to capture the relationship between the focus-structure of a sentence and its prosody. The idea is that the focused elements in a sentence are marked with a feature [+F], whereas the presupposed ones are marked with a feature [-F]. The FPR dictates that the nuclear stress must target an [+F]-element.

(10) *Focus Prominence Rule*

Given two sister nodes C_i (marked [+F]) and C_j (marked [-F]), C_i is more prominent than C_j .

(Zubizarreta 1998: 21)

The coexistence of the FPR and NSR in the grammar does produce any conflict: because the E/CSR assigns emphatic/contrastive stress to any element, no principle prevents an [+F]-element from receiving it.

Some conflicts may arise, however, between the FPR and the NSR. The FPR may force the stress on a non-final position, whereas the NSR requires the stress to fall on the rightmost clausal position. According to Zubizarreta (1998), languages seem to vary in the way they solve this conflict. In languages such as English and French, defocalized [-F] elements are treated as extrametrical in the sense that they are skipped by the NSR. Other languages (such as Spanish and Italian) employ prosodically motivated movement (p-movement), which moves the post-focal [-F]

¹⁶ Actually, this rule consists of two parts: one sensitive to asymmetric c-command relations (C-NSR) and the other sensitive to selectional ordering (S-NSR). Languages differ in the way in which the NSR applies; both C-NSR and S-NSR are active in German and English, but only C-NSR is active in languages such as Spanish and Italian (and Portuguese). Here, I consider only the part of the rule that is relevant to CEP.

elements out of the rightmost clausal position. Then, the NSR applies and puts the stress on the sentence-final position.

4.3.2. Topic

Most authors define the concept of topic in relation to the concept of comment; the topic is what the sentence is about, whereas the comment is what is said about the topic. However, there seems to be no consensus in the literature as to what topic really is. According to Kuroda (2005), three main views on this concept can be found in the literature:

- Topic can be taken as a syntactic concept, referring to a constituent that is placed at the sentential left periphery.
- Topic can be taken as a discourse-theoretical concept, referring to a constituent that expresses old information in the organization of the discourse.
- Topic can be taken as a semantic concept, referring to a constituent that expresses an aboutness relation; it can be familiar or recognizable or presupposed or part of the common ground, but need not be old information.

These views on topic do not correspond to actual theories, but rather to different dimensions of the concept. What usually happens is that linguists have a preferred dimension that they focus on, without denying the relevance of the other dimensions (see Henk 2010).

In the literature, a distinction is also made between a marked topic and a non-marked topic (see Duarte 1987, 1997 and subsequent work, among others). In subject-prominent languages like CEP, a topic is non-marked if it has the grammatical function of the subject (as in (11)). A topic is marked if it does not have the grammatical function of the subject (as in (12)), where the topic is the complement of the verb).

(11) A: a. Porque é que estás tão bem-disposto?
 ‘Why are you in such a good mood?’

B: b. A Cristina já chegou.
the C. already arrived
 ‘Cristina has arrived already.’

(12) Na Cristina, eu nunca mais confio.
in.the C. I never more trust:1SG
 ‘I will never trust Cristina again.’

In the present dissertation, a semantic definition of topic is adopted as a way of rejecting the traditional idea that topic expresses old information (see, among others, Reinhart 1982, Kuroda 2005, and Krifka 2007). Evidence in support of rejecting this idea comes, for instance, from sentences such as (11). In this sentence, the subject *A Cristina* is contained in the focus domain (because the whole sentence is assigned broad information focus). Nevertheless, *A Cristina* is interpreted as the topic of the sentence because it expresses an aboutness relation (i.e., what the sentence is about).¹⁷ In this sense, a topic must be familiar, recognizable, or presupposed, but does not need to be old information.

Another tendency in the literature has been to emphasize the logical function of topic (see, among others, Duarte 1987, 1997, Kuroda 2005, and Martins 1994, in prep.). According to this view, a topic is taken to foreground an individual or class as the subject of the predication, occurring in sentences that express categorical/predicational judgments.¹⁸ A sentence without a topic expresses athetic/descriptive judgment.

¹⁷ In characterizing discourse contexts similar to (11), Gécseg (2006, cited in Henk 2010: 7) proposes that the sentence (in (11)b, in this case) pragmatically asserts something about the speaker by means of a proposition which logically or semantically asserts something about Mary.

¹⁸ In what might be called a theory of judgments, a judgment can be considered a cognitive act expressed by the utterance of a sentence (see Kuroda 1992: 20). There are two types of judgment: categorical judgment and thetic judgment (as proposed by F. Brentano and A. Marty). Following Kuroda (2005), a categorical judgment can be defined as “(...) a cognitive act of attributing a predicate to a subject, a predication of the form conforming to the classical Aristotelian logic” (Kuroda 2005: 25). In turn, a thetic judgment is grounded on perception: “A thetic judgment is a representation of a perceptually apprehended real, recalled, or imagined situation.” (Kuroda 2005: 29-30).

The logical view on topics is of particular relevance to the present study, as it seems to play an important role in explaining some word-order facts in CEP. For instance, broad information-focus sentences with a post-verbal subject always express athetic/descriptive judgment (i.e., the subject is always non-topic):

- (13) *Chegou o João.*
arrived the J.
 ‘João arrived.’

Broad information-focus sentences with a pre-verbal subject are ambiguous (see Martins, in prep.), as they can express either a categorical or athetic judgment.¹⁹ This is illustrated in (14), which is taken from Martins (in prep.). The idea underlying this analysis is that when stage-level predicates (such as *chegar* ‘arrived’) are involved, the ‘apprehension’ of a situation is rooted in visual, auditory, or sensorial perception. Therefore, athetic/descriptive judgment is available only if the speaker is able to perceptually observe the situation.

- (14) *O teu pai já chegou.*
the your father already arrived
 ‘Your father has arrived.’
 ✓ Reading 1: (*thetic*) Your father has arrived already (here, at us, for dinner).
 ✓ Reading 2: (*categorical*) Your father has arrived already (back at his home).

As Readings 1 and 2 show, a broad information-focus sentence involving a pre-verbal subject may express two different types of judgment: a categorical judgment, if the ‘arrive’ stays outside the visual, auditory, or sensorial reach of the speaker (Reading 2) or athetic judgment, if the speaker perceptually observes the situation (Reading 1).

Therefore, it is possible to conclude that the pre-verbal subject position in CEP ([Spec, IP])²⁰ is an ambiguous position: it can be filled by topic elements (i.e., the

¹⁹ Previous studies on CEP (e.g., Duarte 1997, Martins 1994) generally assumed that broad information-focus sentences with pre-verbal subjects always express categorical judgments.

²⁰ There are two competing proposals for the syntactic analysis of pre-verbal subjects in CEP: Costa (2001, 2004a) and Costa and Duarte (2002) claim that pre-verbal subjects A-move to [Spec, IP], whereas Barbosa (1995, 2009) claims that subjects are base-generated in a left-dislocated position
(to be continued)

subject of predication in sentences expressing categorical judgments), but it can also be filled by non-topic elements (i.e., the subject of a sentence expressingthetic/descriptive judgments). On the other hand, the post-verbal position of the subject is non-ambiguous, as it is occupied by non-topic elements (occurring in sentences that expressthetic judgments).

4.3.3. Topicalization/focalization

A last terminological note is in order regarding the use of the terms *topicalization* and *focalization*. The term topicalization has been traditionally used in the generative literature to refer to the movement of a constituent to the left periphery of the sentence. According to this view, topicalization occurs in sentences expressing a topic/comment articulation or a focus-presupposition articulation. For an illustration, see the contrast in (15), taken from Rizzi (1997: 285):

- (15) a. Your book, you should give t to Paul (not to Bill)] [topic-comment articulation]
 b. YOUR BOOK you should give t to Paul (not mine) [focus-presupposition articulation]

In the late 1990s, a terminological shift occurred that reflects the emergence of the cartographic analysis proposed by Rizzi (1997). Within this approach, a clear distinction is made between sentences such as (15)a and (15)b, because preposed topics and preposed foci are considered to occupy different positions in the split-CP. As a consequence, the term topicalization started to designate topic-comment structures alone, whereas the term focalization was used to refer to focus-presupposition structures.

This terminological shift also clarified the status of preposed constituents in Romance languages. Generally, Romance languages express the topic-comment

(as adjuncts to CP/IP). The two hypotheses are sketched in (i)a-(i)b, respectively. In this dissertation, it is assumed that pre-verbal subjects in CEP are in [Spec, IP] (as represented in (i)a).

- (i) a. [IP S V [VP t_S t_V]]
 b. [IP/CP S [IP/CP V [VP *pro* t_V]]

articulation with the construction that Cinque (1990) has called *clitic left dislocation*, involving a resumptive clitic coreferential to the topic. As Duarte (1987, 1997, and subsequent work) shows, CEP has a special behavior in this respect because the topic-comment articulation may also involve a topic that is syntactically connected with an empty category inside the comment (as in the English example in (15)a).²¹ Hence, in CEP, clitic left dislocation coexists with topicalization, as illustrated in (16)a and (16)b:

- (16) a. *Esse livro, ainda não li t.* (topicalization)
that book yet not read
 ‘I have not read that book yet.’
- b. *Esse livro, ainda não o li* (clitic left dislocation)
that book yet not it:CL read
 ‘I have not read that book yet.’

The focus-presupposition articulation is expressed in some Romance languages by a preposed contrastive focus in a construction called focalization (which was also called topicalization before the terminological split). Such a construction has been reported for Italian (cf. Cinque 1990), Spanish (cf. Zubizarreta 1999), Catalan (cf. Sòla 1992), and for CEP (cf. Ambar 1992, Martins 1994, Duarte 1997, and Martins, in prep.). A case in point is provided in (17) (repeated from (5) above).

- (17) *Com estas palavras me despeço.* (focalization)
with these words me:CL say.goodbye:1SG
 ‘I say goodbye WITH THESE WORDS.’

Interestingly, contemporary languages seem to feature a correlation between punctuation and the two constructions under consideration. In particular, the use of a comma after the preposed constituent usually acts as a signal of the topic-comment structure and, concomitantly, as an orthographic means of excluding the focus-presupposition reading.

²¹ For this reason, this construction is also referred to as (English-type) topicalization.

The present dissertation adopts the terminology used in the cartographic approach. The term *topicalization* is used to refer to the construction in (16), in which a topic is syntactically connected with an empty category inside the comment. The term *focalization* is used to refer to constructions such as (17), where the preposed constituent is a contrastive focus.

5. Overview of the dissertation

This dissertation is organized in three different studies on the syntax of relativization. Chapter 2 introduces a construction that, to the best of my knowledge, has not been discussed before in the literature on relativization. This construction (henceforth referred to as *remnant-internal relativization*) involves RRCs that exhibit non-adjacency between the head noun and its modifier/complement, as illustrated in (18) and (19), which are from earlier stages of Portuguese.

- (18) os livros que eu compus da philosophia
the books that I wrote of.the philosophy

'the books of philosophy that I wrote'

(D. Pedro, *Livro dos Ofícios*, 15th century, in Pádua 1960: 70, cited in Martins 2004)

- (19) e esto por prool e verdade de hũa Licença
and this by favor and truth of a license
 que do dito senhor pera ello tenho
that from.the mentioned man for that have:1SG

'... and (I wrote this document) under the benefit and truth of a license from the aforementioned man that I have to (make) it.'

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)

The central claim of Chapter 2 is that remnant-internal relativization provides new evidence in favor of the raising analysis of RRCs (originally proposed by Schachter 1973 and Vergnaud 1974, 1985 and more recently revived by Kayne 1994, Bianchi 1999 and De Vries 2002).

Chapter 3 addresses RRC extraposition. A descriptive point of view is used to show that different languages and different stages of the same language may differ with respect to RRC extraposition properties. The contrasting properties discussed are: (i) the definiteness effect; (ii) extraposition from embedded positions, and (iii)

extraposition from pre-verbal positions. From a theoretical point of view, it is demonstrated no syntactic structure alone can derive the contrasting properties of RRC extraposition. To account for the variation found in the diachronic and cross-linguistic dimension, it is argued that RRC extraposition may involve two different structures: the specifying coordination plus ellipsis (De Vries 2002) and stranding (Kayne 1994). Languages and different stages of the same language may then diverge with respect to the structure they display. From this approach, it follows that the concept of extraposition might be descriptively useful (in unifying a variety of apparently related constructions) but lacks explanatory force because it does not unequivocally correspond to a single syntactic type.

Chapter 4 deals with ARCs. In line with Cinque (1982, 2008) and Smits (1988), it is argued that ARCs do not constitute a uniform syntactic type. This claim is supported by the study of a syntactic change that took place within the history of Portuguese, involving ARCs introduced by the complex relative pronoun *o qual* ‘lit. the which’ (henceforth *o qual*-ARCs). The investigation of this micro-variation demonstrates that the syntactic properties of *o qual*-ARCs have changed significantly over time. The contrasting properties observed are: (i) the possibility of having an additional internal head; (ii) restrictions on extraposition; (iii) restrictions on pied-piping; (iv) the possibility of taking clausal antecedents and (v) split antecedents; (vi) coordination of the *wh*-pronoun with another DP; (vii) illocutionary force; and (viii) the presence of a coordinator. To account for these contrasts, it is proposed that *o qual*-ARCs in CEP involve the head-raising analysis (Kayne 1994, Bianchi 1999), whereas in earlier stages of Portuguese they involve the specifying coordination analysis (De Vries 2006b). The dual approach to the phenomenon straightforwardly derives the variation in the syntax of ARCs found within a language and across languages, both in the synchronic and diachronic dimensions.

The conclusions are presented in Chapter 5.

Chapter 2

Remnant-internal relativization: new evidence for the raising analysis

Contents

1. Introduction.....	29
1.1. Goals	29
1.2. Chapter outline.....	32
2. The syntax of RRCs.....	33
2.1. The definition of <i>RRCs</i>	33
2.2. Competing analyses of RRCs	33
2.2.1. The adjunction analysis	34
2.2.2. The raising analysis	34
2.3. Arguments for the raising analysis	35
2.4. Summary	42
3. The syntax of split noun phrases.....	42
3.1. The definition of <i>split noun phrases</i>	42
3.2. Core properties of split noun phrases	42
3.3. Competing analyses	45
3.3.1. Movement analyses.....	45
3.3.2. Base-generation analyses	54
3.4. Summary	56
4. The syntax of remnant-internal relativization: first take	57
4.1. The definition of <i>remnant-internal relativization</i>	57
4.2. Empirical evidence	59
4.2.1. Split noun phrases in earlier stages of Portuguese.....	59
4.2.2. Remnant-internal relativization in earlier stages of Portuguese	63
4.3. Competing analyses	66

4.4. Problems with the <i>raising analysis of RRCs + movement analysis of split noun phrases</i> combination	71
4.5. Summary	73
5. The proposal	73
5.1. Preliminaries	74
5.1.1. The interpretation of remnant-internal relativization.....	74
5.1.2. The copy theory of movement in the PF side	77
5.1.2.1 Bošković and Nunes (2007).....	77
5.1.2.2 Stjepanović (2007).....	78
5.2. The derivation of remnant-internal relativization	79
5.2.1. Pattern I.....	80
5.2.2. Pattern II	82
5.2.3. Pattern III	88
5.3. Remnant-internal relativization in CEP?	90
5.4. Summary	91
6. Conclusion	93

1. Introduction

1.1. Goals

In the syntactic literature, the notion of a phrase is used to refer to a group of words that behave syntactically (and semantically) as a single unit. More often than not, the elements that make up a phrase are continuous, i.e., they involve words that appear next to one another. However, phrases may also surface in a discontinuous manner, split up into two or more parts. This can be observed in different phrase types (e.g., noun phrases, prepositional phrases) and in different linguistic environments.

This study focuses on discontinuous noun phrases (or split noun phrases) as exemplified in (1)-(4), from Croatian, French, Russian, and Dutch, respectively.¹

- (1) Knjige mi je Marija zanimljive preporučila.
books me has M. interesting recommended

‘Mary has recommended interesting books to me.’

(Fanselow and Ćavar 2002: 66)

- (2) Combien as-tu lu de livres?
how.many have.you read of books

‘How many books have you read?’

(Butler and Mathieu 2004: 2)

- (3) Chja vyshla kniga?
whose:nom came.out book:nom

‘Whose book came out?’

(Fanselow and Féry 2006: 1)

- (4) Wat heb je voor boeken gekocht?
what have you for books bought

‘Which kind of books have you bought?’

(Mark de Vries, p.c.)

These constructions have been documented in various studies (see, e.g., Corver 1990, Devine and Stephens 1999, Fanselow and Ćavar 2002, Butler and Mathieu 2004,

¹ The discontinuous members of the noun phrase are underlined for ease of reading.

Fanselow and Féry 2006, Kariaeva 2009) and have been approached from the syntactic, semantic, pragmatic and prosodic points of view. Of particular interest here is the idea, put forth by Fanselow and Cavar (2002: 69), that split noun phrases arise only in the context of operator movement. This generalization accounts for the typical cases reported in the literature (namely, interrogative *wh*-movement as in (1) and focus/topic movement as in (2)-(3)), but it predicts that other patterns may be found. In particular, it predicts that split noun phrases may also surface in relative clauses.

In this chapter, I will show that this prediction is correct: in earlier stages of Portuguese (and Latin), split noun phrases may arise in RRCs, a phenomenon that will henceforth be referred to as *remnant-internal relativization*. In this construction, an element that is thematically dependent on the head noun (either as a complement or as a modifier) does not appear adjacent to it but rather in an RRC-internal position, as illustrated in (5)-(8), from earlier stages of Portuguese:

(5) Casos que Adamastor contou futuros
cases:MASC.PL that A. told future:MASC.PL

‘(the) future events that Adamastor foresaw’

(L. de Camões, *Os Lusíadas*, Book V, 60, 16th century, cited in Lausberg 1967/1972: §331)

(6) os livros que eu compus da philosophia
the books that I wrote of.the philosophy

‘the books of philosophy that I wrote’

(D. Pedro, *Livro dos Ofícios*, 15th century, in Pádua 1960: 70, cited in Martins 2004)

(7) e qualquer que de nos primeiro morer
and any that of us first die:FUT.SUBJ

‘and whoever of us first die’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1266)

(8) e esto por prool e verdade de hũa Licença
and this by favor and truth of a license
 que do dito senhor pera ello tenho
that from.the mentioned man for that have:1SG

‘... and (I wrote this document) under the benefit and truth of a license from the aforementioned man that I have to (make) it.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)

In (5) and (6), the modifier/complement surfaces in the rightmost position of the noun phrase. For this reason, it can be structurally analyzed as occurring either in an RRC-internal position or in an RRC-external position as a second modifier following the RRC. However, the fact that the modifier/complement may occur in other positions than the rightmost one, namely in the left periphery of the relative clause (see (7)-(8)), shows that it is structurally RRC-internal.

Based on these facts, I submit that remnant-internal relativization can illuminate the long-standing debate between the right adjunction analysis of RRCs (originally proposed by Ross 1967, Chomsky 1977 and Jackendoff 1977) and the raising analysis of RRCs (originally proposed by Schachter 1973 and Vergnaud 1974, 1985, and more recently revived by Kayne 1994, Bianchi 1999 and De Vries 2002), providing evidence in favor of the latter. I argue that this is so for the following reasons.

- If the head and its modifier/complement were base-generated together in an RRC-external position (as proposed by the right adjunction analysis), the pattern in (7)-(8) could not be derived as it would require lowering the modifier/complement to a non-c-commanding position (see Fiengo 1977).
- If the head and its modifier/complement were generated separately (the head being CP-external – as proposed by the adjunction analysis – and the modifier/complement being CP-internal), the semantic dependency between the head and its modifier/complement (requiring that these elements be in a structural relation at some point of the derivation) would not be satisfied.

By contrast, there is a natural explanation for remnant-internal relativization if we assume that the head and its modifier/complement are merged together in the relativization site, as proposed by the raising analysis of RRCs. The fact that the modifier/complement is in an RRC-internal position at some point of the derivation suffices to explain why, under certain circumstances, it may surface CP-internally.

In line with these findings, I propose an analysis of remnant-internal relativization in terms of the copy theory of movement. Based on work by Fanselow and Ćavar (2002), Bošković and Nunes (2007) and Stjepanović (2007), I submit that the non-adjacency between the head and its modifier/complement is an effect of phonological deletion in the PF side of the grammar induced by PF requirements (for stress assignment and EPP).

With this background in mind, the central claim of this chapter is that remnant-internal relativization should be added to the arguments adduced in the literature in favor of the raising analysis of RRCs. The primary empirical evidence for this argument comes from earlier stages of Portuguese. However, data from other languages are also taken into account, namely from Latin and French.

1.2. Chapter outline

Before discussing the syntax of remnant-internal relativization, this chapter will briefly review the competing analyses proposed in the literature to account for RRCs and split noun phrases. Section 2 presents the definition of RRCs (2.1) and introduces the two most salient analyses of RRCs: the adjunction analysis and the raising analysis (2.2). Then Section 2.3 reviews some arguments that have previously been used in the literature to motivate the raising analysis, namely (i) binding theory, (ii) quantifier binding, (iii) scope assignment, (iv) idioms, (v) degree relatives, (vi) the interpretation of adjectival modifiers, and (vii) head-internal relatives. Section 3 is devoted to the syntax of split noun phrases. It starts by introducing the concept of the split noun phrase (3.1) and the core properties of split noun phrases (3.2). Then, it equates the competing analyses available in the literature to account for this phenomenon (3.3).

With these preliminaries in mind, Section 4 focuses on the syntax of remnant-internal relativization. After characterizing this construction (4.1), it offers an overview of data from Latin and earlier stages of Portuguese (4.2). Then, it proceeds to examine the various possible accounts of remnant-internal relativization, showing that the movement-based approaches to splits associated with the raising analysis of RRCs better address the properties of this construction (4.3). The problems faced by the *movement-based approach to splits + raising analysis of RRCs* combination are discussed in Section 4.4.

Finally, Section 5 offers an analysis of remnant-internal relativization in terms of the version of the theory of movement proposed by Bošković and Nunes (2007) (based on previous work by Bošković 2001, 2002, 2004a,b and Nunes 1999, 2004). Section 5.1 provides background, focusing on the interpretation of remnant-internal relativization (Section 5.1.1) and on the basics of the copy theory of movement (in the PF side). Section 5.2 presents the analysis, and Section 5.3 explains the contrast

between Contemporary European Portuguese (CEP) and earlier stages of Portuguese with respect to remnant-internal relativization. Section 6 summarizes the chapter.

2. The syntax of RRCs

2.1. The definition of RRCs

Much of the traditional and generative literature has assumed that relative clauses can be semantically classified as restrictives (RRCs) or appositives (ARCs). RRCs are interpreted intersectively, i.e., as restricting the denotation of the antecedent. ARCs are interpreted as providing additional information about the antecedent. This is illustrated in (9). In the RRC in (9)a, there is another potential group of students that did not participate in the research, whereas in the ARC in (9)b, there is only one group of students in the domain of discourse and no contrast with other students.²

(9) a. The students who participated in the research showed improvement in this area.

b. The students, who participated in the research, showed improvement in this area.

2.2. Competing analyses of RRCs

Much of the debate on the syntax of RRCs has centered on the contrast between the *adjunction analysis* and the *raising analysis*. These analyses are illustrated in a simplified form in (10)a-b, respectively:

(10) a. The book [\emptyset_i I read t_i].

b. The [$book_i$ I read t_i].

The major difference is that the head noun *book* is generated in the matrix clause in the adjunction analysis (see (10)a), but it is raised from within the relative in the

² In recent studies, there seems to be a growing consensus that not all relative clauses fit in the traditional RRC/ARC dichotomy. Some authors (see, e.g., Carlson 1977, Heim 1987 and Grosu and Landman 1998) have identified a “third type” of relative clause: the so-called *degree relative* – or *amount, maximalizing, third type relatives*. For more details on degree relatives, see Section 2.3.

raising analysis (see (10)b). This is why these theories are also called, respectively, *head external analysis* and *head internal analysis* (or *promotion analysis*).

2.2.1. The adjunction analysis

The development of the adjunction analysis goes back to Ross (1967), Chomsky (1977), and Jackendoff (1977).³ For the current discussion, I will consider the version of the adjunction analysis proposed by Demirdache (1991), according to which the RRC is right-adjoined to the NP-projection (see (11)). The head noun originates outside of the RRC, and the relative CP involves the A'-movement of a relative operator. The operator is linked to the head-NP via predication, semantically interpreted as intersective modification.

(11) [_{DP} The [_{NP} [_{NP}book]_i [_{CP} Op_i I read t_i]]]

Importantly, in the adjunction analysis, the head is not directly represented in the RRC and the RRC is c-commanded by the D° head.

2.2.2. The raising analysis

The raising analysis was originally proposed by Schachter (1973) and Vergnaud (1974, 1985). It was later revived by Kayne (1994) within a general antisymmetric framework of syntax. Kayne combines raising with the D-complement hypothesis, which states that the relative clause is the complement of the outer determiner. This approach was worked out in more detail by Bianchi (1999) and De Vries (2002).

For the current discussion, I will adopt the version of the raising analysis proposed by Kayne (1994) (with some of the implementations developed by Bianchi 1999 and De Vries 2002). Under the raising analysis, the head originates inside the relative clause and is raised to [Spec, CP] of the relative clause. Moreover, the D° head selects the relative CP as its complement. The representation is given in (12), respectively, for *that*-relatives, zero relatives and *wh*-relatives (Kayne 1994):

³ For the syntax of relative clauses in Portuguese under the adjunction analysis, see Brito (1991).

- (12) a. [DP [D° the] [CP book_i [C° that [I read t_i]]]]
 b. [DP [D° the] [CP book_i [C° C° [I read t_i]]]]
 c. [DP [D° the] [CP book_i [which t_i]]_j [C° C° [I read t_j]]]

It is worth noting that Kayne (1994) is working within a system that does not permit right-adjunction, so the adjunction analysis of relative clauses is not an option in his framework.

2.3. Arguments for the raising analysis

In this section, I provide a short summary of the strongest arguments that have been adduced in the literature in favor of the raising analysis. For further discussion of arguments and counter-arguments regarding the raising analysis, see Bianchi (1999), Alexiadou et al. (2000), De Vries (2002), Bhatt (2002), Salzmänn (2006), and discussions in Borsley (1997) and Bianchi (2000).

A. Binding theory

Reconstruction was originally proposed in the Government-Binding Theory as a process that occurs in the mapping from S-structure to Logical Form (LF), moving some constituents back to their D-structure positions. It has been considered as a reliable diagnosis for movement because a constituent that has undergone movement behaves if it were in the position occupied before movement at the level of computation at which binding principles apply. These facts can be observed, for example, in interrogative *wh*-movement. In (13)a, the anaphor *himself* has to be c-commanded by its antecedent *John*, and consequently it behaves as if it were in its base position. In (13)b, *John* is interpreted in its base position, (13)b being ruled out as a violation of Principle C of the binding theory.

- (13) a. Which picture of himself_i did John_i buy t?
 b. *Which picture of John_i did he_i buy t?

Based on the idea that reconstruction effects can be a diagnosis for movement, the reconstruction of the relative clause head has been widely discussed by proponents and opponents of the raising analysis. One traditional argument in favor of the raising analysis is the presence of reconstruction effects in sentences like (14):

(14) [The portrait of himself_i that John_i painted t] was extremely flattering.

The adjunction analysis makes the wrong predictions about (14). Because, under this analysis, the head is base generated in a relative clause external position, the anaphor *himself* cannot be bound by *John*_i; hence (14) should be ungrammatical, in violation of Principle A of the binding theory.

In contrast, the pattern of grammaticality of (14) can be explained under the raising analysis. Because the head is base generated in a relative clause internal position, the anaphor embedded in the head can be reconstructed in its base position and, consequently, be bound by the subject of the relative clause.⁴

⁴ Cechetto (2005) claims that when a transitive noun like *picture* is used in these tests, the anaphor can be bound by an NP-internal PRO that sits in the subject position of the NP (see Giorgi and Longobardi 1991). Therefore in sentences like (i), the absence of Principle A violation is not a case of reconstruction because the position in which *himself* occurs is c-commanded by a suitable antecedent for *himself* (PRO).

(i) [_{DP} The [_{NP} PRO_i picture of himself_i] [that John_i likes T most]] (was never on display)

Note, however, that the same effect appears if an unaccusative noun like *naufragio* ‘shipwreck’ is involved, as in (ii), taken from Italian. Because in this case no internal PRO is available, the absence of Principle A violation indicates that reconstruction is at stake.

(ii) Il naufragio della propria_i nave que Gianni_i teme t è quello che
the shipwreck of.the self ship that G. fears is that that
 può avvenire durante la regata principale
can happen:INF during the regatta main

‘The shipwreck of self’s ship that Gianni fears is the one that can happen during the main regatta.’

(Cechetto 2005: 188)

Essentially, the same point is made in De Vries (2002: 80-82), who shows that even when transitive nouns are involved, reconstruction is possible if PRO is interpreted as non-coreferent with the anaphor. Such an interpretation is forced in sentences like (iii), from Dutch; in this case, the story-teller must be someone other than the hearer, Paul. Because PRO cannot bind the anaphor, reconstruction is necessary to avoid a Principle A violation.

(iii) De verhalen over zichzelf die Paul hoorde, waren pure leugens.
the stories about SE-SELF which P. heard were mere lies

‘The story about himself which Paul heard were mere lies.’

(De Vries 2002: 80)

For criticism of the use of this test as evidence for the raising analysis, see Platzack (2000: 267, 299 fn. 3).

B. *Quantifier binding*

Quantifier binding requires that a quantificational noun phrase c-commands a bound pronominal. To test quantifier binding in relative clauses, the relevant configuration involves a pronoun embedded in the head bound by a quantifier inside the relative clause, as in (15), from English, and (16), from Italian.

(15) The picture of his_i mother that every soldier_i kept t wrapped in a sock was not much use to him.

(Salzmann 2006: 22)

(16) La parte del suo_i stipendio che ho anticipato t ad ogni
the part of his salary that have anticipated to every
 impiegato_i verrà sottratta dalla busta paga
clerk come:FUT deducted of.the pay-sheet

‘The part of his salary that I paid in advance to every clerk will be deducted from the pay-sheet.’

(Bianchi 1999: 124)

On the assumption that a pronoun cannot incidentally corefer with a quantified expression, quantifier binding in (15)-(16) requires that the pronouns *his* and *suo* ‘his’ be in the gap position of the relative clause at the relevant level.

The required configuration is obtained by the raising analysis: the head is base-generated inside the RRC and, as a result, can be interpreted in its trace position at LF. Such an explanation is not, however, available for the adjunction analysis: because the head does not originate within the relative clause, it cannot be interpreted in an RRC-internal position.

C. *Scope assignment*

The head of a relative clause can be reconstructed for the purposes of scope assignment. This can be observed in (17):

(17) I called the two patients that every doctor will examine t tomorrow

Sentence (17) can have a distributive reading or a wide-scope reading (see, e.g., Salzmann, 2006: 22). In the distributive reading, each doctor will examine a different set of two patients that every doctor examines; in the wide-scope reading, all doctors

examine the same two patients. Crucially, the distributive reading is only possible if the numeral is reconstructed under the scope of the universal quantifier.

In the raising analysis, this requirement is fulfilled because the head can be reconstructed in its base position, under the scope of the subject of the relative clause. This is not possible under the adjunction analysis: the head originates in the matrix and therefore cannot reconstruct to an RRC-internal position.

D. *Idioms*

This argument is based on idiom chunks (or collocations), such as *to make headway* and *to take advantage of*. In idiom chunks, the verb and the object form a fixed expression and thus must be merged together (see Schachter 1973, Vergnaud 1974).⁵

Crucially, the object of an idiom chunk can surface as the head of a relative clause, as in (18):

(18) The headway that we made was satisfactory.

(attributed to Brame 1968 ms., cited in Schachter 1973)

In the raising analysis, the conditions on the adjacency of the parts of the idiom are met in LF: because *headway* is base generated as the complement of *made* within the relative clause, it can be reconstructed in its base position.

In the adjunction analysis, the grammaticality of (18) is unexpected: the verb and the head are not merged together and therefore cannot become adjacent in LF via reconstruction.

E. *Degree relatives*

It has been argued in the literature that some relative clauses do not fit in the traditional dichotomy of ARCs/RRCs. Examples of these so-called degree relatives are given in (19):

⁵ The plausibility of such an argument is underlined by Bhatt (2002), who states “The argument from idioms is based on an assumption about how idioms are interpreted/stored in the lexicon. We need to assume that the domain of special meaning is local. In other words, two elements that can be arbitrarily far apart from each other in a structure at the point of Merge cannot be given a special meaning” (Bhatt 2002: 47).

- (19) a. I took with me the books that there were on the table.
 b. Mary drunk the wine that there was in the bottle.

The relative clauses in (19) refer to the amount of *books/wine* rather than to the fact that there were *books on the table* or *wine in the bottle*. Moreover, (19)a means that *I took all the books with me* and (19)b means that *Mary drank all the wine in the bottle*.

On the surface, degree relatives are very much like RRCs. However, from the interpretative point of view, it is clear that a degree relative is not combined with the head noun through intersection, as happens in an RRC:

- (20) I took with me the three books which were on the table.

To derive the amount reading of degree relatives, it has been assumed that the material that is syntactically external to the relative clause must be semantically interpreted in a relative clause internal position. In the account of Grosu and Ladman (1998), the head within the relative clause provides a degree variable, and an operation of maximalization applies at the clausal level. The crux of the semantic interpretation of degree relative is then the null degree phrase *d many books* that is placed inside the relative clause:

- (21) a. (books) that there were ___ on the table
 b. (books) that there were (d many books) on the table
 c. $\{d: \exists x[\text{BOOK}(x) \text{ and } |x| = d \text{ and ON-THE-TABLE}(x)]\}$
 (Grosu and Landman 1998)

Under the raising analysis, the amount reading is derived straightforwardly if one assumes that the head is reconstructed inside the relative clause and the abstraction is over a degree variable. The same effect cannot, however, be obtained under the adjunction analysis because the head is merged external to the relative clause.

F. *The interpretation of adjectival modifiers*

Bhatt (2002) argues that certain adjectival modifiers that occur in association with the head can be interpreted as being internal to the relative clause. See, for instance,

sentence (22), which is ambiguous between a high and a low reading of the adjective *first*:

(22) the first book that John said that Tolstoy had written

(Bhatt 2002)

In the high reading, the order in which the books were actually written is irrelevant. The relevant order is the order in which John names the books. In the low reading, the order of John's naming it is irrelevant and what matters is the order in which the books were written.

Bhatt (2002) shows that the ambiguity of (22) can be derived if the head and its modifier are reconstructed into different positions inside the relative clause. He reasons that this is possible under the raising analysis because the head and its modifier originate inside the relative clause and undergo leftward movement; it is therefore possible to decide which copy to interpret in LF.⁶ This not possible under the adjunction analysis because the head is not directly represented inside the relative clause.

G. *Head-internal relatives*

From a cross-linguistic perspective, there is wide variation with respect to the relative positions of the head and the relative clause: the head may precede the relative clause (resulting in a post-nominal or head-initial relative clause), follow the relative clause (resulting in a pre-nominal or head-final relative clause), or appear inside the relative clause (resulting in a circum-nominal or head-internal relative clause).⁷ These different types are, respectively, illustrated in (23), from English; (24), from Mandarin Chinese; and (25), from Dagbani (a Gur language spoken in Ghana):

⁶ For further discussion of this argument, see Heycock (2005).

⁷ Post-nominal and pre-nominal relative clauses are sometimes grouped together under the label *head-external relative clauses* (as opposed to *head-internal relative clauses*).

(23) The book that you gave me was very interesting.

(24) Wǒ bǎ nǐ gěi wǒ de shū diūdiào-le.
I ACC you give I NR book loose-PERF

'I have lost the book that you gave me.'

(Lehmann 1984, cited in De Vries 2002: 16)

(25) A mi [o nà ti saan-so ləgri] la.
you know he SR give stranger-SPC/LIV money PTL

'You know the stranger whom he gave the money.'

(Lehmann 1984, cited in De Vries 2002: 16)

Pre-nominal relative clauses are also found, for instance, in Basque, Korean, Tigré, Ijo, Lahu and Nama. Head-internal relative clauses occur in Quechua, Latin, Lakota and Tibetan, among other languages (see Lehmann 1984, Cole 1987, Culy 1990, Comrie and Kuteva 2008). For an overview, see De Vries (2002: 376ff).

The existence of head-internal relative clauses has been taken as a strong argument in favor of the raising analysis (see Bianchi 1999: 61ff, De Vries 2002: 77, 135ff). This construction can be characterized as a nominalized sentence that modifies a nominal, overt or not, internal to the sentence (Culy 1990), with the following appearance (De Vries 2002: 136):

(26) [_{DP} [_{CP} ... **N** ...] (**D**)]

The fact that, in (26), the head surfaces in the argument position inside the relative clause has led some adepts of the raising analysis to postulate the same base position for the head in head-external relative clauses. In accordance with the principle that derivations are uniform, head-external and head-internal relative clauses would then involve the same derivation. The only extra assumption would be that head-internal relative clauses involve covert (and not overt) movement of the head in LF.

This hypothesis is not, however, available in the adjunction analysis. Under this approach, head-internal and head-external relatives must involve two completely different derivational stories, which may be seen as a drawback in terms of the principle of uniformity of the derivation.

2.4. Summary

Intuitively, the phenomenon of remnant-internal relativization seems to be associated with other structures available in the grammar, namely RRCs and split noun phrases. This section has reviewed the two most salient analyses of RRCs: the adjunction analysis and the raising analysis. Because remnant-internal relativization will be taken as a new argument in favor of the raising analysis, particular attention was given to the arguments that have previously been used in the literature to motivate the raising analysis. The next section discusses the syntax of split noun phrases, the other structure that, on par with RRCs, appears to be involved in remnant-internal relativization.

3. The syntax of split noun phrases

3.1. The definition of *split noun phrases*

The elements of a noun phrase may surface in a continuous or discontinuous manner, as illustrated schematically in (27). In continuous noun phrases, the elements that make up the phrase appear next to one another, while in discontinuous noun phrases (or split noun phrases) the elements can be split into two or more parts.

- (27) a. ... [D N Adj] ... (continuous noun phrase)
 b. ... D ... [N Adj] ... (split noun phrase)
 c. ... [D N] ... [Adj] ... (split noun phrase)

These constructions appear under different designations in the literature – *split noun phrases*, *discontinuous noun phrases*, *partial fronting*, *incomplete category fronting*, *left branch extraction*, *hyperbaton*. In this study I adopt the term *split noun phrases* (and sometimes *discontinuous noun phrases*) to cover any interrupted sequence of elements in a noun phrase that would normally surface in a continuous manner.

3.2. Core properties of split noun phrases

Over the past few years, split noun phrases have been the focus of much attention in the typological and theoretical literature (see, in particular, Devine and Stephens

1999, Fanselow and Ćavar 2002, Butler and Mathieu 2004, Fanselow and Féry 2006, and Kariaeva 2009). I now offer a brief overview of their core properties.

I. Split noun phrases can retain the order of elements found in the continuous counterpart (*simple splits* or *pull-splits*) or can invert this order (*inverted splits*) (see Fanselow and Ćavar 2002: 68). This is illustrated in (28)b and (28)c, from Ukrainian:

- (28) a. Marija maje bahato krisel.
M. has.got many chairs:GEN.PL
 ‘Mary has got many chairs.’
 b. Bahato maje Marija krisel.
 c. Krisel Marija maje bahato.

(Fanselow and Féry 2006: 5)

II. The syntactic distinction mentioned in I. tends to correlate with a prosodic distinction: simple splits tend to be cohesive (i.e., the two parts of the split are integrated into a single intonation phrase), whereas inverted splits tend to be non-cohesive (i.e., the two parts of the split are separated into two intonation phrases) (see Fanselow and Féry 2006).

III. Split noun phrases can stretch across more than two discontinuous parts. A case of tripartite discontinuity is given in (29), from Ukrainian:

- (29) a. Ivan kupyv duže velyku mašynu
John:NOM bought very big:F.SG.ACC car:F.SG.ACC
 ‘John bought a very big car.’
 b. Duže Ivan velyku kupyv mašynu
very John:NOM big:F.SG.ACC bought car:F.SG.ACC
 ‘John bought a VERY BIG car.’

(Kariaeva 2009: 207)

IV. Split noun phrases arise in the context of operator movement only (see Fanselow and Čavar 2002: 69),⁸ in interrogative *wh*-movement, as in (2)-(3) (repeated here as (30)-(31)), or in focus/topic movement, as in (1) (repeated here as (32)).

- (30) Combien as-tu lu de livres?
how.many have.you read of books
 ‘How many books have you read?’

(Butler and Mathieu 2004: 2)

- (31) Čija vyšla kniga?
whose:NOM came.out book:NOM
 ‘Whose book came out?’

(Fanselow and Féry 2006: 1)

- (32) Knjige mi je Marija zanimljive preporučila.
books me has M. interesting recommended
 ‘Mary has recommended interesting books to me.’

(Fanselow and Čavar 2002: 66)

On the areal distribution of split noun phrases, studies of other languages and cross-linguistic systematizations have yielded evidence for the idea that splits are frequently found in the world’s languages, although they are quite uncommon in Western European languages (see Fanselow and Féry 2006). This situation changes radically when other European languages are considered. As Fanselow and Féry (2006) put it: “In Europe, the situation changes dramatically when one crosses the river Rhine or the Isonzo: one enters ‘split country’, which extends to the Pacific Ocean.” In this area, split noun phrases are reported to occur, for instance, in the “Eastern” Germanic languages (Dutch, German, Swedish), Romanian, all Slavic languages, the Baltic

⁸ There is no consensus in the literature regarding the relation that can be established between the parts of the splits. While for Fanselow and Čavar (2002: 69), the parts of the splits necessarily establish an operator-variable relation, for Kariaeva (2009) they can simply stand in an agreement relation, as in sentence (i), from Ukrainian:

- (i) Ivan červonu kupyv mašynu
John:NOM red:F.SG.ACC bought car:F.SG.ACC
 ‘John bought a RED car.’

(Kariaeva 2009: 70)

languages of Lithuanian and Latvian, the Finno-Ugric languages, Albanian, Ancient and Modern Greek, and the Altaic languages.

3.3. Competing analyses

Much of the debate on the syntax of split noun phrases has centered on the contrast between movement and base-generation analyses. Movement analyses propose that the parts of the split are related by movement. Base-generation analyses claim that the parts of the split noun phrase should be analyzed as elements that are structurally independent of each other. The main idea underlying these proposals is outlined in Sections 3.3.1 and 3.3.2.

3.3.1. Movement analyses

The movement analyses of split noun phrases can be grouped together into four main types: (i) simple movement analyses; (ii) regeneration; (iii) remnant movement; and (iv) distributed deletion. These types are listed in A. through D. below and are discussed in turn.⁹

A. *Simple movement analyses*

Simple movement analyses posit that split noun phrases are derived from extraction of an element X out of a constituent Y (see, e.g., van Riemsdijk 1989). In the early period of generative syntax, these movement-based approaches faced a serious problem: they seemed to go against the generalization that movement can only apply to maximal or minimal projections. This can be witnessed from the German sentences in (33), which show that any segment of *keine interessanten neuen Bücher* can be extracted. Under the assumption that noun phrases only involve one maximal projection (NP), the part of the split noun phrase that undergoes movement in (33)c,d only forms a submaximal N'-projection and therefore should not be able to undergo movement.

⁹ In Section 3.3.1. I closely follow Fanselow and Ćavar's (2002) criticism of the movement-based analyses.

- (33) a. Sie hat keine interessanten neuen Bücher gekannt.
she has no interesting new books known
 ‘She did not know any interesting new books.’
 b. [Bücher]_i hat sie [keine interessanten neuen t_i] gekannt.
 c. [Neue Bücher]_i hat sie [keine interessanten t_i] gekannt.
 d. [Interessanten neue Bücher]_i hat sie [keine t_i] gekannt.
 e. [Keine interessanten neue Bücher]_i hat sie t_i gekannt.

(Fanselow and Ćavar 2002: 70)

Research in generative grammar on the structure of DP and its similarities with the structure of clauses (see Abney 1987) in the late 1980s and the 1990s provided a new way of looking at the movement analysis of split noun phrases. One important development was the proposal that split noun phrases may involve the movement of different functional projections of the noun phrase. Under this view, the elaborate syntactic structure of the noun phrase in (33) would look like (34), and the problem mentioned above could easily be solved: split noun phrases involve the leftward movement of different functional projections within the noun phrase.

- (34) [DP [D keine] [AGR-A1-P [AP interessanten] [[AGR-A1 e] [AGR-A2-P [AP neuen] [[AGR-A2 e] [Nom-P Bücher]]]]]]]

(Fanselow and Ćavar 2002: 70)

Despite this welcome development, other problems for the simple movement analyses were reported in the literature, namely (i) the movement of non-constituents; (ii) imperfect splits; (iii) morphological adjustments; and (iv) movement across islands (see Fanselow and Ćavar 2002 for an overview).

The movement of non-constituents was a problem in accounting for simple (or pull) splits, as in (35), from Croatian. Simple movement analyses cannot generate sentences like (35) because there is no constituent that includes P+Det but excludes the noun that could be moved to the left to form a split noun phrase.

- (35) Na kakav je Ivan krov skocio?
on what.kind has I. roof jumped

‘On what kind of roof has Ivan jumped?’

(Fanselow and Ćavar 2002: 71)

Another problem faced by the simple movement analyses concerns the existence of imperfect splits, i.e., split noun phrases that have no well-formed source in a movement account. This is the case of preposition doubling exemplified in (36), from German, where the preposition heading the PP appears in both parts of an inverted split.

- (36) In Schlössern habe ich noch in keinen gewohnt
in castles have I yet in no lived

‘I have not yet lived in any castles.’

(Fanselow and Čavar 2002: 69)

The same phenomenon can be observed in the so-called determiner spreading found in Modern Greek. In this construction, the determiner may show up only in the first part of the split noun phrase (see (37)a) or in both parts (see (37)b), depending on the dialect/register (see Mathieu and Sitaridou 2005).

- (37) a. To KOKINO agorase forema
the:ACC.NEUT.SG red:ACC.NEUT.SG bought dress: ACC.NEUT.SG

- b. To KOKINO agorase to forema.
the:ACC.NEUT.SG red:ACC.NEUT.SG bought the:ACC.NEUT.SG dress: ACC.NEUT.SG

‘She bought the RED dress (not the blue one).’

(Mathieu and Sitaridou 2005: 240)

It is widely recognized in the literature that the phenomena of preposition doubling and determiner spreading are a problem for the simple movement account because there is not enough space in a single continuous phrase for the material occurring in the two discontinuous parts.

A third problem with the simple movement analysis concerns the so-called morphological adjustments (see Fanselow and Féry 2006). This term refers to a surprising property of split noun phrases: the parts of a split noun phrase can take morphologically different shapes than they do in their continuous counterpart. See the contrasts in (38)-(39), from German:

- (38) a. Er hat kein Geld
he has no:WEAK money

‘He has no money.’

b. Er hat keines
he has no:STRONG

(39) a. Er hat kein Geld
he has no:WEAK money

b. Geld Er hat keines
money he has no:STRONG

(Fanselow and Féry 2006: 55)

In German, the morphological shapes of quantifiers and adjectives are dependent on the presence of a noun. If a noun is present, as in (38)a, the negative quantifier *kein* bears a weak inflection; if a noun is not present, as in (38)b, it obligatorily carries a strong inflection. As can be seen in (39)b, when a split noun phrase is involved, the quantifier *kein* obligatorily bears a strong inflection. This seems to militate against a simple movement approach in that the two parts inflect as if they were independent noun phrases.

A fourth problem is that split noun phrases are insensitive to some island constraints (see Fanselow and Ćavar 2002). Taking as an example the case of German, we see that subjects (of non-unaccusative verbs, at least) are generally islands for extraction (see (40)a), but they can nevertheless be split up (see (40)b).

(40) a. *An Maria haben mir keine Briefe gefallen.
to M. have me no letters pleased

‘No letters to Mary have pleased me.’

b. Briefe an Maria gefallen mir keine
letters to M. please me no

‘As for letters to Mary, they do not please me.’

(Fanselow and Ćavar 2002: 72)

Different solutions have been proposed in the literature to circumvent these problems, namely: (i) regeneration; (ii) remnant movement; and (iii) distributed deletion. As will become clear below, some of these approaches are better equipped than others to handle the above-mentioned problems. This is explained, at least to some extent, by the fact that some of them were originally conceived as solutions to different very specific phenomena.

B. *Regeneration*

Van Riemsdijk (1989) proposes to account for some of the above-mentioned properties of split noun phrases by means of a process he refers to as *regeneration*. The core of his proposal is that the movement of X' projections is not precluded in principle. Hence, what makes X' movement rare is not a restriction on Move α , but rather a well-formedness condition that applies to the S-structure, which disallows any X' not dominated by its maximal projection node. Some languages simply ban this configuration at the S-structure level, whereas other languages may resort to a mechanism of repair that allows the regeneration of the missing structure and, in some cases, even the relexicalization of the regenerated structures, as illustrated in (41):

- (41) a. Regeneration: $[_{CP} [_{N'}]_i] [_{C'} \dots] \Rightarrow [_{CP} [_{NP} [_{N'}]_i] [_{C'} \dots]$
 b. Relexicalization: $[_{CP} [_{NP} \text{determiner } [_{N'}]] [_{C'} \dots]$

This approach straightforwardly explains two of the problematic properties of split noun phrases mentioned above: (i) the possibility of moving X' projections is linked to the availability of regeneration (see (41)a), whereas the existence of imperfect splits (e.g., determiner spreading) is correlated with the process of relexicalization (see (41)b). As Hoof (2005) notes, this approach is also able to derive the morphological adjustments discussed above by assuming that the spell out of the strong/weak inflection is not caused by a specific lexical-grammatical feature but is instead postponed until the movement of the first part of the split takes place.

However, the regeneration analysis also has its flaws. First, it fails to account for the movement of non-constituents (as in (35)) and for the movement across islands. Second, the mechanism of regeneration has been criticized on independent grounds for introducing unnecessary complications to the movement analysis (see Fanselow and Ćavar 2002).

C. *Remnant movement*

An alternative approach to the syntax of split noun phrases that deserves special attention is the remnant movement analysis. The normal instantiation of remnant movement starts with the movement of an element X out of a constituent. Then, the

whole constituent, which contains the trace of X, rises to its designated position, as is depicted in (42):

(42) [_{YP} ... Y ... t_i]_j ... X_i t_j

An analysis along these lines is proposed by Androutsopoulou (1997) for the split adjectival construction in Modern Greek, illustrated in (43)b. Under this approach the nominal first rises to the Spec of a Clitic Voice Phrase, which functions as the clause internal topic position and then the whole complex containing the trace of the nominal moves to the Specifier of the Focus Phrase in the left periphery of the clause. For other attempts to explain split constructions in terms of remnant movement, see, e.g., Sekerina (1999) for Russian; Franks and Progovac (1994) and Bašić (2004) for Serbo-Croatian.

- (43) a. *Idha to forema (to) kokino.*
saw:1SG the dress the red
 ‘I saw the red dress.’
- b. To KOKINO *idha (to) forema*
the red saw:1SG the dress
 ‘I saw the RED dress.’

(Androutsopoulou 1997, cited in Butler and Mathieu 2004: 174)

A welcome result of this analysis is that it straightforwardly explains why a non-constituent appears to undergo movement. Note, however, that the remnant movement approach faces serious problems in explaining other properties of split noun phrases, namely, the repetition of phonetic material in imperfect splits, the occurrence of morphological adjustments, and the possibility of having split noun phrases that disregard standard islands for movement (see Fanselow and Ćavar 2002 for further details).

D. Distributed deletion

Assuming the copy theory of movement (see Chomsky 1995), according to which moved elements leave copies behind that are subsequently deleted, Fanselow and Ćavar (2002) argue that split noun phrases are best analyzed in terms of distributed deletion. According to this version of the copy theory of movement, split noun

phrases do not involve the extraction of an element from a constituent. Instead, it is assumed that a complete noun phrase is copied to the left and that the splitting results from the fact that the deletion operation may partially affect both the upstairs and the downstairs copies of the moved constituent.

In a nutshell, the deletion operation works as follows. First, the relevant noun phrase undergoes leftward movement, leaving a copy behind. The copies are then each deleted at PF, as illustrated in (44). If the lower copy is completely deleted, we get a continuous (i.e., non-split) noun phrase (see (44)c); if one element is deleted in the higher copy and the other is deleted in the lower copy, we get a split noun phrase (see (44)d) (cf. Fanselow and Ćavar 2002 for further details).¹⁰

- (44) a. hat er keine Bücher gelesen
 has he no books read

Copying the noun phrase ⇒

- b. keine Bücher hat er keine Bücher gelesen

Full deletion of lower copy (continuous noun phrase) ⇒

- c. keine Bücher hat er ~~keine Bücher~~ gelesen

Partial deletion in both copies (split noun phrase) ⇒

- d. ~~keine~~ Bücher hat er keine ~~Bücher~~ gelesen

(Fanselow and Ćavar 2002: 84)

According to Fanselow and Ćavar (2002), it is the pragmatic structure that determines the occurrence of continuous/split noun phrases. Continuous noun phrases emerge when the noun phrase is linked only to one feature. This is the case of the continuous noun phrase in (44)c, which is associated with a +TOP feature. On the other hand, split noun phrases emerge when the noun phrase is linked at least to two different pragmatic features that cannot be checked in the same structural position. This is the case of (44)d, where the first part of the split bears a +TOP feature, whereas the second part bears a +FOC feature (which, under Fanselow and Ćavar's proposal, is checked in a specific lower focus position).

¹⁰ It is worth noting here that the partial deletion of copies (also known as *scattered deletion*) has been independently argued in the literature. For further details, see, e.g., Bošković and Nunes (2007) Bošković (2001), Nunes (1999), and Wilder (1995).

In more concrete terms, what this means is that split noun phrases involve two instances of movement, schematically represented in (45). The heads H^p and H^q have two different semantic or pragmatic features (p , q) and attract a phrase bearing the corresponding feature.

- (45) a. [H^p ... [H^q ... [XP a^p [b c^q]]]]
 b. [[XP a^p [b c^q]] [H^p ... [[XP a^p [b c^q]] [H^q ... [XP a^p [b c^q]]]]]]]
 (Fanselow and Ćavar 2002: 85)

Assuming that the phonetic realization of copies is regulated by the *Spell-Out Principle* in (46), the structure of split topicalization in (44)d would then look like (47) (where q corresponds to a +TOP feature and p to a +FOC feature).

(46) *Spell-Out Principle*:

Suppose $C = \langle C_1, C_2 \rangle$ is formed because a strong feature of H has attracted XP and suppose that H checks the operator features $f_1 \dots f_k$ of XP . Then the categories bearing $f_1 \dots f_k$ must be spelt out in C_1 .

- (47) [[XP a^p [b] q] [H^q ... [[XP a^p [b] q] [H^p ... [XP a^p [b] q]]]]]]]

Distributed deletion has many advantages over the movement-based approaches considered thus far. First, it explains the apparent movement of non-constituents illustrated in (48) (repeated from (35) above). Under distributed deletion, it is the whole PP (*Na kakav krov*) that undergoes leftward movement; hence, what looks like the movement of non-constituents is in fact the result of partial phonological deletion of different copies.

- (48) Na kakav je Ivan krov skocio?
on what.kind has I. roof jumped
 ‘On what kind of roof has Ivan jumped?’

(Fanselow and Ćavar 2002: 71)

Second, imperfect splits, such as the preposition doubling exemplified in (49) below (repeated from (36)), can be derived by assuming that it is the whole PP that undergoes movement, leaving a copy behind. Then, if the language tolerates multiple realizations of the same element, the deletion process removes portions of the phrases in the copy relation.

(49) In Schlössern habe ich noch in keinen gewohnt
in castles have I yet in no lived

‘I have not yet lived in any castles.’

(Fanselow and Ćavar 2002: 69)

Third, morphological adjustments as in (50) below (repeated from (39)) can be derived from distributed deletion if one assumes that the morphological shape of the determiner or adjective is determined after copying and deletion. According to the proposal of Fanselow and Ćavar (2002), the morphemes are merged into a syntactic representation as abstract entities and, when spelled out, they must meet the lexical and morphological wellformedness conditions for DPs. This explains why the quantifier *kein* bears a strong inflection in (50)b, but not in (50)a. Each part of the split noun phrase is dominated by the DP node and therefore must obey the wellformedness conditions for noun phrases in German.¹¹ In (50)a, the negative quantifier carries a weak inflection because the noun phrase *kein Geld* ‘no money’ contains a noun; in (50)b the quantifier in the second part of the split carries a strong inflection because, after copying and deletion, the noun phrase does not contain a noun.

(50) a. Er hat kein Geld
he has no:WEAK money

b. Geld Er hat keines
money he has no:STRONG

(Fanselow and Féry 2006: 55)

Finally, split noun phrase formation does not respect islands because it does not involve extraction out of a noun phrase; it is the whole constituent that undergoes leftward movement.

As this brief discussion shows, the deletion analysis is better equipped to handle the problems raised above than are the other movement-based approaches. This fact

¹¹ As already mentioned, these conditions dictate that: (i) if a noun is present, the negative quantifier *kein* bears weak inflection; (ii) if a noun is not present, the negative quantifier must carry strong inflection.

follows from the less constrained nature of the movement (not involving extraction of the noun phrase) and from the mechanism of partial deletion. However, it should be noted that distributed deletion is subject to the problem of overgeneration. Some authors, including Bošković (2005) and Kariaeva (2009), point out that an item can, in principle, be spelled out in any location where the copy of a constituent appears. Therefore, additional conditions have to be imposed on the deletion operation in order to constrain the application of the distributed deletion, blocking derivations such as (51)a-b:

- (51) a. *The ~~students~~ were arrested ~~the students~~.
 b. *~~The students~~ were arrested the ~~students~~.
 c. The students were arrested ~~the students~~.

Bošković (2005: 14)

3.3.2. Base-generation analyses

Base-generation analyses claim that the parts of a split noun phrase are merged independently of each other in different slots of the sentence (see Hale 1983, Jelinek 1984 and Fanselow 1988). To this end, different proposals have been made.

In the original version developed by Hale (1983) and Jelinek (1984) for Australian languages like Warlpiri, none of the parts of a discontinuous noun phrase figures as an argument in the sentence. Rather, the true argument is the (possibly phonologically empty) pronominal clitic on the predicate, while the discontinuous noun phrase parts are adjunct modifiers of this argument position.

In Fanselow's (1988) account of split topicalization in German, one part of the split noun phrase is merged as a verbal argument, while the other part might originate as a modifying adjunct and move to the topic position later or be generated there directly.

Under Kariaeva's (2009) approach to split constituents with an adjectival part in Ukrainian and Greek, the adjectival modifier is base-generated in the location in which it is spelled out, either as an adjunct inside the noun phrase (deriving a continuous noun phrase) or as an adjunct inside the VP (deriving a split noun phrase).

As far as the problems mentioned above are concerned, the advantages of a base-generation analysis are clear. First, the problem of the movement of a

non-constituent simply does not arise because discontinuity is not derived from movement. Second, imperfect splits, such as preposition doubling, can be derived by assuming that a preposition shows up in two PPs generated independently of each other. Third, morphological adjustments do not constitute a problem given that the different parts of the split can be generated with different morphologies.

However, this kind of analysis faces serious difficulties. The first difficulty concerns the fact that split noun phrases are sensitive to some island effects. As Ott (2009) notes, if it is true that split noun phrases are insensitive to certain island constraints (as is the case of the subjects of transitive verbs), it is also true that they respect other island types, such as the complex-NP constraint (see (52)), the adjunct-island condition (see (53)) and the coordinate-structure constraint (see (54)). Examples (52)-(54) are taken from Ott (2009).

- (52) * Bücher_i habe ich [_{NP} eine Geschichte dass sie [_{keine} t_i] liest] gehört
books have I a story that she no reads heard
 ‘I have heard a story that she does not read any books.’

- (53) * Bücher_i ist sie schon oft nachhause gegangen,[bevor sie [_{welche} t_i]
books is she already often home went before she some
gelesen hat]
read has
 ‘She often went home before reading some books.’

- (54) * Bücher hat sie bisher [_{nur wenige} t_i und Zeitschriften] gelesen
books has she so-far only few and magazines read
 ‘Thus far, she has only read few books and magazines.’

An additional argument that militates against the base-generation analysis is the preservation of the noun-phrase internal order (see van Riemsdijk 1989, cited in Ott 2009). In German, the adjective ordering illustrated in (55)a is unmarked, whereas the order in (55)b is only acceptable with a strong focal stress on the preposed adjective. Crucially, if the noun phrase *schnelle amerikanische Autos* is split-topicalized, the order among the adjectives must be preserved (see (56)). As Ott (2009) concludes, this means that in some sense the two parts of the split are merged together and then split apart.

(55) a. Hans mag schnelle amerikanische Autos
H. likes fast American cars

‘Hans likes American fast cars.’

b. ?? Hans mag amerikanische schnelle Autos
H. likes American fast cars

(Ott 2009: 67)

(56) a. [Amerikanische Autos]_i mag Hans nur [schnelle t_i]
American cars likes H. only fast

‘Hans only likes American fast cars.’

b. ??[Schnelle Autos]_i mag Hans nur [amerikanische t_i]
Fast cars likes Hans only American

(Ott 2009: 67)

Besides the objections already raised in the literature, additional objections may be brought against the base-generation analyses, including the following.

- Under the base-generation analyses, split noun phrases are not derived from continuous noun phrases. Instead, continuous and split noun phrases involve two different derivations, a fact that can be seen as a drawback for those who are committed to a transformational view of grammar.
- It is standardly assumed (at least in transformational-generative approaches to grammar) that topicalization, focalization and questions involve the movement of the fronted/preposed element. Because this constituent is typically the first part of the split, it seems intuitively unnatural to assume that this part of the split is base-generated in the spell out position.
- If the parts of the split noun phrase can be merged independently of each other, the question arises of how to constrain the merge positions of the fragments. At least for some languages, it seems clear that the parts of the split cannot be freely merged.
- If the parts of the split noun phrase can be merged independently of each other, the question arises of how to derive the semantic dependency (and, in some cases, the selectional relation) established between the split parts.

3.4. Summary

To conclude, split noun phrases have been intensively studied from a syntactic perspective and have been alternatively analyzed in terms of movement and in terms

of base generation. From the outline presented above, it is clear that none of the analyses is exempt from problems. This may be in part related to the intriguing properties of split noun phrases, which paradoxically show some signs of movement and some signs of non-movement. Another aspect that should be considered is that there may be non-trivial discrepancies among the different objects included under the term *split noun phrases*, a fact that may explain some of the “contradictory properties” reported in the literature.

4. The syntax of remnant-internal relativization: first take

In this section, I review some empirical evidence and theoretical analyses that might illuminate our understanding of split relativation, before turning to my own proposal in Section 5. After defining the concept of remnant-internal relativization (Section 4.1), I offer an overview of the phenomenon of split noun phrases and remnant-internal relativization in earlier stages of Portuguese (Section 4.2). Then, I examine the various possible accounts of the phenomenon, showing that the movement-based approach to splits, combined with the raising analysis of RRCs, is the combination that best handles the properties of remnant-internal relativization (Section 4.3). I conclude by presenting some drawbacks of simple movement analyses (Section 4.4).

4.1. The definition of *remnant-internal relativization*

To my knowledge, the term *remnant-internal relativization* has not been previously introduced in the literature. It is proposed here to describe relative constructions where the head noun and its modifier/complement appear discontinuously. This is illustrated schematically in (57). Example *a.*, which involves adjacency between the head noun and its modifier/complement, displays the regular word order (in head-initial relative clauses); example *b.*, which does not involve adjacency between these elements, displays a split noun phrase.

- (57) a. [S-matrix ... [N modifier/complement RRC] ...]
 b. [S-matrix ... [N [RRC ... modifier/complement ...]] ...]

Some concrete instances of this phenomenon are given in (58)-(60), from earlier stages of Portuguese (repeated from (5)-(7)). As can be observed below, the head and an element thematically dependent on it (either as a complement or as a modifier) appear discontinuously.

(58) Casos que Adamastor contou futuros
cases:MASC.PL that A. told future:MASC.PL

‘(the) future events that Adamastor foresaw’

(L. de Camões, *Os Lusíadas*, Book V, 60, 16th century, cited in Lausberg 1967/1972: §331)

(59) os livros que eu compus da philosophia
the books that I wrote of.the philosophy

‘the books of philosophy that I wrote’

(D. Pedro, *Livro dos Ofícios*, 15th century, in Pádua 1960: 70, cited in Martins 2004)

(60) e qualquer que de nos primeiro morer
and any that of us first die:FUT.SUBJ

‘and whoever of us first die’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1266)

The fact that relative clause constructions may generate split noun phrases has scarcely been noted in the literature. Pinkster (2005) reports that noun phrase discontinuity may arise in relativization constructions, as in (61), from Latin:

(61) mittit rogatum vasa ea quae pulcherrima apud eum viderat.
sent:3SG ask:PART vessels the which most.beautiful in him had.seen

‘He sent to ask for the loan of the most beautiful vessels he had seen at his house.’

(Cicero, *In Verrem* 4.63, cited in Pinkster 2005, glosses mine)

Some Latin grammars also mention this possibility. I quote:

The relative sometimes takes an adjective after it, which properly belongs to the antecedent.

(Zumpt 1832: 237)

The Relative Cause frequently attracts into itself an Adjective belonging to the antecedent, especially if that Adjective is a Superlative.

(Hale and Buck 1966: 157).

The lack of more studies reporting remnant-internal relativization may in part explain why the theoretical impact of this phenomenon remains unexplored. I only found one vague allusion to this fact in a footnote of Fanselow and Féry (2006), which I transcribe below:

In Old Occitan, relative clause formation leads to discontinuity (see Pinkster 2005). Depending on one's theory of relative clause formation, this construction (exemplified in (i)) would also involve a discontinuous noun phrase.

- (i) la justícia que grant áig a mandar (Old Occitan)
the legal.power which great I have to.dispose
 'The great legal power which I have at my disposal.'

(Fanselow and Féry 2006: 7)

4.2. Empirical evidence

This section presents new empirical evidence of great importance for the study of remnant-internal relativization. The data are mainly drawn from earlier stages of Portuguese, but evidence from Latin and French is also considered. Section 4.2.1 provides background, showing how remnant-internal relativization fits in with the more general phenomenon of split noun phrases attested in earlier stages of Portuguese. Then, Section 4.2.2 focuses on the core properties of remnant-internal relativization.

4.2.1. Split noun phrases in earlier stages of Portuguese

Pinkster (2005) and Devine and Stephens (2006), among others, report that split noun phrases¹² are frequently attested from the earliest Latin texts until Late Latin.¹³ This is illustrated, for instance, in the contrast given in (62): in (62)a, the phrase *legiones*

¹² These authors use the term *hyperbaton* to designate split XPs that were otherwise treated as a unit with respect to Case theory and θ -theory (see Mathieu and Sitaridou 2005). This term comes from a combination of *hyper* and the aorist participle of *baino* 'to come', which means 'passed over' (Mathieu 2004). It is primarily conceived as a rhetorical device that uses the inversion of the normal word order to produce a discourse effect.

¹³ In Latin, the operation of "splitting" arises not only in noun phrases but also in such constituents as prepositional phrases and conjuncts. Given the limited scope of this study, I will focus only on the occurrence of split noun phrases here, paying special attention to those involving discontinuity between the noun and a post-nominal modifier or complement.

novas ‘two legions’ is continuous, whereas in (62)b, the phrase *duas legiones novas* ‘two new legions’ is split into two parts.

(62) a. *Facite hoc meum consilium legiones novas non improbare*
Suppose this my policy legions new non reject

‘Suppose that the new legions do not reject my policy’

(Cicero, *Philippics* 12.29, cited in Devine and Stephens 2006: 531)

b. *Caesar duas legiones inciteriore Gallia novas conscripsit*
Caesar two legions in hither Gaul new enrolled

‘Caesar enrolled two new legions in Hither Gaul’

(Caesar, *De Bello Gallico* 2.2, cited in Devine and Stephens 2006: 531)

According to Devine and Stephens (2006: 531), the split noun phrase in (62)b arises in the context of scrambling: *duas legiones* ‘two legions’ has been scrambled, leaving the modifier *novas* ‘new’ in a focus position. The same phenomenon can be found in sentences involving topicalization, as in (63); here, the noun *vinaceos* ‘dregs’ has raised to a topic position to the left of the frequency adverb *cotidie* ‘daily’, stranding the adjective *recentis* ‘fresh’.

(63) *Vinaceos cotidie recentis succernito*
dregs daily fresh sift:IMP.FUT.2SG

‘Sift the fresh dregs daily.’

(Cato, *De Agri Cultura* 25.1, cited in Devine and Stephens 2006: 531)

Interestingly, Devine and Stephens (2006) also report the occurrence of split noun phrases in interrogative *wh*-contexts. This is illustrated in (64). Here the *wh*-expression *quod supplicium* ‘what punishment’ is split into two parts: the *wh*-pronoun occurs in the left periphery of the clause, and the noun *supplicium* ‘punishment’ appears in the rightmost sentential position.

(64) *quod tandem excogitabitur in eum supplicium...?*
what then will.be.thought.up in him punishment

‘What punishment, I ask you, will be thought up for the man...?’

(Cicero, *Pro Rabirio Perduellionis Reo* 28, cited in Devine and Stephens 2006: 584)

Although split noun phrases are frequently attested from the earliest until Late Latin texts, this situation changed quite drastically in the development from Latin to

Romanc languages. Western European Romance languages are often characterized as not allowing split noun phrases or by allowing them only in a very restricted way (see Pinkster 2005, Fanselow and Féry 2006).

Importantly, some exceptions to this generalization have been reported in the literature. Pinkster (2005) refers to the occurrence of split noun phrases in Old French (see (65)). Butler and Mathieu (2004) take the French construction in (66) (repeated from (2))¹⁴ as involving a split noun phrase.

(65) la hautece i sera tote de mon empire

‘All the high-ranking people of my empire will be there’

(Dole 3486, Jensen 1990: 543, cited in Pinkster 2005: 5)

(66) Combien as-tu lu de livres?

how.many have.you read of books

‘How many books have you read?’

(Canonical: ‘Combien de livres as-tu lus?’)

(Butler and Mathieu 2004: 2)

Similar facts have been documented for earlier stages of Portuguese. Martins (2004) shows that split noun phrases are attested in earlier stages of Portuguese.¹⁵ Some of the examples cited in her paper are given in (67)-(69).

(67) Noticia fecit pelagio romeu de fiadores
notice made Pelagio Romeu of guarantors

‘Pelagio Romeu made a notice of guarantors’

(12th century, Martins 1999)

(68) Boscadas as rrazoões dos que livros fezerom desta estoria
found the reasons of the ones that books did of.this story

‘Found the reasons of the ones that made books of this story ... (Once the reasons of the ones that wrote this story are found...)’

(15th century, Fernão Lopes, Cr. de D. João I, Parte I, cap. VIII, cited in Martins 2004)

¹⁴ For a different analysis of the French construction in (66), see Fanselow and Féry (2006).

¹⁵ To be more precise, Martins (2004) also reports the possibility of finding split noun phrases in CEP under some restricted constructions. However, as will become clear in Section 5.3, the split noun phrases found in CEP are much more restricted than are those found in earlier stages of Portuguese and, as a consequence, are not crucial for the argument developed in this study.

- (69) Em que nos mostra esta regra que fame ham da
in that us:CL shows this rule that hunger have of.the
palavra de Deus aquelas que desejam de a ouujr.
word of God those that want of it:CL listen:INF

‘This rule shows us that those who want to listen to God’s word have hunger for it.’

(Regra de Santo Agostinho. *In* Pádua 1960, p. 159, cited in Martins 2004)

In (67)-(69) the split noun phrases correspond to the sequence: *head noun ... PP*, but this is not necessarily so. The head noun in the first split-part may be associated with other elements, such as an adjective (see (70)) or a determiner (see (71)). Moreover, the constituent in the second split-part can be a PP, as in (67)-(70), or an adjective, as in (71).

- (70) diz que se deus o matar de fome q ma bocado a deus
says that if god him:CL kills of hunger that bad piece have god
 de comer dele
to eat:INF of.him

‘[He] says that if God kills him by hunger, God will eat a bad piece of him (= God will go through a hard time with him).’

(17th century, CARDS4002)

- (71) Outros fauores se lhe tem feito extraordinarios.
other favors SE:CL him:CL has done extraordinary

‘Other extraordinary favors have been done for them.’

(17th century, Coelho 1987: 192)

Just as was observed for Latin, split noun phrases in earlier stages of Portuguese can also arise in interrogative *wh*-movement contexts, as illustrated below.

- (72) a. que origem lhe havemos de dar mais nobre?
what origin to.him:CL have:IPL of give:INF more noble

‘What more noble origin shall we give him?’

(TYC; Matias Aires, *Reflexão sobre a Vaidade dos Homens*; 18th century)

- b. que susto ou que dano nos pode vir maior?
what fright or what damage to.us:CL can come:INF bigger

‘What bigger fright or damage could be done to us?’

(TYC; José da Cunha Brochado, *Cartas J. C. Brochado*; 1696-1703)

- (73) a. Quantas castas há de nomes?
how.many types has of nouns
 ‘How many types of nouns are there?’
 (TYC; Jerónimo Contador de Argote, *Regras da Lingua portugueza*; 1725)

- b. Quantas figuras há de Dicção?
how.many figures has of diction
 ‘How many figures of diction are there?’
 (TYC; Jerónimo Contador de Argote, *Regras da Lingua portugueza*; 1725)

4.2.2. Remnant-internal relativization in earlier stages of Portuguese

This section demonstrates that relativization combines with other constructions (such as topicalization, scrambling, focalization, and interrogative *wh*-movement) in generating split noun phrases. By way of illustration, see sentence (74), from Latin, where the head and its adjectival modifier appear discontinuously.

- (74) Inter jocos quos inconditos jaciunt
amidst jests which rude utter:3PL
 ‘Amidst the rude jests which they utter.’
 (Titius Livius, cited in Zumpt 1832: 237)

Instances of remnant-internal relativization are also attested in earlier stages of Portuguese, as illustrated in (75):

- (75) Casos que Adamastor contou futuros
cases that A. told future
 ‘(the) future events that Adamastor foresaw’
 (L. de Camões, *Os Lusíadas*, Book V, 60, 16th century, cited in Lausberg 1967/1972: §331)

On the basis of the data from earlier stages of Portuguese inspected thus far, the core properties of remnant-internal relativization can be described as follows.

I. Remnant-internal relativization involves post-nominal relative clauses (i.e., relative clauses with the *head + relative clause* order).¹⁶

II. Remnant-internal relativization involves bipartite discontinuity.

III. Remnant-internal relativization involves pull-splits, i.e., splits that retain the order of elements found in the continuous counterpart (see Fanselow and Cavar 2002). This is illustrated in (76):¹⁷

(76) a. os livros que eu compus da philosophia
 the books that I wrote of.the philosophy

 b. os livros da philosophia que eu compus
 the books of.the philosophy that I wrote

IV. In the first split-part, the head noun may appear alone (see (75)) or associated with other elements, namely a definite article (see (77)), an indefinite article (see (78)) or an adjectival modifier (see (79)).

(77) os livros que eu compus da philosophia
 the books that I wrote of.the philosophy

 ‘the books of philosophy that I wrote’

 (D. Pedro, *Livro dos Ofícios*, 15th century, in Pádua 1960: 70, cited in Martins 2004)

(78) e pasarã huũ rrio que perhy core dagoa doce
 and they.crossed a river that through.there flows of.water sweet

 ‘and they crossed a river of sweet water that flows through there’

 (*Carta de Pero Vaz de Caminha*, fol. 3v., 15th century, cited in Martins 2004)

¹⁶ As indicated in Section 2.3 (point G.), relative clauses can be syntactically classified as post-nominal (or head-initial); pre-nominal (or head-final); or circum-nominal (or head-internal). Because what underlies such a classification is the spell out position of the head, remnant-internal relativization might be taken as involving post-nominal relatives.

¹⁷ The sequence given in (76)b is not attested in Old Portuguese texts with the exact words that parallel the example (76)a. However, because the construction is well attested in all periods of the history of Portuguese, (76)b was constructed to show the relevant contrast.

- (79) *da qual cousa ellas dizem que som hisentas e que nũa*
from.the whichthing they say that are free and that never
a pagarõ per príuelegio antigo que tẽem do papa
it.AC paid:3PL by privilege old that have:3PL from.the pope
 ‘... as for it, they say that they never paid it because they have an old privilege from the pope.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1426)

V. The second split-part can be an adjectival modifier (see (74)-(75)) or a PP (see (80)-(82)). The PPs can be either modifiers (see (80)) or complements of the noun (as in (81) and (82)).¹⁸

- (80) *e pasarã huũ rrio que perhy core dagoa doce*
and they.crossed a river that through.there flows of.water sweet
 ‘and they crossed a river of sweet water that flows through there’
 (*Carta de Pero Vaz de Caminha*, fol. 3v., 15th century, cited in Martins 2004)

- (81) *eram sobrinhas da molher que faleseo de lamsarote rodrigues*
they.were nieces of.the wife that died of L. R.
 ‘they were nieces of the wife of Lamsarote Rodrigues who died’
 (Coelho 1987; *Inquisição de Évora: dos primórdios a 1668*; 1612)

- (82) *como eu me encontro num estado miseravel pella falta que*
as I myself:CL am in.a state miserable by.the lack that
há do vinho
has of.the wine
 ‘...as I am in a miserable state by the lack of wine that is there (= because of the wine shortage)’
 (CARDS 4002, year 1890)

VI. The second split-part may surface in the rightmost clausal position (as in (80)-(82)) or in a non-final position following the relativizer, as in (83)-(87).

¹⁸ In (81), *molher* ‘wife’ is a relational name, and in (82), *falta* ‘lack’ is a deverbal noun derived from the verb *faltar* ‘to lack’.

- (83) e qualquer que de nos primeiro morer
and any that of us first die:FUT.SUBJ
 ‘and whoever of us first die’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1266)
- (84) e esto por prool e verdade de hũa Licença
and this by favor and truth of a license
 que do dito senhor pera ello tenho
that from.the mentioned man for that have:1SG
 ‘... and (I wrote this document) under the benefit and truth of a license from the
 aforementioned man that I have to (make) it.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)
- (85) e o deradeiro que delas fiquar posa amte
and the last that of.them stay:FUT.SUBJ can:PRES.SUBJ before
 de sua morte nomear a terceira
of his death appoint:INF the third
 ‘and the last of them that stays alive can appoint the third (person) before his death.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1528)
- (86) Nas bombas que de fogo estão queimando
in.the bombs that of fire are burning
 ‘ [the Cyclops’ art is shown] in the bombs of fire that (they) are burning’
 (Ramos 1982; L. de Camões, *Os Lusíadas*, Book V, 90, 16th century)
- (87) quando alevantarão Hum por seu capitão, que peregrino
when chose a by their captain that foreign
 Fingio na cerva espirito divino
pretended in.the doe spirit divine
 ‘when they choose for captain a foreign (person) who pretended that there was divine
 spirit in his doe’
 (L. de Camões, *Os Lusíadas*, Book I, 26, 16th century, cited in Dias 1933/1970: §462)

4.3. Competing analyses

Sections 2.2 and 3.3 have outlined the most salient analyses proposed in the literature to account for RRCs and split noun phrases, which are summarized in Table 1:

Table 1. Competing analyses of RRCs and split noun phrases

	Competing analyses
RRCs	raising analysis adjunction analysis
Split noun phrases	movement analysis base-generation analysis

Building on these approaches, the present section aims to identify the type of syntactic analysis that best conforms to the properties of remnant-internal relativization. If we take remnant-internal relativization to involve both an RRC and a split noun phrase, it emerges from Table 1 that there are four logically possible combinations:

- Adjunction analysis of RRCs + movement analysis of split noun phrases.
- Adjunction analysis of RRCs + base generation analysis of split noun phrases.
- Raising analysis of RRCs + movement analysis of split noun phrases
- Raising analysis of RRCs + base generation analysis of split noun phrases

In what follows, I will examine the four hypotheses and show that the combination *raising analysis of RRCs + movement analysis of split noun phrases* best captures the properties of remnant-internal relativization in earlier stages of Portuguese.

A. Adjunction analysis of RRCs + Movement analyses of split noun phrases

Under the adjunction analysis, the antecedent is base-generated in a relative clause-external position. If we combine the adjunction analysis with the movement analysis of split noun phrases, we have to postulate that the head noun and its modifier/complement are base-generated in a relative clause-external position and that the split noun phrase is derived via the rightward movement of the modifier/complement (which ends up right-adjoined to the DP node), as in (88):

$$(88) [_{DP} [_{NP} [_{NP} \text{head } t_{\text{modifier/complement}}] [_{CP} \text{RRC}]] \text{modifier/complement}]$$

A concrete example is given in (89), where the PP *da filosofia* ‘of philosophy’ is taken to undergo rightward movement to a DP-right-adjoined position.

- (89) [DP OS [NP [NP livros t_k] [CP Op_j [C' que [IP eu_i [I' compus_i [VP t_i [V' t_i t_j]]]]]]]]]]
the books that I wrote
 [PP da philosophia]_k
of.the philosophy

Although the combination *adjunction analysis of RRCs + movement analysis of split noun phrases* accounts for the sentences with the modifier/complement in the rightmost position, it fails to account for the contexts involving a non-final modifier/argument, as in (90) (repeated from (83)).

- (90) e qualquer que de nos primeiro morer
and any that of us first die:FUT.SUBJ
 ‘and whoever of us first die’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1266)

The pattern in (90) could be interpreted as resulting from rightward movement of the PP *de nos* ‘of us’, placing it between the relativizer *que* ‘that’ and the adverb *primeiro* ‘first’. Apart from several problems that this derivation raises, the strongest objection is that it would involve lowering to a non-c-commanding position (see Fiengo 1977) and hence should be rejected in view of such requirements as the Proper Binding Condition or the Empty Category Principle.

B. *Adjunction analysis of RRCs + Base-generation analysis of split noun phrases*

Under this scenario the head noun and its modifier/complement are generated separately in two different syntactic positions: the head noun is merged CP-externally, whereas the modifier/complement is merged CP-internally.

The basic assumption underlying this proposal is that adjuncts/arguments may not have their source inside the noun phrase with which they are associated. This hypothesis has been proposed for adjectives by Baker (2003) and for PPs by Koster (1987), among others.

Baker (2003) proposes that adjectives can be generated in any syntactic position (such as VP) as long as that position permits free adjunction. Koster (1987: 197) claims that PPs introduced by *of* do not necessarily have their source inside a noun phrase. According to him, if this were always the case, (91)a would have an ungrammatical source, as shown in (91)b.

- (91) a. *Of the students in the class*; I like [Mary t_i] better than anyone else
 b. * I like [Mary of the students in the class] better than anyone else
 (Koster 1987: 197)

There are at least three objections that can be raised against analyzing remnant-internal relativization along these lines. First, in contrast to Koster's example (see (91)), when remnant-internal relativization is involved, the head and the PP could have a grammatical source, as shown in (92)b (repeated from (76)).

- (92) a. os livros que eu compus da philosaphia
 the books that I wrote of the philosophy
 b. os livros da philosaphia que eu compus
 the books of the philosophy that I wrote

Second, the idea that the second split-parts function as event modifiers (being adjoined to the VP) rather than noun modifiers is not plausible. Although some modifiers can be related to entities and events, not all have this ability. For instance, in (92), if the book of philosophy was written by me, it is the book that is of philosophy and not the event of writing. Thus, the modifier is clearly interpreted in relation to the noun (and not in relation to the event).

Finally, it seems reasonable to assume that the semantic dependency between the head and its modifier/complement require these elements to be in a structural relation at some point at the derivation. This cannot be achieved, however, under the *adjunction analysis of RRCs + base-generation analysis of split noun phrases* hypothesis, because a modifier/complement merged in a relative clause-internal position cannot be structurally related to a head in a relative clause-external position.

C. *Raising analysis of RRCs + Movement analyses of split noun phrases*

This combination provides us with a natural explanation of why the head and the modifier/complement can split. In a nutshell, let us assume that the head and its modifier/complement are merged in an RRC-internal position and that head

movement optionally involves pied-piping.¹⁹ If so, the head and its modifier/complement appear consecutively (see (93)a); if it does not, the modifier/complement is stranded and the noun phrase splits into two parts (see (93)b).

- (93) a. [_{DP} OS [_{CP} [livros_k da philosophia]_k [_C que [_{IP} eu_j [_{I'} compus_i [_{VP} t_j [_{V'} t_i t_k]]]]]]]]]
 the books of.the philosophy that I wrote
- b. [_{DP} OS [_{CP} livros_k [_C que [_{IP} eu_j [_{I'} compus_i [_{VP} t_j [_{V'} t_i [t_k da philosophia]]]]]]]]]
 the books that I wrote of.the philosophy

Although this simple story needs to be qualified, it offers us a good starting point for a more elaborate analysis of remnant-internal relativization.

D. Raising analysis of RRCs + Base-generation analysis of split noun phrases

As already mentioned, base-generation approaches assume that the members of the split do not map onto a phrasal constituent at any point of the derivation. There are four logical ways of deriving remnant-internal relativization along these lines:²⁰ (i) the two parts of the split are merged in the spell out position; (ii) the two parts of the split are not merged in the spell out position; (iii) only the first part of the split is merged in the spell out position; or (iv) only the second part of the split is merged in the spell out position.

Hypotheses (i) and (iii) are incompatible with the raising analysis of RRCs. While the raising analysis of RRCs requires the head to be base-generated in the relativization site, the base-generation analyses of split noun phrases in (i) and (iii) require the head to be generated in [Spec, CP].

Hypotheses (ii) and (iv) are at first sight compatible with the raising analysis of RRCs as they assume that the spell out position of the head differs from its base position. There are, however, good reasons to doubt that these hypotheses are on the

¹⁹ The notion of pied-piping was first introduced by Ross (1967). The term refers to a phenomenon whereby a particular movement operation, designated to displace an element X, in fact displaces a larger phrase in which X is embedded. This is the case of (93)a, where the movement operation designated to displace the head *livros* ‘books’, in fact displaces the larger constituent *livros da philosophia* ‘books of philosophy’.

²⁰ Here, the discussion is confined to bipartite remnant-internal relativization, which is the only type attested in the data inspected in my research.

right track. As already discussed in point B. above, in order to be interpreted in relation to the head, the modifier/complement has to be in a structural relation with the head at some point of the derivation, and this requirement cannot be fulfilled if the head and the modifier/complement are not merged together.

4.4. Problems with the *raising analysis of RRCs + movement analysis of split noun phrases combination*

On the basis of the discussion in the preceding section, it is possible to conclude that the *Raising analysis of RRCs + Movement analyses of split noun phrases combination* is much better equipped to handle remnant-internal relativization in general than the other theories are.

However, as already mentioned, the simple movement approach to remnant-internal relativization needs to be improved because, as it stands, it cannot derive the following properties of remnant-internal relativization.

I. Internal structure of the first split-part

The first split-part may contain just the head (see (94), repeated from (5)) or be associated with other elements, such as an adjectival modifier (see (95), repeated from (79)).

(94) Casos que Adamastor contou futuros
cases that A. told future

‘(the) future events that Adamastor foresaw’

(L. de Camões, *Os Lusíadas*, Book V, 60, 16th century, cited in Lausberg 1967/1972: §331)

(95) da qual cousa ellas dizem que som hisentas e que nũa
from.the which thing they say that are free and that never
 a pagarõ per príuelegio antigo que tõem do papa
it:AC paid:3PL by privilege old that have:3PL from.the pope

‘... as for it, they say that they never paid it because they have an old privilege from the pope.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1426)

This property is not surprising at all. In Latin and in earlier stages of Portuguese, “simple” split noun phrases also allow for it. See (96)-(98) (repeated from above),

where the first split-part exhibits the head combined with other elements, namely a quantifier (96), a determiner (97) and an adjective (98).

- (96) Caesar duas legiones inciteriore Gallia novas conscripsit
Caesar two legions in/hither Gaul new enrolled

‘Caesar enrolled two new legions in Hither Gaul’

(Caesar, *De Bello Gallico* 2.2, cited in Devine and Stephens 2006: 531)

- (97) Outros fauores se lhe tem feito extraordinarios.
other favours SE:CL him:CL has done extraordinary

‘Other extraordinary favours have been done to them.’

(17th century, Coelho 1987: 192)

- (98) diz que se deus o matar de fome q ma bocado a
says that if god him:CL kills of hunger that bad piece have

deus de comer dele
god to eat:INF of.him

‘[He] says that if God kills him by hunger, God will eat a bad piece of him (= God will go through a hard time with him).’

(17th century, CARDS4002)

These complex first split-parts are problematic for simple movement approaches because they seem to involve non-constituent movement. Take, for instance, sentence (95) above; if we assume that the PP originates within the DP in a structure like [DP D [NP N PP]] and that the adjectival modifier also originates within the DP as a specifier of a functional projection (see, e.g., Cinque 1994), then the problem is that there is no constituent that includes the head and the adjective but excludes the PP and that can undergo leftward movement, displaying remnant-internal relativization.

II. The non final-position of the modifier/complement

As noted above, the simple movement analysis is not able to derive the non-final position of the modifier/complement in sentences like (99) (repeated from above):

- (99) e esto por prool e verdade de hũa Licença
and this by favor and truth of a license

que do dito senhor pera ello tenho
that from.the mentioned man for that have:1SG

‘... and (I wrote this document) under the benefit and truth of a license from the aforementioned man that I have to (make) it.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)

Here, the PP *do dito senhor* ‘from the aforementioned man’ does not appear in the final sentential position, but instead in the left periphery of the embedded clause. If remnant-internal relativization is taken to involve head movement to [Spec, CP] and the stranding of the modifier/complement, it remains a mystery why, in sentences like (99), the PP is not placed in the rightmost clausal position.

4.5. Summary

This section was devoted to the discussion of remnant-internal relativization, the major focus of this chapter. It began with a comprehensive description of split noun phrases and remnant-internal relativization in earlier stages of Portuguese (with some references to Latin). Then the most important competing theories that could, in principle, account for remnant-internal relativization were systematically compared. Four main approaches were outlined, based on the combination raising/adjunction analysis of RRCs and movement/base-generation analysis of split noun phrases. The empirical and theoretical evaluation of these proposals showed that the combination *raising analysis of RRCs + movement-based approach to splits* best captures the properties of remnant-internal relativization in earlier stages of Portuguese. At the end, it was shown that even this combination is not completely satisfactory as it can derive neither the internal structure of the first split-part nor the non-final position of the modifier/complement.

5. The proposal

Building on Fanselow and Ćavar’s (2002) proposal for discontinuous phrases (see Section 3.3.1), I submit that the problems raised in Section 4.4 can be circumvented if we take remnant-internal relativization as an effect of phonological deletion. As I will show, this hypothesis gains strength when considered in the light of the insights of Bošković and Nunes (2007) on the copy theory of movement.

However, before going on to the analysis of remnant-internal relativization, it is useful to make a short excursus on the following topics: (i) the interpretation of remnant-internal relativization (Section 5.1.1) and (ii) the version of the copy theory of movement proposed by Bošković and Nunes (2007) (Section 5.1.2). With this

background in mind, Section 5.2 provides a step-by-step analysis of remnant-internal relativization in earlier stages of Portuguese. Then, Section 5.3 briefly discusses the contrast between remnant-internal relativization in CEP and in earlier stages of Portuguese.

5.1. Preliminaries

5.1.1. The interpretation of remnant-internal relativization

It has been observed in the literature that the members of split noun phrases differ in their information structure status, a property that Predolac (2009) refers to as *split information structure*.

The same seems to be true of the instances of remnant-internal relativization attested in earlier stages of Portuguese. As will become clear below, in neutral declarative sentences displaying broad information focus, remnant-internal relativization emerges when the modifier/complement is assigned emphatic/contrastive focus.

This interpretation is available if the second split-part appears in either a non-final or a final clausal position. The first hypothesis is instantiated in (100), where the modifier/complement surfaces in a non-final position, immediately after the complementizer. In this case, *do dito senhor* ‘of the aforementioned man’ is interpreted as an identificational focus (in the sense of Kiss 1998) as it presupposes a set of relevant entities for which the predicate can hold and exhaustively identifies the proper subset of this set for which the predicate actually holds. (100) can thus be paraphrased as in (101).

- (100) e esto por prool e verdade de hũa Licença
and this by favor and truth of a license
 que do dito senhor pera ello tenho
that from.the mentioned man for that have:1SG

‘... and (I wrote this document) under the benefit and truth of a license from the aforementioned man that I have to (make) it.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)

- (101) Of a set of relevant licenses it is true for the license from the aforementioned man and no other else that I have it.

If the second split-part surfaces in the rightmost position, various meaning facets of emphatic/contrastive focus related to different syntactic environments may be available. The notion of contrastiveness may be expressed in overtly contrastive statements, as in (102). In this case, the focused constituent (*livros da philosophia* ‘(books) of philosophy’ explicitly contrasts with (*obras minhas que som fundadas sobre bem falar* ‘(writings of mine) on the art of speaking’. It is clear, then, that the paragraph context explicitly indicates the existence of a contextually salient set of alternatives.

- (102) Por a qual cousa aficadamente te amoesto, meu Ciceram, que nom soamente aquelas obras minhas que som fundadas sobre bem falar, mas ainda

os livros que eu compus da philosophia,
the books that I wrote of.the philosophy

que som ja iguaaes a elas, tu os leas com boa deligencia.

‘This is why I strongly urge you, my dear Cicero, to read with care not only my writings on the art of speaking well, but also the books that I wrote on philosophy, which are now about as extensive.’

(Piel 1948; D. Pedro, *Livro dos Officios*; 15th century)

However, just as in example (100) above, the emphatic/contrastive focus in the rightmost position can be interpreted as an identificational focus. The exhaustive interpretation associated with *das suas qujntãas e casaaes* ‘of their farms and hamlets’ in (103) is evident upon the paraphrase in (104).

- (103) E aos prazos que as Egreias e Moesteiros
and to.the contracts that the churches and monasteries

qujserem ffazer das suas qujntãas e casaaes
want:FUT.SUBJ make:INF of.the their farms and hamlets

‘(and give authority) to the contracts that the churches and the monasteries may make on their farms and hamlets.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1441)

- (104) Of the set of the relevant things that a church or a monastery can contract, it should be given legal authority to the contracts that they make on their farms and hamlets, and nothing else.

Moreover, the emphatic/contrastive focus may simply signal the speaker-oriented emphasis, i.e., the relative weight that the speaker/writer wants to attach to a particular element in the sentence (see Enkvist 1980: 135). In this sense, it adds a

surplus value to the interpretation, “(...) not at the level of the proposition, but of speech modality or a metalinguistic level, where information stemming from the speaker coordinates becomes relevant” (Remberger 2010). A case in point is provided in (105), where the writer emphasizes the sort of “shortage” that leads him to such a miserable state without referring (implicitly or explicitly) to any other type of “shortage”.

- (105) como eu me encontro num estado miseravel pella falta que
 as I myself:CL am in.a state miserable by.the lack that
 há do vinho
 has of.the wine
 ‘... as I am in a miserable state by the lack of wine that is there (= because of the wine shortage)’

(CARDS 3017, year 1890)

The fact that the emphatic/contrastive focused element may occur either in a final or in a non-final position is not surprising, considering that CEP also displays different strategies of emphatic/contrastive focus marking (cf. Chapter 1, Section 4.3.1). Hence, I take earlier stages of Portuguese to be like CEP in that emphatic/contrastive focused constituents can be marked syntactically through movement of the focused element to the left periphery, as proposed by Martins (1994). This explains why the second split-part in (100) surfaces in a non-final position, after the relativizer *que* ‘that’.²¹ I furthermore assume that earlier stages of Portuguese pattern with CEP in the possibility of marking emphatic/contrastive focus only prosodically. This explains why, in examples (102)-(105), the second split-part surfaces in a final position: it receives stress prominence in the rightmost position via the NSR.

²¹ As will become clear in Section 5.3 below, there is one important difference between CEP and earlier stages of Portuguese. In earlier stages of Portuguese, an emphatic/contrastive focused constituent may surface in a focus-dedicated position in the left periphery of both a dependent and a non-dependent clause, whereas in CEP a focus-dedicated position is only available in non-dependent clauses. This is in accordance with Martins (1994, 2001), who shows that the structures of preposed contrastive foci were fully productive in the history of Portuguese until the 17th century, when their frequency began decreasing significantly. Note in addition that, according to Zubizarreta (1999), Spanish still allows for preposed contrastive focus in embedded domains.

5.1.2. The copy theory of movement in the PF side

This section briefly presents two approaches that very important to the analysis of remnant-internal relativization. The first is the version of the copy theory of movement proposed by Bošković and Nunes (2007). The second is the analysis of post-verbal subjects in Serbo-Croatian put forth by Stjepanović (2007). In presenting these approaches, I pay particular attention to the theoretical devices used to account for the syntax-phonology interaction.

5.1.2.1 Bošković and Nunes (2007)

Following Chomsky's (1995) approach to movement, where the raising of elements leaves copies behind that are subsequently deleted, Bošković and Nunes (2007) (based on previous work by Bošković 2001, 2002, 2004a,b and Nunes 1999, 2004) argue that *traces* (i.e., copies that are structurally lower in the syntactic representation) may be phonetically realized as the result of a copy deletion mechanism that works at PF.

The basic idea is that PF has a preference for pronouncing the highest copy of a chain, but a lower copy may be pronounced to avoid a PF violation. Technically, this statement requires further clarification.

The first point in need of clarification concerns the preference for deleting lower copies. The explanation Nunes (1999, 2004) provides for this fact is based on an economy principle that prefers fewer applications of deletion in later computations of the phonological component. As J. Nunes puts it:

Exploring the null hypothesis regarding the copy theory of movement, the above proposal thus takes the position that both heads of chains and traces should in principle be subject to phonetic realization. According to the logic of the proposal, there is nothing intrinsic to lower copies that prevents them from being pronounced. If Chain Reduction [i.e. deletion of constituents of a non-trivial chain, AC] proceeds in such a way that only a trace survives, the derivation may eventually converge at PF. The fact that in most cases such a derivation yields unacceptable sentences is taken to follow from *economy* considerations, rather than convergence at PF. Since the highest chain link is engaged in more checking relations, it will require fewer application of F[ormal]F[eature]-Elimination than lower chain links, thereby being the optimal candidate to survive Chain Reduction and be phonetically realized, all things being equal.

(Nunes 2004: 33)

This brings us to the second point needing clarification: why the deletion of a lower copy in PF is just a preference and not the only option. The reason is clear: a lower copy can be produced to avoid a PF violation. Some of the PF factors that have been considered to induce the pronunciation of lower copies are stress assignment processes, intonational requirements and morphological restrictions on identical elements.

However, PF constraints may also block full copy deletion. In this case, a last-resort mechanism is admitted on the PF side: the so-called *scattered deletion*. It consists of the deletion of different pieces of different copies, as represented in (106):

(106) [X ~~Y~~ⁱ ... [~~X~~ Y]ⁱ]

Scattered deletion has been successfully applied to a range of languages and phenomena, namely to cliticization in Bulgarian and Macedonian (Bošković 2001), participle-auxiliary order in Bulgarian (Lambova 2004) and split phrases (Fanselow and Ćavar 2002).

5.1.2.2 Stjepanović (2007)

Stjepanović (2007) proposes that the adoption of the copy theory of movement proposed by Bošković and Nunes (2007) provides a principled account of several patterns of word order in Serbo-Croatian. Under this account, post-verbal subjects are recast as one of the possible effects of copy deletion on the PF side. The most relevant aspects of this proposal are outlined below.

Serbo-Croatian is characterized by great freedom of word order. Depending on the information structure involved, sentences containing a new information focus can exhibit the following word orders: SVO, SOV, OSV and OVS. For the sake of illustration, consider the OVS word order in (107)b, which can be produced as an answer to the question in (107)a:

- (107) a. [Who is catching the mouse?]
 b. Miša hvata mačka
mouse catches cat
 ‘A cat is catching a mouse.’

(Stjepanović 2007: 235)

Stjepanović (2007) shows that in sentences with the neutral intonation pattern, the constituent bearing new information focus (such as *mačka* ‘cat’ in (107)b) follows elements that represent old information. In order to explain the final sentential position of focalized elements, the author proposes that (i) the subject moves in overt syntax even in the cases where it surfaces post-verbally; (ii) the subject surfaces post-verbally because a lower copy of the chain is pronounced in PF.

The factor that Stjepanović takes to induce the pronunciation of the lower copy is sentential stress assignment. The basic tenets of her analysis are as follows.

- In the output of syntax, sentences have a focus structure whereby each element is associated with an [F]-feature. For example, (107)b would have the syntactic output in (108) (taken from Stjepanović 2007: 236):

(108) [_{AgrSP} miša [_{AgrSP} mačka [_{TP} mačka hvata [_{AgrOP} miša hvata [_{V1} mačka
 [-F] [+F] [+F] [-F] [-F] [-F] [+F]
 hvata [_{V2} hvata miša]]]]]]]]
 [-F] [-F] [-F]

- In Serbo-Croatian, the NSR applies just after spell out, assigning prominence to the rightmost/lowest sentential constituent (Zubizarreta 1998, 1999).
- If the element that receives the nuclear stress has an [+F]-feature, no problem arises.
- If the rightmost element is [-F], a conflict situation emerges between the NSR and the FPR. To resolve this conflict, Serbo-Croatian renders defocalized [-F] elements extrametrical for the application of the NSR.
- A lower copy must be pronounced if it carries the nuclear neutral stress. If this PF requirement is not satisfied, the derivation does not converge.

In summary, post-verbal subjects in Serbo-Croatian (as in (107)b) are a result of lower copy pronunciation. This is due to the requirements on sentential stress assignment, which force the copy associated with the nuclear stress to be pronounced.

5.2. The derivation of remnant-internal relativization

The analysis of remnant-internal relativization presented in this section is based on three central claims: (i) remnant-internal relativization constructions and regular

relative constructions are derived from a continuous noun phrase; (ii) remnant-internal relativization has the function of focus-marking the second split-part (with emphatic/contrastive focus); and (iii) non-adjacency between the head and its modifier/complement is determined by conditions of the phonological component (and not of syntactic movement per se). Following Bošković and Nunes (2007) and Stjepanović (2007), it is argued that copies (or parts of the copies) that are structurally lower in the syntactic representation can be phonetically realized to satisfy PF requirements on stress assignment.

To start with, let us consider the distributional patterns that need to be derived from the present analysis, which will be discussed in turn.

Pattern I. Regular relativization

The head and its modifier/complement are adjacent. See (109):

- (109) *os livros da philosophia que eu compus*
the books of.the philosophy that I wrote

Pattern II. Remnant-internal relativization (with a final modifier/complement)

The head and its modifier/complement are not adjacent; the modifier/complement surfaces in the rightmost sentential position. See (110):

- (110) *os livros que eu compus da philosophia*
the books that I wrote of.the philosophy

Pattern III. Remnant-internal relativization (with a non-final modifier/complement)

The head and its modifier/complement are not adjacent; the modifier/complement surfaces in a non-final position, following the relativizer. See (111):

- (111) *os livros que da philosophia eu compus*
the books that of.the philosophy I wrote

5.2.1. Pattern I

Along the lines of the copy theory of movement, let us assume that the head and its modifier/complement are merged together in the relativization site. Then, in order to

satisfy the *wh*-feature on C° , the entire noun phrase is copied and merged in [Spec, CP], leaving a copy behind:

- (112) [os [livros da filosofaphia]ⁱ [que eu compus [livros da filosofaphia]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

Next, let us suppose, in line with Nunes (2004 and subsequent work), that the syntactic object in (112) cannot be linearized because it is not in accordance with Kayne's (1994) Linear Correspondence Axiom (LCA) (see Chapter 1, Section 4.1). The LCA dictates that at PF, an element cannot asymmetrically c-command and be asymmetrically c-commanded by the same element in a structure. Because the two instances of *os livros da filosofaphia* 'the books of philosophy' are non-distinct, the verb *compus* 'wrote' is required to precede and be preceded by the same element. This induces a violation of the asymmetry, canceling the derivation.

Under the assumption that deletion for purposes of linearization only targets constituents, the deletion of copies may then yield the outputs in (113)a-b, with a single application of deletion, and in (113)c-d, with two applications of deletion.

- (113) a. [os [livros da filosofaphia]ⁱ [que eu compus [~~livros da filosofaphia~~]ⁱ]]
 b. [os [~~livros da filosofaphia~~]ⁱ [que eu compus [livros da filosofaphia]ⁱ]]
 c. [os [~~livros da filosofaphia~~]ⁱ [que eu compus [~~livros da filosofaphia~~]ⁱ]]
 d. [os [~~livros da filosofaphia~~]ⁱ [que eu compus [livros ~~da filosofaphia~~]ⁱ]]

If there were no convergence problems resulting from these reductions, the derivations would be eligible for economy compariso, and the derivations yielding *c.* and *d.* would be excluded for employing more operations of deletion than necessary. That is, economy principles would ensure that deletion applies as few times as possible.

An economy-based explanation is also available to cancel the derivation yielding *b.* Because the formal features associated with the higher copy have already been rendered invisible for PF upon checking, the deletion of the lower copy will employ fewer applications of deletion in later computations of the phonological component. Therefore, all else being equal, *a.* is the most economical way of deriving (112).

Under these conditions alone, remnant-internal relativization would never be derived. The most economical derivation in (113)a would always be preferred and, as a result, Pattern I would be the only option.

5.2.2. Pattern II

However, as in any economy approach, if the most economical option does not lead to convergence, a less economical option may be chosen. According to the mainstream version of the copy theory of movement adopted here, lower copies can be pronounced if there is an independent well-formedness PF requirement that precludes the pronunciation of a higher copy. The same line of reasoning applies to scattered deletion: if full deletion does not satisfy PF requirements, deletion may apply within different chain links.

In the light of these assumptions, the derivation yielding remnant-internal relativization (with the modifier/complement in a final position) can also be a legitimate outcome. In this case, the phonological system resorts to scattered deletion, with part of the noun phrase pronounced in the higher copy and part in the lower one, as shown in (114):

- (114) [os [livros da ~~philosophia~~ⁱ] [que eu compus [livros da philosophia]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

I submit that this is the method to which the computational system resorts in order to accommodate the following PF requirements (to be detailed below).

PF requirement I. In non-corrective contexts, emphatic stress must be rightmost.

PF requirement II. The EPP feature on C° dictates that the relative head must be pronounced in the higher copy.

A. PF requirement I

Sticking to the principle that synchrony can inform historical approaches to languages, I presume that earlier stages of Portuguese are like CEP in that (narrow)

information focus always appears in the rightmost position of the sentence²² (cf. Costa 1998 (for CEP) and Martins 2002 (for earlier stages of Portuguese)).

Moreover, I assume that earlier stages of Portuguese pattern with CEP in that an emphatic/contrastive focus in non-corrective contexts²³ must either move to a dedicated left peripheral focus-position (see (111)) or be clause-final (just like narrow information focus); its location in a non-final position leads to ungrammaticality (see (115)b-(116)b).^{24 25}

- (115) a. Comi prego no prato raspado.
 ate:1SG steak in.the dish scraped
 ‘I ate scraped steak in the dish.’
- b. *Comi prego RASPADO no prato.
 ate:1SG steak scraped in.the dish

²² Recall from Chapter 1 (Section 4.3.1) that this is due to the fact that the (narrow) information-focused constituent bears a nuclear neutral stress assigned via the NSR (Zubizarreta 1998, 1999).

²³ Corrective contexts aim to set right a poorly transmitted or wrongly received part of a message. In this case, the emphatic stress can fall on any item: a phrase (as in (i)), a word (as in (ii)) or even an individual syllable of a word: *I said DEfensive, not OFfensive* (Enkvist 1980: 135).

- (i) [A] a. O rapaz de olhos azuis é bonito.
 the boy of eyes blue is handsome
 ‘The boy with blue eyes is handsome.’
- [B] b. O rapaz DE OLHOS VERDES é bonito.
 the boy of eyes green is handsome
 The boy WITH BLUE EYES is handsome (the one with blue eyes is absolutely stunning!).’
- (ii) [A] a. Vamos para Paris.
 go:3PL to P.
 ‘We are going to Paris (±and stay there for a while).’
- [B] b. Vamos A Paris.
 go:3PL to P.
 ‘We are going to Paris (±and come back soon).’

²⁴ Examples (115) and (116) are adapted from Martins (2004).

²⁵ Making use of the notational conventions in Zubizarreta (1999), the emphatic/contrastive focus identified by the E/CSR rule is indicated by small caps, whereas the (rightmost) emphatic/contrastive focus identified by the NSR is underlined (see, e.g., Zubizarreta 1999: 4229).

- (116) a. Paguei de multa cem euros.
paid of fine one.hundred euros
 ‘I paid a fine of one hundred euros.’
- b. *Paguei CEM EUROS de multa.
paid one.hundred euros of fine

The contrast between the *a.* and *b.* examples is not surprising under the system developed thus far. As already mentioned, if emphatic/contrastive focus is assigned by prosodic prominence alone, it is freely assigned in corrective contexts but not in non-corrective ones.²⁶ In the latter case, which corresponds to examples (115)-(116), the nuclear stress always targets the rightmost constituent (being assigned by the NSR).

Then, the question arises as to how the ambiguity between (narrow) information focus and emphatic/contrastive focus in the rightmost sentential position is resolved in CEP. Following Frota (1998, 2002 and much related work), I propose that the ambiguity just mentioned is eliminated by differences in peak alignment (or choice of pitch accent) (see Hualde 2002 for a brief overview). As Frota demonstrates, CEP-declarative sentences with emphatic/contrastive focus in the last word are systematically distinguished from neutral declaratives. If the last word carries a neutral prosodic/information focus, it is pronounced with a falling contour through the last stressed syllable from a preceding peak (H+L*). In contrast, if an emphatic/contrastive focus is intended, the last word is pronounced with a circumflex

²⁶ Remarkably, the idea that emphatic stress assignment may be constrained by construction-specific conditions is not new. Culicover and Winkler (2008) propose a solution along these lines to account for the comparative inversion structure (CI) in (i)a. They show that CI is a focus construction that prosodically marks its subject, places it at the right edge of the intonational phrase and requires a contrastive focus reading. To account for these properties, they argue that the subject is in [Spec, IP] in (i)b but not in (i)a. This difference follows from the markedness constraint in (ii), which applies mandatorily in CI.

- (i) a. Sandy is much smarter than is the professor.
 b. Sandy is much smarter than the professor is.

(Culicover and Winkler 2008: 1)

- (ii) *Right Edge Alignment Constraint of Contrastive Focus in CI*

Each contrastively focused constituent is right-aligned in ip [intonational phrase].

(Culicover and Winkler 2008: 22)

contour (rise followed by fall), with a peak over the stressed syllable. Frota (1998) illustrates this difference with the one-word utterance *casaram* ‘they got married’. As depicted in (117),²⁷ neutral and emphatic contrastive foci are distinguished by different intonational contours.

(117)

- a. Neutral focus b. Emphatic/contrastive focus



The PF facts just discussed nicely illuminate our general understanding of remnant-internal relativization, providing us with the tools to handle the pattern in II (repeated here as (118)):

(118) [os [livros ~~da~~ philosophia]ⁱ [que eu compus [livros da philosophia]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

First, recall from Section 5.1.1 that remnant-internal relativization emerges in non-corrective contexts where the second split-part is assigned emphatic/contrastive focus. In this case, emphatic/contrastive focus is marked prosodically via the NSR (the neutral focus and the emphatic/contrastive focus being distinguished by different intonational contours).

Adopting Stjepanović’s (2007) view of the relation between PF deletion and stress assignment, the deletion of the higher copy of *da philosophia* in (109) can be explained as follows.

- In the output of syntax, (118) has the focus structure in (119):

²⁷ For ease of reproduction, the contour presented in (117) is taken from Hualde (2002).

(119) [os [livros da philosaphia]ⁱ [que eu compus
 [-F][-F] [-F][-F] [-F] [-F] [-F]
 [livros da philosaphia]ⁱ]]
 [-F] [+F] [+F]

- In earlier stages of Portuguese (as in CEP), the NSR applies just after spell out, assigning prominence to the rightmost/lowest sentential constituent (i.e., to *philosaphia* in (119)).
- The element that receives the nuclear stress has an [+F]-feature; hence, the FPR does not conflict with the NSR.
- The higher copy of *da philosaphia* is deleted, and the lower one is pronounced. Not pronouncing this copy would cause a PF violation given that no main stress would be assigned.

Importantly, the present approach also explains why the modifier/complement in Pattern I (repeated here as (120)) cannot be pronounced in the lower copy.

(120) [os [livros da philosaphia]ⁱ [que eu compus [~~livros da~~ ~~philosaphia~~]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

The explanation of this pattern is as follows.

- In the output of syntax, (120) has the focus structure in (121):

(121) [os [livros da philosaphia]ⁱ [que eu compus
 [-F][-F] [-F][-F] [-F] [-F] [+F]
 [livros da philosaphia]ⁱ]]
 [-F] [-F] [-F]

- In earlier stages of Portuguese (just like in CEP) the NSR applies just after spell out, assigning prominence to the rightmost/lowest sentential constituent (i.e. to *philosaphia*).
- Because the rightmost element is [-F], a conflict situation emerges between the NSR and the FPR. To resolve this conflict, CEP renders defocalized [-F]

elements extrametrical for the application of the NSR, and *compus* ‘wrote’ is assigned the neutral nuclear stress.²⁸

- Because there is no PF requirement precluding the pronunciation of the higher copy of *da philosophia* ‘of philosophy’, the most economical option in (120) is derived (with deletion of the lower copy).

B. PF requirement II

The question that remains to be answered now is why the head noun is not pronounced in the lower copy as well; in other words, why the more economical option of full copy deletion is not allowed and the last-resort mechanism of scattered deletion is used instead.

Adopting as my point of departure Chomsky’s (2000) view on EPP, I assume that core functional categories, such as *v*, *T* and *C*, can have an EPP feature, which requires that their Specifier position be filled. This is the case of the relative *C* in CEP and in earlier stages of Portuguese. In this language, there are no head-internal relative clauses,²⁹ a fact that clearly indicates that the relative *C* must have its Specifier position filled.

Following Laudau (2007) (and much related work), I take the EPP to be a PF constraint that requires the Specifier of a functional category to have its Spec filled with an overt element at PF. This idea has been implemented in different ways in the literature. In line with Laudau (2007), I will assume that (i) the EPP is a selectional feature governing PF configurations; (ii) it is locally satisfied by some element with

²⁸ According to Zubizarreta (1998), languages vary in the way they resolve the conflict between the FPR and the NSR. As already mentioned in Chapter in (Section 4.3.1), in languages like English and French defocalized [-F] elements are treated as extrametrical in the sense that they are skipped by the NSR. Other languages (like Spanish, Italian and CEP) employ prosodically motivated movement (p-movement), which moves the post-focal [-F] elements out of the rightmost clausal position. If the hypothesis put forth in this analysis proves correct, it means that these two mechanisms are not necessarily incompatible; instead, they may coexist in the same language (e.g., CEP) as different ways of resolving the conflict between the FPR and the NSR. Under this view, CEP differs from French and English in two aspects: (i) allowing p-movement of defocalized [-F] constituents and (ii) requiring extrametrical material to be deleted in PF.

²⁹ As already noted in Section 2.3 (point G.), these are relative clauses in which the head appears inside the relative clause (see Section 2.3 for more details).

phonological content; and (iii) the satisfaction of the EPP selectional requirement is not evaluated until PF.³⁰

The approach to EPP just outlined gives us a good basis for explaining why the head is pronounced in the higher copy in Pattern II (see (122), repeated from above): because the relative C has an EPP feature, failure to pronounce the head in [Spec, CP] would represent a PF violation.

(122) [os [livros ~~da~~ ^{philosophia}]ⁱ [que eu compus [livros da philosophia]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

5.2.3. Pattern III

To account for remnant-internal relativization constructions with the second split-part in a non-final position (as in (123), repeated from above), I propose a remnant movement approach interpreted in terms of the copy theory of movement (see Bošković and Nunes 2007: 65).

(123) os livros que da philosophia eu compus
the books that of.the philosophy I wrote

In order to derive this distributional pattern, I make the following assumptions.

- The modifier/complement is interpreted as emphatic/contrastive focus, which is marked syntactically.
- In earlier stages of Portuguese, a designated Focus position (see Rizzi 1997 and much related cartographic work) exists in the left periphery of dependent and non-dependent clauses.³¹
- The (emphatic/contrastive) focused constituent (*da philosophia* ‘of philosophy’ in (123)) undergoes leftward movement to check a Focus feature in FocP.³²

³⁰ It is worth clarifying that, under this view, the EPP does not trigger movement on its own. It is always parasitic on an anchoring feature (in this case the *wh*-feature), which triggers movement.

³¹ To account for the presence of a Focus projection in relative clauses, I assume, in line with Bianchi (1999), that the complementizer *que* ‘that’ is spelled out in Force^o and that there is a Focus projection below Force that hosts interrogative and focalized phrases in its Spec.

- The noun phrase (containing a copy of the moved modifier/complement) moves to [Spec, CP] (or [Spec, ForceP] in the new terms) for *wh*-checking.

Under these assumptions, at the output of syntax the structure of (123) would be as in (124):³³

(124) [os [livros da philosophia]ⁱ [que [da philosophia]^j eu compus [livros [da philosophia]^j ⁱ]]]

The particular configuration dictated by remnant movement requires PF deletion to apply to two different chains: the PP chain formed by the movement of *da philosophia* ‘of philosophy’ and the noun phrase chain formed by the movement of *livros da philosophia* ‘books of philosophy’. Note that if the lower copy of these two chains were deleted (see (125)), the resulting structure would not be linearized due to the presence of more than one copy of *da philosophia* ‘of philosophy’.

(125) [os [livros da philosophia]ⁱ [que [da philosophia]^j eu compus [~~livros [da philosophia]^j ⁱ]]]~~

Adopting the *representational* hypothesis proposed in Bošković and Nunes (2007: 65), let us assume that spell out sends the whole structure in (124) to the phonological component. Let us furthermore suppose that the functional category Foc has an EPP feature that requires its Spec to be filled with phonological material (and that failure to pronounce the material in Spec would cause a PF violation).

Given these requirements, Chain Reduction inspects the PP chain and instructs the phonological component to delete the occurrence of *da philosophia* ‘of philosophy’ that is a sister of *livro* ‘book’. As there are two elements that satisfy this description, Chain Reduction ends up deleting the two copies that satisfy this

³² From a typological point of view, the internal movement of head-related elements is not surprising. Basilico (1996) shows that in some languages internally-headed relative clauses display movement of the head to a sentence-internal position.

³³ For simplicity, I am abstracting away other syntactic movements involved in this sentence.

instruction, as represented in (126) (see Bošković and Nunes 2007 for more technical details).

- (126) [os [livros ~~da~~ ^{philosophia}]ⁱ [que [da philosophia]^j eu compus [livros [~~da~~ ^{philosophia}]^j]ⁱ]]

As for the noun phrase chain formed by the movement of *livros da philosophia* ‘books of philosophy’, let us adopt a solution similar to that proposed for Pattern II. According to this proposal, the EPP is a selectional feature governing PF configurations, which is locally satisfied by some element with phonological content. Because the relative C has an EPP feature, failure to pronounce phonological material in [Spec, CP] represents a PF violation. The higher pronunciation of the head noun is thus necessary to avoid a PF violation:

- (127) [os [livros ~~da~~ ^{philosophia}]ⁱ [que [da philosophia]^j eu compus [~~livros~~ [^{philosophia}]^j]ⁱ]]

5.3. Remnant-internal relativization in CEP?

In the discussion of remnant-internal relativization in earlier stages of Portuguese, the reader might have wondered whether remnant-internal relativization is still possible in CEP and, if so, whether there is any contrast with the patterns attested in the history of Portuguese. Although this is clearly beyond the scope of the current study, suffice it to say that in CEP, remnant-internal relativization is possible with the modifier/complement in the rightmost position, as in (128)a-(131)a. The occurrence of the modifier/complement in the left periphery of the relative clause leads to ungrammaticality, as illustrated in (128)b-(131)b.

- (128) a. uma filha que eu tenho pequena
 a daughter that I have young
 lit. ‘a young daughter that I have’
 b. *uma filha que pequena eu tenho
 a daughter that young I have

(example *a.* is cited in Martins 2004)

- (129) a. uns touros que aqui tinha agrestes
some bulls that here had:1SG wild
 lit. ‘some wild bulls that I had here’
 b. *uns touros que agrestes aqui tinha
some bulls that wild here had:1SG
 (example *a.* is from the TV-show *Liga dos Últimos*)
- (130) a. uma casa que eu comprei de cinco assoalhadas
a house that I bought of five rooms
 lit. ‘a house with five rooms that I bought’
 b. *uma casa que de cinco assoalhadas eu comprei
a house that of five rooms I bought
- (131) a. um rapaz que eu conheço de Leiria
a boy that I know of L.
 lit. ‘a boy from Leiria that I know’
 b. *um rapaz que de Leiria eu conheço
a boy that of L. I know

Somewhat tentatively, I would like to suggest that the contrast between CEP and earlier stages of Portuguese can be reduced to an independent change that took place in the history of Portuguese, namely the loss of a focus-dedicated position in the left periphery of embedded clauses. In earlier stages of Portuguese there is a focus position available in the left periphery of dependent and non-dependent clauses, whereas in CEP this position is no longer available in dependent clauses or, at least, in some types of dependent clauses. As already mentioned in footnote 21 above, this hypothesis is consistent with the quantitative results of Martins (1994, 2001), which indicate that the frequency of preposed contrastive foci structures decreased significantly after the 17th century.

5.4. Summary

Focusing on the derivation of remnant-internal relativization, this section aimed to derive the following three patterns:

Pattern I. Regular relativization (involving adjacency between the head noun and its modifier/complement).

Pattern II. Remnant-internal relativization (with a modifier/complement in the rightmost clausal position).

Pattern III. Remnant-internal relativization (with a modifier/complement in the left periphery of the RRC).

Under the approach developed in this section, these patterns are derived as follows.

Pattern I. The head and its modifier/complement are copied and merged in [Spec, CP] to check the *wh*-feature on C° ; Chain Reduction applies at PF, deleting the lower copy.

Pattern II. The head and its modifier/complement are copied and merged in [Spec, CP] to check the *wh*-feature on C° ; scattered deletion applies at PF, deleting the modifier/complement in the higher copy and the head in the lower one.

Pattern III. First, the modifier/complement moves to [Spec, Foc] to check a focus feature on Foc° ; then, the noun phrase (containing a copy of the moved modifier/complement) moves to [Spec, CP] (or [Spec, Force]) for *wh*-checking. Chain Reduction applies, deleting: (i) two copies of the modifier/complement (one in [Spec, CP] and the other in the merge position); and (ii) the head in the lower copy.

The PF requirements claimed to regulate the deletion of copies in remnant-internal relativization were the following.

- The stress rule assignment dictates that the emphatic/contrastive focus (in non-corrective contexts) must surface in the rightmost clausal position.
- The EPP feature on C° dictates that the relative head must be pronounced in the higher copy.
- The EPP feature on F° dictates that the (emphatic/contrastive) focused constituent must be pronounced in [Spec, Foc].

Finally, it was shown that CEP contrasts with earlier stages of Portuguese in not allowing remnant-internal relativization with the modifier/complement in a non-final position. According to the hypothesis formulated in this study, this is due to the loss of a focus-dedicated position in the left periphery of embedded clauses.

6. Conclusion

The term *remnant-internal relativization* is proposed in this chapter to describe RRCs where the head noun and some modifier/complement related to it appear discontinuously, as schematically represented in (132):

(132) [S-matrix ... [N [RRC ... modifier/complement ...]] ...]

On the basis on empirical data from earlier stages of Portuguese, two distributional patterns of remnant-internal relativization were identified:

- (i) Remnant-internal relativization (with a modifier/complement in the rightmost clausal position), as in (133):

(133) os livros que eu compus da philosaphia
the books that I wrote of.the philosophy

- (ii) Remnant-internal relativization (with a modifier/complement in the left periphery of the RRC), as in (134):

(134) os livros que da philosaphia eu compus
the books that of.the philosophy I wrote

The main goal of this chapter was to show that the phenomenon of remnant-internal relativization provides important new evidence for the raising analysis of RRCs.

The first step of this inquiry was to demonstrate that the adjunction analysis of RRCs could not account for the properties of remnant-internal relativization. In particular, the following properties were shown:

- If the head and its modifier/complement were base-generated together in an RRC-external position, the pattern in (ii) could not be derived, as it would require lowering of the modifier/complement to a non-c-commanding position.
- If the head and its modifier/complement were generated separately (the head being CP-external and the modifier/complement being CP-internal), the semantic dependency between the head and its modifier/complement, which

requires that these elements be in a structural relation at some point of the derivation, would not be satisfied.

In a second step, an analysis of remnant-internal relativization in terms of the copy theory of movement was carried out that not only accounts for the properties of remnant-internal relativization but has the welcome effect of deriving regular and non-remnant-internal relativization from essentially the same syntactic structure.

Under the view advocated in the present chapter, the contrast between regular and remnant-internal relativization is due to the deletion operations that take place in the PF side of the grammar. In a nutshell, the idea is that both constructions are derived from a structure like (135):

- (135) [os [livros da filosofaphia]ⁱ [que eu compus [livros da filosofaphia]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

The most economical derivation generates regular relativization, as in (136). This is due to an economy principle (proposed by Nunes 1999, 2004), which prefers fewer applications of deletion in later computations of the phonological component. In the case at hand, the *wh*-feature of the higher copy of *livros da filosofaphia* ‘books of philosophy’ has already been rendered invisible for PF (and LF) purposes upon checking and therefore need not be deleted by later computations of the phonological component.

- (136) [os [livros da filosofaphia]ⁱ [que eu compus [~~livros da filosofaphia~~]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

However, a lower copy can be pronounced if there is an independent well-formedness PF requirement that precludes the pronunciation of a higher copy. This is what happens when the modifier/complement surfaces in the rightmost position. Because the pronunciation of the higher copy of *da filosofaphia* ‘of philosophy’ would violate a PF requirement on stress assignment, scattered deletion applies, yielding (137):

- (137) [os [~~livros da filosofaphia~~]ⁱ [que eu compus [livros da filosofaphia]ⁱ]]

The distributional pattern of remnant-internal relativization exemplified in (138) requires essentially the same basic steps. The difference is that an additional chain link is involved, formed by the movement of the PP *da philosophia* ‘of philosophy’ to the left periphery of the relative clause. It is only after this first movement takes place that the noun phrase (containing a copy of the moved modifier/complement) undergoes movement. The phonological component blindly scans the structure to delete the copy of *da philosophia* ‘of philosophy’ that is a sister of *livro* ‘book’, and the copy in the left periphery of the relative clause is the one that gets pronounced. Then, given the presence of an EPP feature in the relative C, the lower copy of the head *livros* ‘books’ is deleted. See (138):

- (138) [os [livros ~~da philosophia~~]ⁱ [que [da philosophia]^j eu compus [livros [da philosophia]^j]ⁱ]]

Abstracting away from particular derivations, the global picture that emerges is that earlier stages of Portuguese (and CEP, in a more restricted way) had at their disposal constituent discontinuity as a way of syntactically marking emphatic/contrastive focus. Just like clefts, remnant-internal relativization (and, more generally, split noun phrases) appears to constitute a syntactic environment capable of codifying emphatic/contrastive focus. Under this view, it is not surprising that emphatic/contrastive focus in remnant-internal relativization can be additionally marked by prosodic prominence (see (133)) or syntactic movement (see (134)). This squares up nicely with the observation (already noted in Chapter 1, Section 4.3.1) that different focus marking devices may conspire to encode emphatic/contrastive focus.

In summary, the thesis developed throughout this chapter is that remnant-internal relativization provides new evidence for the raising analysis of RRCs. As such, it should number among the arguments commonly adduced in the literature in favor of the raising analysis of RRCs.

Chapter 3

Variation and Change in the Syntax of Restrictive Relative Clause Extraposition

Contents

1. Introduction.....	99
1.1. Goals	100
1.2. Chapter outline.....	101
2. The syntax of RRC extraposition.....	102
2.1. Competing analyses	102
2.1.1. Rightward adjunction analyses	102
2.1.2. Stranding analyses	104
2.1.3. Coordination analyses.....	105
2.2. Unitary vs. non-unitary approaches to extraposition	106
3. RRC extraposition in Portuguese: previous scholarship.....	108
4. Properties of RRC extraposition in contrast	112
4.1. Properties of RRC extraposition in CEP.....	112
4.1.1. Summary	122
4.2. Properties of RRC extraposition from a cross-linguistic perspective.....	123
4.2.1. Summary	128
4.3. Properties of RRC extraposition in earlier stages of Portuguese.....	129
4.3.1. Summary	137
5. The analysis of RRC extraposition in CEP.....	138
5.1. The stranding analysis.....	138
5.2. RRC extraposition and movement to the left periphery	140
5.3. RRC extraposition and short scrambling	142
5.3.1. Subject and object scrambling in CEP	142
5.3.2. The involvement of short scrambling in RRC extraposition	146
5.3.3. The trigger for short scrambling in RRC extraposition	148

5.3.3.1	Short scrambling and information structure in CEP.....	148
5.3.3.2	Short scrambling and information structure in RRC extraposition.....	152
5.3.4.	Deriving the intervening material	156
5.3.4.1	Deriving the occurrence of adverbs and PPs in the intervening position.....	157
5.3.4.2	Blocking the occurrence of other constituents in the intervening position	169
5.3.5.	Summary	174
5.4.	Deriving the properties of RRC extraposition in CEP	175
5.4.1.	Summary	181
5.5.	Problems with the stranding analysis.....	182
5.5.1.	Summary	192
5.6.	Conclusion	193
6.	The analysis of RRC extraposition in earlier stages of Portuguese.....	193
6.1.	The specifying coordination analysis.....	194
6.1.1.	Koster (2000)	194
6.1.2.	De Vries (2002).....	200
6.2.	Deriving the properties of RRC extraposition in earlier stages of Port.	204
6.2.1.	Summary	208
6.3.	Problems with the specifying coordination plus ellipsis analysis.....	209
6.4.	Conclusion	214
7.	Comparative remarks	215
7.1.	Diachronic change in RRC extraposition	216
7.2.	Cross-linguistic contrasts	222
8.	Conclusion	224

1. Introduction

The term *extraposition* is used ambiguously in the literature on syntactic theory. It is used in a pre-theoretical sense to refer to the non-adjacency between two parts of a constituent and in a theoretical sense to refer to a specific type of movement (typically rightward movement). It is usually conceived as a very general phenomenon that affects both relative clauses and a wide range of constituents (e.g., conjuncts, result clauses, appositions, comparative clauses, prepositional phrases and complement clauses) (see De Vries 2002: 236-237 for an overview).

This chapter focuses on the extraposition of restrictive relative clauses (RRCs). In this context, the term *extraposition* will be used in a pre-theoretical sense to refer to an RRC that does not appear adjacent to the antecedent, instead being separated from it by material that belongs to the matrix clause, as depicted in (1):

(1) [... [antecedent] ... RRC]

An example of this construction is given in (2). Example *a.*, which exhibits adjacency between the antecedent and the RRC, displays the regular word order. Example *b.*, which exhibits non-adjacency between the antecedent and the RRC, demonstrates an extraposed RRC. In the contexts of extraposition, there are elements that intervene between the antecedent and the RRC (e.g., the verb, the preposition, and the adverb in (2)b). These elements will henceforth be referred to as *intervening material* and will be underlined for ease of reading.

(2) a. A man [_{RRC} that I met last year] came in yesterday.

b. A man came in yesterday [_{RRC} that I met last year].

(adapted from Givón 2001: 208)

In the traditional account of RRC extraposition, the RRC is said to undergo rightward movement to a right-adjoined position (see, e.g., Reinhart 1980 and Baltin 1984). However, within more recent developments in generative grammar, rightward movement is excluded from the grammar altogether (see Kayne 1994) or at least from core syntax (see Chomsky 1995, 2000).

These developments in generative grammar have obviously energized the debate on the structure of RRC extraposition. The challenge was not an easy one, especially for the adepts of Kayne's antisymmetry framework; it was necessary to determine a syntactic analysis of extraposition that excluded not only rightward movement but also rightward adjunction.

In this context, various solutions are put forward in the literature. For instance, Kayne (1994) proposes that extraposition involves leftward movement of the antecedent and stranding of the relative clause, and De Vries (1999, 2002) proposes an analysis of extraposition in terms of coordination. Notably, these two analyses have the advantage of being compatible with the head-raising analysis of RRCs.

At the same time, other solutions are offered in the literature, which cannot be accommodated in the traditional Y-model of grammar. This is the case for the analysis suggested by Fox & Nissenbaum (1999) and Fox (2002). Eliminating the distinction between covert and overt operations, they claim that RRC extraposition should be analyzed as involving covert quantifier raising (QR) of the antecedent, followed by late merging of the RRC.

Despite the wide range of analyses already available in the literature, extraposition is still a rather poorly understood phenomenon. However, this is not to say that little has been written about it. In contrast, over the last years, several studies have gradually added details to the picture, but most of these studies primarily focus on Germanic languages, especially English, German, and Dutch. Unfortunately, little has been said about RRC extraposition in Romance languages, and as a consequence, the theoretical impact of cross-linguistic variation remains largely unexplored.

1.1. Goals

The major goal of this chapter is to contribute to a better understanding of the syntax of RRC extraposition by discussing new empirical evidence from earlier stages of Portuguese and Contemporary European Portuguese (CEP), as well as data from other languages.

From a descriptive point of view, the specific goals of this chapter are as follows:

- To establish clear properties that distinguish RRC extraposition in CEP and earlier stages of Portuguese.
- To correlate the variation documented in the diachronic dimension with that found in the cross-linguistic dimension.

From a theoretical point of view, the specific goals of this chapter are as follows:

- To demonstrate that the variation found in the syntax of RRC extraposition is not compatible with a uniform account of the phenomenon.
- To argue for a dual approach to RRC extraposition, according to which RRC extraposition may involve the specifying coordination (plus ellipsis) analysis (De Vries 2002) or the VP-internal stranding analysis (Kayne 1994).
- To demonstrate that languages (and different stages of the same language) may demonstrate divergence with respect to the specific type they display.
- To show that the concept of extraposition is descriptively useful but lacks explanatory force because it does not unequivocally correspond to a single syntactic type.

1.2. Chapter outline

This chapter is structured as follows. Section 2 presents the most relevant competing analyses of extraposition and introduces the distinction between unitary and non-unitary approaches to the phenomenon. Section 3 is a state-of-the-art survey of what is known about RRC extraposition in CEP. In Section 4, a comparative approach is adopted, showing that different languages and different stages of the same language may contrast with respect to the properties of RRC extraposition. This leads to the postulation of two different strategies of RRC extraposition: the specifying coordination (plus ellipsis) (De Vries 2002) and the VP-internal stranding (Kayne 1994) strategies. Sections 5 and 6 demonstrate how the contrastive properties of RRC extraposition in CEP and earlier stages of Portuguese are derived from the dual approach advocated here. Then, Section 7 focuses on the contrasts found in the diachronic and cross-linguistic dimension, and Section 8 concludes the chapter.

2. The syntax of RRC extraposition

2.1. Competing analyses

Generally speaking, the existing approaches to RRC extraposition can be divided into three different groups (see also De Vries 2002, for an overview):

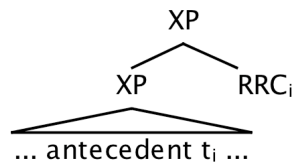
- *Rightward adjunction analyses*
 - Rightward movement (e.g., Reinhart 1980, Baltin 1984).
 - Base-generated rightward adjunction (e.g., Culicover and Rochemont 1990).
 - Rightward adjunction (after covert movement) (e.g., Fox and Nissenbaum 1999, Fox 2002).
- *Stranding analyses*
 - Leftward movement and stranding (e.g., Kayne 1994).
 - Leftward movement (plus deletion) (e.g., Wilder 1995).
- *Coordination analyses*
 - Specifying coordination (e.g., Koster 2000).
 - Specifying coordination (plus ellipsis) (e.g., De Vries 2002).

2.1.1. Rightward adjunction analyses

The basis of the rightward adjunction analysis is that the extraposed RRC is right-adjoined to some maximal projection. Within this approach, some variants can be identified, according to the exact point of the derivation where the rightward adjunction takes place.

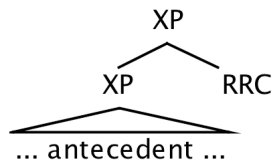
The rightward movement analyses (e.g., Reinhart 1980, Baltin 1984) assume that the RRC is base-generated next to the antecedent. Then, the RRC undergoes rightward movement and right-adjoins to some maximal projection. This is schematically represented in (3).

(3)



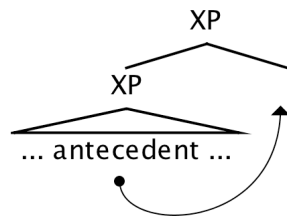
The rightward adjunction analyses (e.g., Culicover and Rochemont 1990) propose that there is no syntactic link between the antecedent and the extraposed RRC. The latter is base-generated in some right-adjoined position, as shown in (4):

(4)

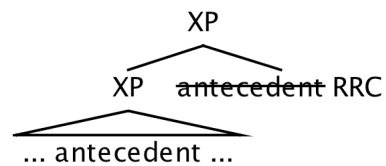


The rightward adjunction (after covert movement) analyses (see Fox and Nissenbaum 1999, Fox 2002) advance a radical alternative to the standard assumptions of overt/covert movement, suggesting that covert operations such as QR can precede overt operations. First, the antecedent undergoes covert movement QR to a rightward position (see (5)a). Then, the RRC is right-adjoined to the antecedent in the post-QR position (see (5)b). Phonology will determine that the antecedent is pronounced in its pre-QR position.

(5) a. QR (covert)



b. RRC merger (overt)



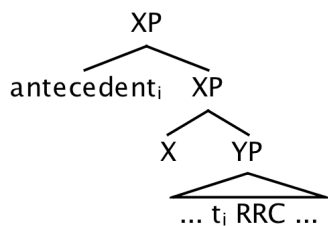
2.1.2. Stranding analyses

The rightward adjunction analyses are challenged by theoretical developments of generative grammar, such as Kayne's (1994) antisymmetric framework. Under this approach, all syntactic representations are asymmetrical in nature, and the linear order is determined by hierarchical relations.¹ As a result, rightward adjunction (and rightward movement) is excluded from the theory of grammar.

As has happened with other phenomena traditionally analyzed as involving rightward positioning (e.g., rightward adjuncts, heavy NP shift, and post-verbal subjects), new proposals of RRC extraposition compatible with Kayne's theory emerged in the literature.

One of the possibilities explored is that RRC extraposition involves leftward movement and stranding (see Kayne 1994). In this paradigm, the antecedent and the RRCs are generated together. Then, the antecedent undergoes leftward movement, stranding the RRC in its base-position. This is schematically represented in (6):

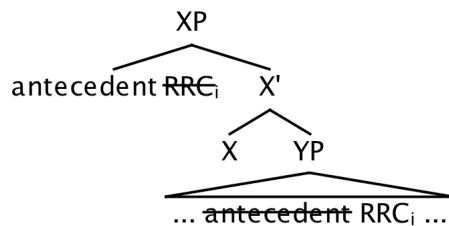
(6)



Another possibility is the leftward movement (plus deletion) analysis (see Wilder 1995), in which the entire noun phrase containing the RRC undergoes leftward movement, leaving a copy behind. Then, there is a deletion of the RRC in the higher copy and a deletion of the antecedent in the lower one (i.e., scattered deletion), as shown in (7):

¹ This is formulated as the Linear Correspondence Axiom (LCA) in Kayne (1994: 5-6). See Chapter 1 (Section 4.1) for additional details.

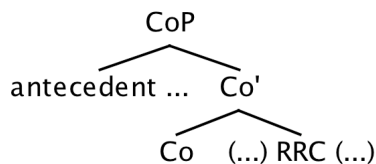
(7)



2.1.3. Coordination analyses

Coordination analyses (e.g., Koster 2000, De Vries 2002) crucially rely on the assumption that extraposition involves the same structure as coordination. Assuming the structure of coordination in Munn (1993) and Kayne (1994) (see also Johannessen 1998), these analyses propose that the antecedent is merged within the Spec of an abstract head, whereas the extraposed RRC is merged in the complement position of the head, as schematically represented in (8):²

(8)



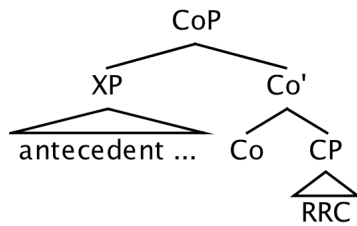
In this structure, the second conjunct specifies (i.e., adds information about) the anchor, hence the term *specifying coordination*.

At least two variants of this approach can be identified. According to Koster (2000), the second conjunct contains only the RRC, which is attached at the relevant line of projection, as sketched in (9):

² The structure in (8) involves a semantically specialized abstract head; it constitutes an asymmetric relationship of specification between the two conjuncts. Koster (2000) symbolically represents this relator by a colon; De Vries (2002) employs an ampersand plus a colon '&:'. Here, I will simply use the more general denotation Co for coordinative head.

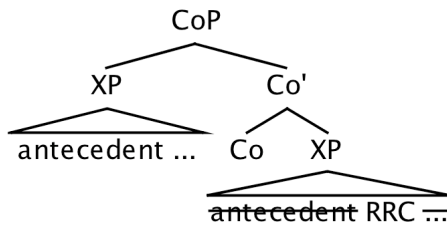
Additionally, note that the use of the parentheses with ellipsis '(...)' in the complement position of the structure in (8) aims to capture the possible presence of additional material in this position, depending on the specific analysis adopted.

(9)



According to De Vries (2002), the second conjunct has the same categorial status as the first conjunct. It repeats the material contained in the first conjunct, adding the extraposed RRC in its canonical position. Then, the repeated material is phonologically deleted, as shown in (10):

(10)



In summary, although this overview of the analyses of extraposition is extremely brief and incomplete, it allows us to conclude that the emergence of different analyses of extraposition was, at least to some extent, motivated by theory-internal concerns. I will return to the stranding and coordination analyses of RRC extraposition in greater detail in Sections 5.1 and 6.1 below.

2.2. Unitary vs. non-unitary approaches to extraposition

There are additional sources of variation in the approaches to extraposition available in the literature. Aside from the variation regarding the specific syntactic structure adopted, analyses may also contrast with respect to their general applicability. In this respect, two different lines of research can be identified: unitary and non-unitary approaches.

Unitary approaches claim that a single syntactic structure can cover a wide range of construction types involving extraposition (within a language and/or across languages). This is, for instance, the case of the coordination analysis proposed by

Koster (2000) and De Vries (1999, 2002). These authors claim that extraposition is not a sub-strategy of relative clauses but is instead a very general phenomenon that applies to a wide range of constructions. Specifically, De Vries (2002) proposes that every construction that may be divided into a first and a second part (*duplex constructions*) allows for the extraposition of the second part (e.g., conjuncts, relative clauses, result clauses, appositions, comparative clauses, PP complements of N, complement clauses of N, and PP complements of A). The same is true of simplex phrases, such as complement clauses of V and heavy NPs. De Vries (2002) proposes that all of these construction types may involve extraposition and that in all these cases, extraposition can be dealt with in terms of specifying coordination.

By contrast, non-unitary approaches claim that more than one syntactic type is necessary to account for the extraposition involved in the different construction types. There are many variants of non-unitary approaches in the literature. Some approaches emphasize the contrast between adjunct and complement extraposition. This is, for instance, the case of the analysis put forth by Fox and Nissenbaum (1999) and Fox (2002), who propose that complement extraposition involves the rightward movement of the complement, whereas relative clause extraposition involves post-QR merging of the relative clause.

Other analyses claim that extraposition is not a unitary phenomenon, even if only relative clauses are considered. This is the case of Smits (1988), who highlights the cross-linguistic variation found in this domain:

The ways in which extraposition seems conditioned is certainly surprising. For, intuitively at least, it is hardly probable for a specific rule of the grammar [...] to be subject to some condition in one language, with that same condition having no relevance whatsoever in another, otherwise closely related language, and vice versa, without any apparent reason. The incoherence of the sprawling array of conditions [...] [in the possibilities of RC-extraposition across languages] alone suggests, perhaps, that extraposition of RCs is only apparently a monolithic phenomenon. That is to say, it suggests that only some proportion of cases we find is the result of the rule of extraposition [...] which moves clauses from their NPs to the right end of the clause.

(Smits 1988: 183)

To account for the different properties of RC-extraposition, Smits (1988) proposes a non-uniform approach to the phenomenon, in which extraposed RRCs are derived from rightward movement (being nevertheless subject to specific conditions in

different languages), whereas appositive relative clauses (ARCs) are detached from the antecedent and base-generated in a rightward position.

Perhaps more surprisingly, non-uniform approaches have also been proposed to account for even more specific constructions, such as RRC extraposition. On the basis of comparative evidence from German and English, Inaba (2005) claims that, in spite of superficial similarities, RRC extraposition in these languages exhibits a remarkable contrast that has been ignored in previous research; RRC extraposition is a syntactic operation in English, whereas it involves phonological movement in German.

In short, two opposing views dominate the analysis and discussion on the syntax of extraposition. On one hand, unitary approaches claim that the concept of extraposition corresponds to an explanatory pattern in the sense that it can be accounted for by a single syntactic structure. In contrast, non-unitary approaches claim that the concept of extraposition may have descriptive adequacy (in unifying apparently related constructions) but lacks explanatory force because it does not correspond to a single construction type. It was also shown that non-unitary approaches may differ with respect to: (i) the range of syntactic construction types covered and (ii) their universal vs. language-specific scope.

3. RRC extraposition in Portuguese: previous scholarship

Examples of RRC extraposition have been reported in a number of languages: English (see, e.g., Ziv and Cole 1974, Quirk et al. 1985, Givón 2001), Dutch (see, e.g., Koster 2000, De Vries 2002), German (see, e.g., Haider 1996, 1997, Kiss 2005) Italian (see Cardinaletti 1987), and Spanish (see Brucart 1999). See also Smits (1988) for an overview.

However, as far as CEP is concerned, only sparse allusions to RRC extraposition are found in the literature (see Brito and Duarte 2003: 661, Barbosa et al. 2005, Barbosa 2009, Smits 1988: 407).

On the basis of the contrast in (11), Brito and Duarte (2003: 661) claim that *RRCs cannot be easily extraposed in CEP*. However, the authors do not specify what they mean by the use of the adverb *easily*, nor do they provide any example of RRC extraposition in CEP.

- (11) a. Uma pessoa que tu conheces telefonou.
a person that you know phoned
 ‘A person that you know phoned.’
- b. Telefonou uma pessoa que tu conheces.
phoned a person that you know
- c. *Uma pessoa telefonou que tu conheces.
a person phoned that you know

Furthermore, while discussing the position of the subject in CEP, Barbosa et al. (2005) and Barbosa (2009) allude to RRC extraposition. Specifically, they show the impossibility of RRC extraposition from an indefinite subject in a pre-verbal position, as in (12):

- (12) *Um homem apareceu que deseja falar contigo.
a man showed.up that wants talk:INF with.you
 ‘A man showed up that wants to talk to you.’

(Barbosa 2009)

This restriction is claimed to correlate with the Null Subject Parameter. The authors contend that Null Subject Languages (NSLs) such as CEP do not allow extraposition from pre-verbal indefinite subjects, whereas non-NSLs such as English and French allow it. According to Barbosa et al. (2005) and Barbosa (2009), this can be explained if we assume that the subject occupies different positions in NSLs and in non-NSLs. In the former, the subjects are left-dislocated, whereas in the latter, they are in [Spec, IP]. I will come back to this analysis in Section 5.4, where I show that the correlation between the possibility of extraposition from pre-verbal subjects and the Null Subject Parameter simply does not hold. Indeed, in earlier stages of its history, Portuguese was an NSL and allowed for the extraposition of indefinite subjects from pre-verbal positions.

In his overview of relative clauses in Germanic and Romance languages, Smits (1988: 407) also refers to some properties of RRC extraposition in CEP. Considering RRC extraposition and ARC extraposition together, he proposes the following descriptive principles:

(13) *Principles of RC-extraposition in CEP* (Smits 1988: 407)

1. Extraposition of relative clauses belonging to the subject is impossible both for restrictives and appositives.
2. Extraposition of relative clauses belonging to objects is possible for restrictives only.
3. Definiteness properties of the restrictive antecedent do not influence the possibilities of extraposition.

As will become clear in Section 4.1, these principles are simply not correct and must be revised.

Finally, other references can be found in the literature, but they only concern extraposition of ARCs (see Brito 2004: 402, Peres and Mória 1995: 367). This issue will be addressed, though in a slightly different context, in Chapter 4 (Sections 3.2.1.2. and 5.2.).

As for earlier stages of Portuguese, the RRC extraposition is almost totally neglected in the grammars and studies on the history of Portuguese. However, there are a few exceptions (see Dias 1933/1970, Mattos e Silva 1989, A. Costa 2004).

Dias (1933/1970: 329) mentions that an RRC can be separated from its antecedent if no ambiguity arises. He illustrates this possibility with the examples in (14)-(16):

- (14) *aquelle se chamará bom prelado que tiver letras, reputação,
that SE:CL call:FUT good prelate that have:FUT.SUBJ letters, reputation
e virtudes
and virtues*

‘People will identify a good prelate as the one who has education, a good reputation and virtues.’

(Dias 1933/1970: 329; Heitor Pinto, *Diálogos*; 16th century)

- (15) *que naquelle coração não ha vestigio de justiça, onde
that in.that heart not has trace of justice where
a avareza tem feyto sua morada
the meanness has done its home*

‘[In a sermon, Pope St. Leo says] that there is no trace of justice in the heart where the meanness is deeply rooted.’

(Dias 1933/1970: 329; Heitor Pinto, *Diálogos*; 16th century)

- (16) Esse, é meu amigo, que moe no meu moinho.
that is my friend that mills in.the my mill

‘The one who mills (something) in my mill is my friend.’

(old saying, cited in Dias 1933/1970: 329)

Mattos e Silva (1989) also reports the possibility of RRC extraposition in sentences such as (17)-(18):

- (17) depois que naquela hora o seu filho ficara sen féver
after that in.that hour the his son stay:PMQP without fever
 en que hi o nosso Salvador e nosso meestre Jesu Cristo dissera que
in that there the our Savior and our master J. C. say:PMQP that
 era são.
was:3SG healed

‘[and the priest realized that it was true] after his son stopped having fever in that hour that our Savior and Master Jesus Christ said that he was healed.’

(Mattos e Silva 1989: 766; *Diálogos de São Gregório*; 14th century)

- (18) que en aquela hora morrera en que el vira estando longe
that in that hour die:PMQP.3SG in that he saw be:GER away
 dele que lhi saira a alma do corpo.
from.him that to.him:CL fall.out:PMQP the soul of.the body

‘[and he realized that] he died in that hour in which he saw (being away from him) that his soul fell out of his body.’

(Mattos e Silva 1989: 766; *Diálogos de São Gregório*; 14th century)

Discussing the syntactic properties of relative constructions in the 14-15th century Portuguese, A. Costa (2004) alludes to the possibility of RRC extraposition in sentences such as (19):

- (19) E mando que se outra mada parecer que eu mândasse
and order:1SG that if another will appear:FUT.SUBJ that I order:IMPERF.SUBJ
 fazer ante dessta que quebre e ñ ualha
make:INF before this that±be.annulled:PRES.SUBJ. and not be.valid:PRES.SUBJ.

‘And, if another will appear that I ordered to be made before this one, I order it to be annulled and not valid.’

(Martins 2001, cited in A. Costa 2004: 414, year 1383)

Additional references to non-adjacency phenomena, namely those concerning the so-called *relatif de liaison* or *connecting relative*, can be found in the literature (see Dias 1933/1970: 269, Said Ali 1931/1971: 107, A. Costa 2004: 418-419). However,

because these constructions clearly do not involve RRCs, their discussion is postponed until Chapter 4 (Section 3.2.1.2).

In short, this survey of previous research plainly demonstrates that much of the syntax of RRC extraposition in Portuguese still awaits a proper description. The next section intends to contribute to filling this gap by offering a description of RRC extraposition in CEP and in earlier stages of Portuguese. Although much of the discussion focuses on the contrast between CEP and earlier stages of Portuguese, evidence from other languages also plays an important role in keeping with a comparative and universalist approach to the phenomenon.

4. Properties of RRC extraposition in contrast

In this section, I identify three main properties of RRC extraposition and show how CEP (Section 4.1), other languages (Section 4.2), and earlier stages of Portuguese (Section 4.3) behave with respect to them.

4.1. Properties of RRC extraposition in CEP

CEP allows for RRC extraposition. By way of illustration, see examples in (20)-(22):

(20) Ainda por cima, dá-se conta de que as obras não têm licença camarária

‘As if it wasn't enough, he/she realizes there is no council permit for the construction works’

e faz diligências na Câmara das Caldas da Rainha que levam
and makes ±measures at.the Town.Hall of.the C. da R. that lead

ao seu embargo.
to.the its embargo

‘and takes some measures at Caldas da Rainha Town Hall that block them.’

(CETEMPúblico - Primeiro milhão)

(21) Encontrei uma pessoa ontem que não via há muito tempo.
met:1SG a person yesterday that not saw:1SG has much time

‘I met a person yesterday that I haven’t seen for a long time.’

(http://coisasbelasesujas.blogspot.com/2004_09_01_archive.html)

- (22) Houve alguém no meio da noite que decidiu agarrar uma cana
was someone in.the middle of.the night that decided grab a pole
 que supostamente seria do Aranha (...)
that supposedly be:IMP of.the A.

‘There was someone who, in the middle of the night, grabbed a pole that supposedly belonged to Aranha.’

(<http://www.pescador.com.pt/livre/viewtopic.php?f=9&t=1772>)

However, RRC extraposition in CEP is limited by a number of restrictions, which are listed in A.-C. below and discussed in turn:

A. *The definiteness effect*

B. *Restriction on extraposition from embedded positions*

C. *Restriction on extraposition from pre-verbal positions*

A. *The definiteness effect*

In CEP, RRC extraposition exhibits a definiteness effect reminiscent of that found in existential constructions with the verb *haver* ‘to have’ (see Duarte and Oliveira 2003: 224, fn. 32). Applied to the RRC extraposition, the definiteness effect is a restriction against the occurrence of ‘definite’ noun phrases as antecedents of extraposed RRCs. This restriction is illustrated by the contrasts in (23)-(24), involving (respectively) a subject and a direct object as the antecedent. As these examples show, RRC extraposition is fine with indefinite antecedents (see *a.* examples) but impossible with definite ones (see *b.* examples).

[subject]

- (23) a. Chegou um rapaz ontem que te quer conhecer.
arrived a boy yesterday that you:CL wants meet:INF

‘A boy arrived yesterday that wants to meet you.’

- b. *Chegou o rapaz ontem que te quer conhecer.
arrived the boy yesterday that you:CL wants meet:INF

‘The boy arrived yesterday that wants to meet you.’

[object]

- (24) a. Encontrei um rapaz no cinema que perguntou por ti.
met:1SG a boy at.the cinema that asked for you

‘I met a boy at the cinema that asked for you.’

- b. *Encontrei o rapaz no cinema que perguntou por ti.
met:1SG the boy at.the cinema that asked for you
 ‘I met the boy at the cinema that asked for you.’

Importantly, if no extraposition is involved, both definite and indefinite antecedents are allowed in these contexts, as illustrated in (25)-(26):

- (25) Chegou ontem um/o rapaz que te quer conhecer.
arrived yesterday a the boy that you:CL wants meet:INF

- (26) Encontrei no cinema um/o rapaz que perguntou por ti.
met:1SG at.the cinema a the boy that asked for you

However, RRC extraposition is not limited to indefinites in a narrow, grammatical sense (i.e., noun phrases with an overt indefinite article). The X-position in (27) may be filled, for example, by *um livro* ‘a book’ or *três livros* ‘three books’ but not by *o livro* ‘the book’ or *aqueles livros* ‘those books’ (see (27)-(28)).

- (27) Foi/foram publicado(s) X recentemente que vale a pena ler.
was/were published X recently that is.worth read:INF
 ‘X that is/are worth reading was/were recently published.’

- (28)
- $$X = \left\{ \begin{array}{l} \textit{um livro} \text{ ‘a book’} \\ \textit{três livros} \text{ ‘three books’} \\ \textit{alguns livros} \text{ ‘some books’} \\ \textit{muitos livros} \text{ ‘many books’} \\ \textit{livros} \text{ ‘books’} \end{array} \right\} *X = \left\{ \begin{array}{l} \textit{o livro} \text{ ‘the book’} \\ \textit{aqueles livros} \text{ ‘those books’} \\ \textit{todos os livros} \text{ ‘all the books’} \\ \textit{cada livro} \text{ ‘each book’} \end{array} \right\}$$

Notably, the noun phrases that can fill the X-position in (27) can be grouped together under the class of *weak noun phrases* (as opposed to *strong noun phrases*), in the sense of Milsark (1974).³ Therefore, the descriptive generalization that captures the

³ Milsark (1974) distinguishes between weak determiners (e.g., *a, some, many, several, and few*), which can occur in *there*-insertion contexts, and strong determiners (e.g., *the, every, each, most, all*), which cannot appear in these contexts (see (i)). He further claims that weak determiners are not quantifiers but cardinality words, whereas strong determiners are quantificational.

- (i) a. There is/are a/some/many/three fly/fies in my soup.
 b. *There is/are the/every/all/most fly/flies in my soup.

relation between RRC extraposition and the definiteness effect can be formulated as in (29):

(29) *The definiteness effect and RRC extraposition*

In CEP, RRC extraposition can only take place from weak noun phrases.

B. *Restriction on extraposition from embedded positions*

In CEP, it is impossible to extrapose from embedded positions. Examples (30)-(34) demonstrate that RRC extraposition cannot take place from the object of a preposition. In these examples, the PP containing the antecedent is either an indirect object (see (30)-(32)) or an oblique complement (see (33)-(34)). As shown in the *b.* examples, the sentences are fine with the normal (non-extraposed) order.

[indirect object]

- (30) a. *O João deu o livro [a uma rapariga] ontem que estava na festa.
the J. gave the book to a girl yesterday thatwas at.the party
 ‘João yesterday gave the book to a girl that was at the party.’

- b. O João deu o livro ontem a uma rapariga que estava na festa.
the J. gave the book yesterday to a girl thatwas at.the party

- (31) a. *Agradececi [a alguns amigos] no jantar que me ajudaram
thanked:1SG to some friends at.the dinner that me:CL helped
 nesta fase difícil.
in.this phase difficult

‘At the dinner, I thanked some friends that helped me in this difficult phase.’

- b. Agradececi no jantar a alguns amigos que me ajudaram
thanked:1SG at.the dinner to some friends that me:CL helped
 nesta fase difícil.
in.this phase difficult

- (32) a. *Oferececi [a uma amiga] esse livro que vive no Algarve.
offered to a friend that book that lives in.the A.

‘I offered this book to a friend that lives in Algarve.’

- b. Oferececi esse livro a uma amiga que vive no Algarve.
offered that book to a friend that lives in.the A.

[oblique complement]

- (33) a. *Esqueci-me [de um livro] ontem que era fundamental.
forgot.me:CL of a book yesterday that was fundamental
 ‘Yesterday, I forgot a book that was fundamental.’
- b. Esqueci-me ontem de um livro que era fundamental
forgot.me:CL yesterday of a book that was fundamental
- (34) a. *O João candidatou-se [a uma câmara] nesse ano que fica
the J. ran.se:CL for a town.council in.that year that is
 no distrito de Bragança.
in.the district of B.
 ‘That year João ran for a position on a town council that is located in the district of Bragança.’
- b. O João candidatou-se nesse ano a uma câmara que fica
the J. ran.SE:CL in.that year for a town.council that is
 no distrito de Bragança.
in.the district of B.

The same constraint is found when the antecedent of an extraposed RRC is embedded in a non-argumental PP:

- (35) a. *Conduzi [numa estrada] ontem que estava em péssimo estado.
drove:ISG in.a road yesterday that was in horrible condition
 ‘Yesterday I drove in a road that was in horrible condition.’
- b. Conduzi ontem numa estrada que estava em péssimo estado.
drove:ISG yesterday in.a road that was in horrible condition

As expected, extraposition is also impossible from a PP within a DP, as in (36):

- (36) a. *Vi [a filha de um rapaz] ontem que joga no Benfica.
saw:1SG the daughter of a guy yesterday that plays in.the B.
 ‘Yesterday, I saw the daughter of a guy that plays for Benfica.’
 (intended reading: *um rapaz* ‘a guy’ is the antecedent of the extraposed RRC)
- b. Vi ontem a filha de um rapaz que joga no Benfica.
saw:1SG yesterday the daughter of a guy that plays in.the B.

C. Restriction on extraposition from pre-verbal positions

C1. Restriction on extraposition from pre-verbal subjects

Extraposed RRCs can take post-verbal subjects as their antecedents, as illustrated in (37)a and (38)a. However, if the subject is construed pre-verbally, the sentence is out, as shown in (37)b and (38)b.

- (37) a. Ontem explodiu uma bomba em Israel que causou 5 mortos.
yesterday exploded a bomb in I. that caused 5 deaths
 ‘Yesterday a bomb exploded in Israel that caused 5 deaths.’
- b. *Ontem uma bomba explodiu em Israel que causou 5 mortos.
yesterday a bomb exploded in I. that caused 5 deaths
- (38) a. Chegou um senhor ontem que fez muitas perguntas sobre ti.
arrived a man yesterday that made many questions about you
 ‘A man arrived yesterday who asked many questions about you.’
- b. *Um senhor chegou ontem que fez muitas perguntas sobre ti.
a man arrived yesterday that made many questions about you

In the context of RRC extraposition, the verbs typically found with post-verbal subjects are unaccusatives, as in (37)-(38): *explodir* ‘to explode’ in (37) is an internally caused unaccusative verb; *chegar* (‘to arrive’) in (38) is an unaccusative verb of inherently directed motion (see Duarte 2003). Other unaccusative-related constructions, such as passive sentences, are also compatible with RRC extraposition, as illustrated in (39). Note again that if the subject is construed pre-verbally, the sentence is out:

- (39) a. Foi capturado um indivíduo esta noite que é responsável pelo
was captured a man this night that is responsible by.the
 assalto ao banco.
robbery to.the bank

‘A man who is responsible for the bank robbery was captured tonight.’

- b. *Um indivíduo foi capturado esta noite que é responsável pelo
a man was captured this night that is responsible by.the
 assalto ao banco.
robbery to.the bank

However, RRC extraposition from the subject is not limited to the spectrum of constructions related to unaccusativity. The subject of unergative verbs may also occur in this context; see (40)a-(43)a. It is nevertheless worth noting that not all unergatives can smoothly surface with an extraposed RRC without a propitious context.

- (40) a. Telefonou um rapaz ontem que queria informações sobre a tua casa.
phoned a boy yesterday that wanted details about the your house

‘A boy phoned yesterday who wanted details about your house.’

- b. *Um rapaz telefonou ontem que queria informações sobre a tua casa.
a boy phoned yesterday that wanted details about the your house

- (41) a. Dormiu uma rapariga ontem em minha casa que está a tirar
slept a girl yesterday at my house that is A do:INF
 o mesmo curso que tu.
the same degree that you

‘A girl who is doing the same degree as you slept at my house yesterday.’

- b. *Uma rapariga dormiu ontem em minha casa que está a tirar
a girl slept yesterday at my house that is A do:INF
 o mesmo curso que tu.
the same degree that you

- (42) a. Trabalha um senhor na minha empresa que nunca viu o mar.
works a man in.the my company that never saw the sea

‘A man works in my company who has never seen the sea.’

- b. *Um senhor trabalha na minha empresa que nunca viu o mar.
a man works in.the my company that never saw the sea

- (43) a. Tossiu um bebé na sala de recobro que deve estar a precisar
coughed a baby in.the room of recovery that must be:INF in need:INF
 de ajuda.
of help
 ‘A baby who must be in need of help coughed in the recovery room.’
- b. *Um bebé tossiu na sala de recobro que deve estar a
a baby coughed in.the room of recovery that must be:INF in
 precisar de ajuda.
need:INF of help

Extrapolation from the subject is also possible when verbs with oblique complements are involved:

- (44) a. Vivem alguns portugueses em Paris que nunca foram à Torre Eiffel.
live some Portuguese.people in P. thatnever went to.the T. E.
 ±‘Some Portuguese people who have never been to the Eiffel Tower live in Paris.’
- b. *Alguns portugueses vivem em Paris que nunca foram à Torre Eiffel.
some Portuguese.people live in P. thatnever went to.the T.E.
- (45) a. Entrou um homem na sala que deve ser o orador convidado.
entered a man in.the room that might be:INF the speaker invited
 ‘A man who might be the invited speaker entered the room.’
- b. *Um homem entrou na sala que deve ser o orador convidado.
a man entered in.the room that might be:INF the speaker invited

The only verb classes that seem not to allow extrapolation from the subject are direct transitive verbs (i.e., verbs with a direct object) and ditransitive verbs; see (46)-(47).⁴

⁴ In CEP, non-monoargumental verbs only allow for the VSO order in particular syntactic or semantic environments (see Martins, in prep.). Importantly, the VSO order is available in the syntactic environment displayed in (46) and (47), where a root sentence containing the verb in the imperfect is articulated with an adverbial clause that locates the situation described by the VSO sentence in the speaker’s perceptual field (see Martins, in prep.). The possibility of the VSO order in this syntactic environment is illustrated in (i) and (ii) below:

- (i) Ontem quando entrei no parque de estacionamento,
yesterday when entered:1SG in.the parking lot
 roubavam três rapazes um carro.
stole:3PL.IMPERF three boys a car
- (ii) Quando cheguei ao aeroporto, vendiam três rapazes o
when arrived:1SG at.the airport sold three boys the
 Borda D’Água.
 B. D.

- (46) *Ontem quando entrei no parque de estacionamento,
yesterday when entered:1SG in.the parking lot
 roubavam três rapazes um carro que são amigos do meu filho.
stole:3PL.IMPERF three boys a car that are friends of.the my son
 ‘Yesterday when I entered the parking lot, there were three boys that are my son’s friends stealing a car.’

- (47) *Quando cheguei ao aeroporto, vendiam três rapazes o
when arrived:1SG at.the airport sold:3PL.IMPERF three boys the
Borda D’Água que não tinham mais de cinco anos.
B. D. that not had:3PL.IMPERF more of five years
 ‘When I arrived at the airport, there were three boys that were less than five years old selling the *Borda D’Água* [an almanac].’

In summary, RRC extraposition from the subject obeys two different constraints in CEP: (i) it can take place from post-verbal subjects but not from pre-verbal ones; and (ii) it is compatible with all types of verbs except direct transitives and ditransitives.

C2. Restriction on extraposition from other pre-verbal constituents

In CEP, extraposed RRCs can take a *wh*-constituent (see (48)-(49)), a preposed emphatic/evaluative phrase (in the sense of Raposo 1995 and Ambar 1999) (see (50)) and a preposed focus (in the sense of Zubizarreta 1999 and Martins in prep.) (see (51)-(52)) as an antecedent.

[*wh*-constituent]

- (48) Quantas pessoas apareceram que não foram convidadas?
how.many people showed.up that not were invited
 ‘How many people showed up who were not invited?’
- (49) Quantas pessoas é que tu conheces que não viram este jogo?
how.many people is that you know that not saw this game
 ‘How many people you know that did not see this game?’

[emphatic/evaluative phrase]

- (50) Muito whisky o João bebeu que estava fora do prazo!
a.lot.of whisky the J. drank that was out of.the expiry.date
 ‘João drank a lot of whisky that was expired!’

[preposed focus]⁵

- (51) Poucas pessoas conheço que fazem interpolação, mas todas elas
few people know:1SG that make interpolation but all they
 produzem coisas deste tipo.
produce things of.this type

‘I know few people who make interpolation, but all of them produce things like this.’

- (52) Nem uma única pessoa apareceu que estivesse interessada em
not a single person showed.up that was:FUT.SUBJ interested in
 colaborar.
collaborate:INF

‘Not even a single person showed up that was interested in collaborating.’

- (53) Uma estranha doença o atingiu que lhe retirou toda a alegria.
a strange disease him:CL hit that to.him:CL took all the joy

‘He caught a strange disease that deprived him of all joy.’

- (54) Nada fiz que pusesse em causa a tua decisão.
nothing made that put:IMPERF.SUBJ in question the your decision

‘I have done nothing that would call your decision into question.’

- (55) Só um homem havia na terra que sabia tudo.
only a man had in.the earth that knew everything

‘There was only one man in the earth that knew everything.’

⁵ As already mentioned in Chapter 1 (Section 4.3.1, fn. 9), the term *preposed focus* is used here to refer to a contrastive-focused constituent that appears in the left periphery of the sentence, as in (i):

- (i) COM ESTAS PALAVRAS me despeço.
with these words me:CL say.goodbye:1SG

‘I say goodbye WITH THESE WORDS.’

The fact that the preposed constituent in (51)-(55) above has a contrastive focus status (and not, for instance, a topic status) is confirmed by three independent indicators. First, consider evidence from clitic positioning. Sentences involving preposed focus display proclisis (i.e., the order clitic-verb), whereas sentences involving topicalization or left-dislocation display enclisis (i.e., the order verb-clitic) (see Martins, in prep.). As shown in example (53), when a clitic is involved, the sentences display proclisis, which indicates that the dislocated constituent is a preposed focus. Second, Duarte (1987: 49; 1996: 328; 1997: 583) and Carrilho (2005: 88) demonstrate that topicalization is sensitive to the referential properties of nominal expressions. In particular, they show that indefinite negatives cannot become topics. However, an indefinite negative is possible in sentences such as (54), which indicates that the preposed constituent is not topicalized. Additionally note that focalization is compatible with negative words, as reported by Zubizarreta (1999: 4241) for Spanish. Finally, a preposed focus may contain a focus operator (such as *só* ‘only’ in (55)), which cannot easily be associated with a topicalized constituent.

However, RRC extraposition is incompatible with topicalization;⁶ see the contrasts in (56) and (57):

- (56) a. *Pessoas que não tinham bilhete, apareceram às centenas!*
people that not had ticket showed.up by hundreds
 ‘People who did not have a ticket showed up by the hundreds.’
- b. **Pessoas, apareceram às centenas que não tinham bilhete!*
people showed.up by hundreds that not had ticket
- (57) a. *Pessoas que praticam yoga, também conheço.*
people that practice yoga also know:1SG
 ‘I also know people who practice yoga.’
- b. **Pessoas, também conheço que praticam yoga.*
people also know:1SG that practice yoga

(Ernestina Carrilho, p.c.)

4.1.1. Summary

This section has shown that RRC extraposition in CEP is subject to specific conditions having to do with, among other things, the position of the antecedent and the contrast between weak and strong noun phrases. The restrictions reviewed thus far are summarized in Table 1. In this table, the use of a plus ‘+’ denotes that RRC extraposition can occur in the relevant context; a minus ‘-’ indicates that it cannot.

Table 1. RRC extraposition in CEP

	CEP	
A. <i>Extraposition from strong noun phrases</i>	-	
B. <i>Extraposition from embedded positions</i>	-	
C. <i>Extraposition from pre-verbal constituents</i>	subjects	-
	<i>wh</i> -constituents	+
	emphatic/evaluative phrases	+
	preposed foci	+
	topics	-

⁶ As mentioned in Chapter 1 (Section 4.3.3), the term *topicalization* is used in this dissertation to refer to topic-comment structures where the topic is syntactically connected with an empty category inside the comment (see Duarte 1987 and subsequent work).

4.2. Properties of RRC extraposition from a cross-linguistic perspective

Interestingly, in a brief survey of the behavior of extraposition in different languages, it becomes clear that the restrictions that hold for RRC extraposition in CEP do not universally apply. Some examples of languages that behave differently with respect to these restrictions are given in A.-C. below.

A. *The definiteness effect*

Not all languages exhibit the definiteness effect found in CEP. Extraposition from strong noun phrases is not possible in Italian, French and Spanish, but it is possible, for example, in English,⁷ Dutch and German:

[Italian]

- (58) *Ho regalato quel libro a Carlo che mi avevi consigliato tu.
have:1SG given that book to C. that to.me had recommed you

‘I gave Carlos that book that you recommended me.’

(Cardinaletti 1987: 44, n. 4)

[French]

- (59) *La radio a été volée, que tu m’as donné.
the radio has been stolen that you to.me.have given

‘The radio that you gave me was stolen.’

(Smits 1988: 332)

[Spanish]

- (60) *Escribió la columna en la prensa la semana pasada
wrote:3SG the column in the press the week last

en la que se quejaba amargamente de su situación.
in the that SE:CL complained:3SG bitterly about his situation

‘Last week he wrote in the press the column in which he complained bitterly about his situation.’

(Brucart 1999: 465)

⁷ Note however that, according to Diesing (1992: 144, n. 23), there is variability in speakers' grammaticality judgments of extraposed RRCs with definite noun phrases as antecedent. See also Ziv and Cole (1974: 781) and Baltin (2006: 243).

- (61) *De repente, apareció el individuo en la reunión
suddenly showed.up the man in the meeting
 que parecía sacado de una película de terror
that seemed taken from a movie of horror

‘Suddenly the man showed up that seemed to be taken from an horror movie.’

(Brucart 1999: 465)

[English]

- (62) The woman came in yesterday that I told you about.

(Givón 2001: 206)

- (63) Those students will pass this course who complete all of their assignments on time.

(Baltin 2006: 243)

- (64) That loaf was stale that you sold me.

(Quirk et al. 1985: 1397)

- (65) She rapidly spotted the book right on my desk that I had been desperately searching for all the morning.

(Quirk et al. 1985: 1398)

[Dutch]

- (66) Ik heb de man gezien die zijn tas verloor.
I have the man seen who his bag lost

‘I have seen the man who lost his bag.’

(De Vries 2002: 65)

[German]

- (67) als sie endlich selbst über die Musik erzählen darf,
when she finally herself about the music tell may
 die sie macht.
that she makes

‘when she finally is allowed to speak herself about the music that she makes.’

(Tübinger Baubank des Deutschen/Schriftsprache, cited in Strunk 2007)

B. Restriction on extraposition from embedded positions

The restriction on extraposition from embedded positions does not equally apply to all languages. It is reported in the literature that extraposed RRCs can take the object of a preposition as the antecedent, for example, in English, Dutch and German:

[English]

(68) John is going to talk [to someone] tomorrow who he had a lot of faith in.

(Kayne 1994: 126)

[Dutch]

(69) Ik heb [op een plek] gelopen waar jij ook bent geweest.
I have on a spot walked where you also have been

‘I have walked on a spot where you also have been.’

(De Vries 2002: 244)

[German]

(70) weil er auf eine Frau gewartet hat, die einen Porsche fährt
since he for a woman waited has who a P. drives

‘since he has been waiting for a woman who drives a Porsche.’

(Meinunger 2000: 208)

De Vries (2002: 246) also reports that RRC extraposition in Dutch may take place from a PP within a DP:

[Dutch]

(71) Ik heb [de papieren van *de man*] gecontroleerd die een rode jas droeg.
I have the papers of the man checked who a red coat wore

‘I have checked the papers of the man who wore a red coat.’

C. Restriction on extraposition from pre-verbal positions

C1. Restriction on extraposition from pre-verbal subjects

Barbosa et al. (2005) and Barbosa (2009) report that the extraposition of RRCs from a pre-verbal position is possible in some languages (e.g., English and French). However, it is impossible in Spanish, Catalan, and Italian (as well as CEP); see (72)-(73). As already mentioned in Section 3, these authors claim that this contrast correlates with the Null Subject Parameter.⁸

⁸ As mentioned above, in Section 5.4, I show that this correlation between the Null Subject Parameter and the possibility of RRC extraposition is simply not correct and must be revised.

- (72) a. A man arrived that wants to talk to you. [English]
 b. Un homme est arrivé qui veut te parler. [French]
- (73) a. *Un hombre apareció que dice que quiere hablar contigo. [Spanish]
 b. *Un home va venir que volia parlar amb tu. [Catalan]
 c. *Un uomo è arrivato che vuole parlarti. [Italian]

Dutch and German pattern with English and French with respect to this property, as illustrated in (74)-(75):

[Dutch]

- (74) Iemand heeft me een boek gegeven die ik niet ken.
someone has me a book given who I not know
 ‘Someone gave me a book who I do not know.’

(De Vries 2002: 244)

[German]

- (75) weil eine Frau gehustet hat, die mit einem Porsche kam
since a woman coughed has who with a P. came
 ‘since a woman coughed who came with a Porsche.’

(Meinunger 2000: 208)

C2. Restriction on extraposition from other pre-verbal constituents

As with CEP, some Germanic languages allow RRC extraposition from a *wh*-constituent (see (76) from English and (77) from Dutch), an emphatic/evaluative phrase (see (78) from English and (79) from Dutch), and a preposed focus (see (80) from English and (81) from Dutch).

[*wh*-constituent]

- (76) a. Who do you know that you can really trust?
 b. Which argument do you know that Sandy thought was unconvincing?

(Kiss 2002)

- (77) Hoeveel kinderen ken jij die niet van snoepjes houden?
how.many children know you that not of sweets like
 ‘How many children do you know that do not like sweets?’

(Smits 1988: 195)

[emphatic/evaluative phrase]

- (78) People loose their eyesight when they don't take support of the STD's and much more things can happen that are far worse than loosing your eye sight.

(<http://genital-herpes-warts.com/genitalherpes/genitalherpes-6555.html>)

- (79) Heel veel mensen hebben een verre reis geboekt die daar eigenlijk
very many people have a far trip booked who there actually
 niet het geld voor hebben.
not the money for have

‘A lot of people booked a long journey who in fact didn’t have enough money for it.’

(Mark de Vries, p.c.)

[preposed proposed focus]⁹

- (80) %Not even one painting did I see which would please Laura.

(Smits 1988: 195)

- (81) Alleen die bloemen kon hij benoemen, die zijn moeder hem vroeger
only those flowers could he identify that his mother him formerly
 had aangewezen
had pointed.out

‘Only those flowers could he identify, that his mother had once pointed out to him.’

(Smits 1988: 380)

Beatrice Santorini (p.c.) reports to me that RRC extraposition from emphatic/evaluative phrases and preposed foci is also attested in the diachrony of English. By way of illustration, see (82)-(84):

[emphatic/evaluative phrase]

- (82) Many more such worthie iniunctions and honourable ordinances I obserued, which are hardly worth pen and inke the describing

(PPCEME, JOTAYLOR-E2-P1,3,89.C2.411)

[preposed focus]

- (83) Two or three things I recollected when it was too late, that I might have told you;

(PPCMBE, AUSTEN-180X,171.245)

- (84) One thing I had almost forgot which the mention of the girls brought into my minde

(PPCEME, LOCKE-E3-P1,37.74)

⁹ Smits (1988) uses the symbol ‘%’ to indicate that it is a highly formal and marked construction.

However, just like CEP, English does not allow RRC extraposition from topics; see (85)c below:

- (85) a. I like micro brews that are located around the Bay Area.
 b. Micro brews that are located around the Bay Area, I like.
 c. *Micro brews, I like that are located around the Bay Area.

(Kiss 2002)

The same seems to be true of Dutch, as illustrated in example (86) (involving hanging topic left dislocation):¹⁰

- (86) *Die meisjes, ik ken ze niet die uit Lissabon komen.
those girls I know them not that from L. come

lit. 'Those girls, I don't know them, that are from Lisbon.'

(Mark de Vries, p.c.)

4.2.1. Summary

Although this overview has several limitations in terms of cross-linguistic coverage (because it primarily draws on data reported in the literature), it offers important empirical evidence showing that languages do not behave uniformly with respect to RRC extraposition.

One important conclusion that emerges from the data reported in Sections 4.1 and 4.2 is that CEP contrasts sharply with some Germanic languages (e.g., English and Dutch) as far as the properties of RRC extraposition are concerned. An overview of the contrasting properties is provided in Table 2. In this table, the use of a plus indicates that RRC extraposition can occur in the relevant context; a minus indicates that it cannot.

¹⁰ There is a possible terminological confusion here. Recall that there is a difference between the traditional notion of topicalization and the topic position in a cartographic sense (see Chapter 1, Section 4.3.3). Earlier claims (e.g., Smits 1988, De Vries 2002) that RRC extraposition can take place from a topic(alization) position (say, [Spec, CP]) must not be understood as extraposition from an aboutness topic. Rather, it concerns the extraposition from a constituent in first position. As shown in the main text, such constituents are always affected by focus in some way or another (e.g., *wh* and contrastive foci). Therefore, it may be better to speak of focalization rather than topicalization in these cases.

Table 2. A comparative approach to RRC extraposition (CEP vs. English and Dutch)

	CEP	English	Dutch
A. <i>Extraposition from strong noun phrases</i>	-	+	+
B. <i>Extraposition from embedded positions</i>	-	+	+
	subjects	+	+
C. <i>Extraposition from pre-verbal constituents</i>	<i>wh</i> -constituents	+	+
	emphatic/evaluative phrases	+	+
	preposed foci	+	+
	topics	-	-

Another interesting conclusion is that Romance languages do not behave in a uniform manner. On the basis of the limited data that I collected from the literature, it is possible to identify the following contrasts:

Table 3. A comparative approach to RRC extraposition in Romance languages

	CEP	Italian	Spanish	French
A. <i>Extraposition from strong noun phrases</i>	-	-	-	-
B. <i>Extraposition from pre-verbal subjects</i>	-	-	-	+

Strikingly, French exhibits a peculiar behavior. It contrasts with other Romance languages in allowing extraposition from a pre-verbal position (see Table 3), but it also contrasts with Germanic languages in not allowing extraposition from strong noun phrases. This seems to be a rather puzzling set of restrictions, but from this, it emerges (at least for now) that not all Romance languages behave equally and that there may be other factors that additionally contribute to the contrasts presented in Table 3.

Having made this excursus into the properties of RRC extraposition across languages, let us now turn to Portuguese and see whether RRC extraposition may also be subject to variation in the diachronic dimension.

4.3. Properties of RRC extraposition in earlier stages of Portuguese

In this section, I show that CEP and earlier stages of Portuguese behave differently with respect to RRC extraposition. The historical data from Portuguese that support this view are presented in A.-C. below.

A. *The definiteness effect*

Earlier stages of Portuguese pattern with CEP in allowing extraposition from weak noun phrases, as illustrated in (87)-(88):

[subject]

- (87) Junto das casas [...] sta hũa llata ante a porta que dara
near to.the houses is a ±grapevine before the door that give:FUT
 hũs anos pollos outros çinquo allmudes de vinho
some years by.the others five a. of wine

‘Near the houses there is a grapevine before the door that on average will give five *allmudes* [medieval agrarian measure] of wine.’

(Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1528)

[object]

- (88) e mãdo. huno casale ad Monasterium in quo morat Michael de souto
and leave:1SG an hamlet to monastery in which lives M. d. S.

‘and I leave an hamlet to the monastery in which Michael de Souto lives.’

(Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1252)

However, unlike CEP, earlier periods of Portuguese allow for extraposed RRCs with strong noun phrases as antecedents, as illustrated in (89)-(92):¹¹

[subject]

- (89) As chagas erã muytas de que se uertia muyta sangue.
the sores were many of that SE:CL shed a.lot.of blood

‘There were many sores from which a lot of blood was being shed.’

(Brocardo 2006; *Livro de Linhagens do Conde D. Pedro*; 14th century)

- (90) mas aquelle dia sem falha aveo que forom i todos
but that day without fail came that went there all

‘but the day everyone went there came without fail.’

(Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)

¹¹ Brucart (1999) reports that extraposition from strong noun phrases is also possible in earlier stages of Spanish (see (i)).

(i) Aquel decimos ser mejor médico, que mejor cura y más enfermos sana.
that say:1PL be:INF better doctor that better heals and more patients cures
 ‘We say that the better doctor is the one who heals (the diseases) better and cures more patients.’

(Luis de Granada, *Introducción al símbolo de la fe*, 1583, cited in Brucart 1999: 466)

- (91) de tal homẽ como aquel será que esta spada ha de trazer
of such.a man as that be:FUT that this sword will carry:INF
 ‘of such a man as the one who will carry this sword will be.’
 (Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)

[object]

- (92) «Vede lo escudo aqui que demandades».
see:IMP the shield here that look.for:2PL
 ‘See here the shield that you are looking for.’
 (Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)

B. Restriction on extraposition from embedded positions

Unlike CEP, earlier stages of Portuguese allow for extraposed RRCs with the object of a preposition as the antecedent, as illustrated in (93)-(94) (example (94) is repeated from (18) above).

- (93) e logo lhj abriu [de todo] mão que sseu era
and immediatly to.him:CL opened of everything hand that his was
 ‘and immediately he gave him (= lit. opened hand of) everything that he had.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1339)

- (94) que [en aquela hora] morrera en que el vira estando longe
that in that hour die:PMQP.3SG in that he saw be:GER away
 dele que lhi saira a alma do corpo.
from.him that to.him:CL fall.out:PMQP the soul of.the body
 ‘[and he realized that] he died in the hour in which he saw (being away from him) that his soul fell out of his body.’
 (Mattos e Silva 1989: 766; *Diálogos de São Gregório*; 14th century)

- (95) depois que [naquela hora] o seu filho ficara sen féver
after that in.that hour the his son stay:PMQP without fever
 en que hi o nosso Salvador e nosso meestre Jesu Cristo dissera
in that there the our Savior and our master J. C. say:PMQP
 que era são.
that was:3SG healed
 ‘[and the priest realized that it was true] after his son stopped having fever in that hour that our Savior and Master Jesus Christ said that he was healed.’
 (Mattos e Silva 1989: 766; *Diálogos de São Gregório*; 14th century)

- (96) que [de mui poucos] sabemos que bebessem vinho
that of very few know:1PL that drink:PRES.SUBJ wine
 ‘[The sobriety and moderation of our kings is so praised] that we know of very few ones who drank wine.’
 (TYC; Francisco Rodrigues Lôbo, *Côrte na Aldeia e Noites de Inverno*; 1619)

C. Restriction on extraposition from pre-verbal positions

C1. Restriction on extraposition from pre-verbal subjects

Earlier stages of Portuguese and CEP behave alike in allowing RRC extraposition from post-verbal subjects, as shown in (87) above. However, in contrast to CEP, earlier stages of Portuguese allowed for extraposed RRCs with pre-verbal subjects as antecedents, as illustrated in (97)-(102) ((98) is repeated from (19) above):

- (97) se Algê A eles veer que diga que llj eu
if someone to them come:FUT.SUBJ that says that to.him:CL I
 Alguna cousa diuíá
some thing owed
 ‘[And] if someone who says that I owed him something comes towards them...’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1275)
- (98) E mando que se outra mãda parecer que eu mândasse
and order:1SG that if another will appear:FUT.SUBJ that I order:IMPERF.SUBJ
 fazer ante dessta que quebre e nõ ualha
make:INF before this that±be.annulled:PRES.SUBJ. and not be.valid:PRES.SUBJ.
 ‘And, if another will appear that I ordered to be made before this one, I order it to be annulled and not valid.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1383)
- (99) que cayam. e cayades na pea que filhos e
that fall:PRES.SUBJ.3PL and fall: PRES.SUBJ.2PL in.the punishment that children and
 netos deuẽ a caer. que contra béenço de padre uéerem
grandchildren should A fall:INF that against blessing of father come:FUT.SUBJ
 ‘[and I order] that they and you receive the punishment that the children and grandchildren who go against their father’s blessing should receive.’
 (CIPM; *Os Doc. em Português da Chancelaria de Afonso III*; year 1278)

- (100) se alguu for asy de mia parte como d' estraya que a
if someone be:FUT.SUBJ either of my side as of strange that to
 uos queyra cõtrastar seya maldito
you want:PRES.SUBJ go.against:PRES.SUBJ be:PRES.SUBJ damn
 ‘if there is someone either from my side of from a strange side that wants to go
 against you, (I want him) damned.’
 (Maia 1986; *História do Galego Português*; year 1274)
- (101) Como Galuam se salvou e como a donzella disse que algũus
how G. escaped and how the damsel said that some
 a _____ creiam que a nom creiam.
her:CL believed that her:CL not believed
 ‘How Galuam escaped and how the damsel said that some (people) believed her that
 (actually) did not believed her.’
 (Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)
- (102) ca de muitos que ja i seerom nunca i tal foi que
because of many that already there were never there such was that
 i nom fosse morto
there not be:IMPERF.SUBJ killed
 ‘because among the many people who have been there, there was no one who has
 not been killed.’
 (Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)

C2. Restriction on extraposition from other pre-verbal constituents

Earlier stages of Portuguese pattern with CEP in allowing extraposition from a *wh*-constituent (see (103)-(104)), a preposed emphatic/evaluative phrase (see (105)), and a preposed focus (see (106)-(107)).

[*wh*-constituent]

- (103) Que caso pod'esse ser / em que tanto sopesais?
what case can.this be:INF in that so.much ±think:2PL
 ‘What case can this be that you think so much about?’
 (Camões 1999; Gil Vicente, *Processo de Vasco Abul*; year 1516)

- (104) Já sei que [...] me perguntará qual Mestre conheço eu
already know:1SG that me:CL ask:FUT which master know I
 que tenha toda esta erudição
that have:PRES.SUBJ. all this erudition
 ‘I already know that you will ask me which master I know that has all this erudition.’
 (TYC; Luís António Verney, *Verdadeiro Método de Estudar*; 1746)

[emphatic/evaluative phrase]

- (105) Muitos letrados sei eu (disse Solino) que não são moços
many lettered know I said S. that not are young.boys
 ‘I know many lettered men (said Solino) who are not young.’
 (TYC; Francisco Rodrigues Lôbo, *Côrte na Aldeia e Noites de Inverno*; year 1619)

[preposed focus]

- (106) pois d’el-rrei dom Fernando nêhũa coisa teemos que contar
because of.the.king D. F. none thing have:1PL that tell:INF
 até a morte d’este rrei dom Pedro.
until the death of.this king D. P.
 ‘because we do not have anything to tell about the king Dom Fernando until the death of the king Dom Pedro.’
 (Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)
- (107) Todos falam da economia, e pouca gente tenho visto que
all speak of.the economy and few people have:1SG seen that
 tenha uma idéia distinta desta ciência
have:PRES.SUBJ a idea clear of.this science
 ‘Everyone talks about the economy, but I have seen few people who have a clear idea about this science.’
 (TYC; Marquesa de Alorna, *Inéditos - Cartas e Outros Escritos*; 18th century)

By contrast, if a topic is involved, RRC extraposition does not seem to be possible in earlier stages of Portuguese, at least in the corpus inspected thus far.¹²

¹² I found only one example that could be taken as involving RRC extraposition from topic; see (i):

- (i) Esta barca onde vai agora / que assim está apercebida?
this boat where goes now QUE this.way is equipped
 (Camões 1999; Gil Vicente, *Auto da Barca do Inferno*; year 1517)

Note, however, that (i) may instead involve a coordinate clause, introduced by the coordinating conjunction *que*, meaning *since, as*: ‘Where does this boat go, as it is so well equipped?’ In this respect, it is also worth pointing out that Martins (2002) suggests that topicalization (as opposed to focalization) may not be a grammatical option in earlier stages of Portuguese.

C3. *New contexts of extraposition*

Aside from the contexts demonstrated thus far, there is another important source of RRC extraposition in the history of Portuguese that is not available in CEP: IP scrambling.¹³

IP scrambling is an optional syntactic process whereby a constituent scrambles past the verb. This is illustrated in the contrast provided in (108). In (108)a, the OV order involves a scrambled object (marked in boldface), whereas in (108)b, the VO order involves a non-scrambled object surfacing in its base-position.

- (108) a. sse pela uêtujra uos alguê **a dita** **vÿa** enbargar
if by chance you:CL someone the mentioned vineyard blocks:IMPERF.SUBJ
 ‘and if by chance someone blocks the vineyard from you’
- b. sse pela uêtujra uos alguê enbargar a dita vÿa
if by chance you:CL someone blocks IMPERF.SUBJ the mentioned vineyard
 (Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1296, cited in Martins 2002: 234)

According to Martins (2002), IP scrambling consists of the movement of various types of constituents to multiple Specifier positions available in the IP domain. This movement has a prosodic/discourse motivation; it allows the scrambled constituent to escape the default focus stress (and the information focus interpretation). Martins (2002) claims that the prosodic/discourse approach to scrambling explains why it imposes no restrictions on the categorial status of the scrambled constituent. By way of illustration, consider examples (109)-(111), where the scrambled constituent is (respectively) a PP, an adjectival phrase, and a past participle.

- (109) de quê lhe **ssobre elle** embargo poser
from whoever him:CL over it obstruction put:FUT.SUBJ.3SG
 ‘[protecting him] from whoever tries to block it [the land] from him.’
 (Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1509, cited in Martins 2002: 244)

¹³ Following common practice, I will distinguish two types of scrambling in this study: *short scrambling* (i.e., scrambling to VP) and *medial scrambling* (i.e., scrambling to IP (TP)). There is another type of scrambling (*long distance scrambling*, involving movement across a CP boundary), which will not be addressed here (see Takano 1998 and references therein).

- (110) *todollos adubiyos que lhes conpridoiros e neçsareos forem*
all.the fertilizers thatthem:CL due andnecessary be:FUT.SUBJ.3PL
 ‘[The renters will put on the land] all sorts of fertilizers that the land may need.’
 (Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1476, cited
 in Martins 2002: 245)
- (111) *com os lauradores que as ssemeadas teuerẽ*
with the farmers that them:CL cultivated have be:FUT.SUBJ.3PL
 ‘with the farmers who keep the lands cultivated’
 (Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1472, cited
 in Martins 2002: 245)

Crucially, IP scrambling in earlier stages of Portuguese can generate RRC extraposition, as illustrated in (112)-(113):

- (112) *que llj eu Alguna cousa diuía que nõ seia escripto*
that to.him:CL I some thing owed that not be:PRES.SUBJ written
en Esta mãda
in this will
 ‘(And if there arrives someone who says) that I owed him something which is not
 written in this will ...’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1275)
- (113) *E pera todas as cousas e cada hũa delas fferaser que uerdadeyro*
and to all.the things and each one of.them make:INF that real
e lĩdemo procurador pode e deue fferaser
and legitimate proxy can and should make:INF
 ‘And to make all the things and each one of them that a real and legitimate proxy
 can and should make.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1317)

The scrambling of *alguma coisa* ‘lit. some thing’ in (112) is confirmed by the relative position of this constituent with respect to the verb and the clitic. According to Martins (2002), clitics in clauses with interpolation set the border between left-dislocated/focused constituents and scrambled constituents. Hence, in (112), because *alguma coisa* ‘some thing’ is interpolated (i.e., occurs between the proclitic and the verb), it is necessarily a scrambled constituent.

4.3.1. Summary

This section has shown that earlier stages of Portuguese are less restrictive than CEP with respect to RRC extraposition. This can be seen in Table 4, which summarizes the relevant contrasting properties.

**Table 4. A comparative approach to RRC extraposition
(CEP vs. earlier stages of Portuguese)**

	CEP	Earlier stages of Portuguese
A. <i>Extraposition from strong noun phrases</i>	-	+
B. <i>Extraposition from embedded positions</i>	-	+
	subjects	+
	<i>wh</i> -constituents	+
C. <i>Extraposition from pre-verbal constituents</i>	emphatic/evaluative phrases	+
	preposed foci	+
	topics	-

On the basis of a cross-linguistic comparison, another relevant conclusion that can be drawn from this study is that earlier stages of Portuguese are (to a large extent) Germanic-like, unlike CEP, as shown in Table 5.

**Table 5. A comparative approach to RRC extraposition
(a cross-linguistic overview)**

	CEP	Earlier stages of Portuguese	English	Dutch
A. <i>Extraposition from strong noun phrases</i>	-	+	+	+
B. <i>Extraposition from embedded positions</i>	-	+	+	+
	subjects	+	+	+
C. <i>Extraposition from pre-verbal constituents</i>	<i>wh</i> -constituents	+	+	+
	emphatic/evaluative phrases	+	+	+
	preposed foci	+	+	+
	topics	-	-	-

The contrasts outlined in Table 5 are not accidental and clearly call for an explanation. In Section 4.2.1 above, the hypothesis was raised that some of these contrasts may correlate with the Null Subject Parameter. However, such a hypothesis must be discarded in the face of the data discussed in Section 4.2. Because Portuguese

has always been an NSL over the course of its history, the fact that earlier periods of Portuguese (unlike CEP) allowed for extraposition out of pre-verbal subjects shows that the contrasting properties cannot be explained via the Null Subject Parameter. I will return to this issue in Section 5.4.

Alternatively, in the next two sections, I suggest that the contrasts in Table 5 can be straightforwardly explained by a dual approach to RRC extraposition. In particular, I contend that the diachronic (and cross-linguistic) data considered thus far provide strong empirical evidence in favor of the hypothesis that different stages of the same language (and languages in general) may resort to different strategies of RRC extraposition. Specifically, I propose that extraposed RRCs in earlier stages of Portuguese differ in their structure and derivation from extraposed RRCs in CEP. The former are derived from specifying coordination (plus ellipsis) (De Vries 2002), whereas the latter are derived from stranding (Kayne 1994).

5. The analysis of RRC extraposition in CEP

In this section, I propose that the properties of RRC extraposition in CEP can be accounted for in terms of the stranding analysis proposed by Kayne (1994). Section 5.1 establishes the basic tenets of the stranding analysis of extraposition. It also introduces the key to the present proposal: RRC extraposition in CEP always involves leftward movement of the antecedent, either via movement to the left periphery or via short scrambling. Section 5.2 demonstrates how RRC extraposition in CEP can be derived from movement to the left periphery, and Section 5.3 shows how it can be derived from short scrambling. Then, Section 5.4 demonstrates how this theoretical apparatus accounts for the contrasting properties of RRC extraposition outlined in Section 4.1. Finally, in Section 5.5, I examine nine arguments that have been adduced in the literature against the stranding analysis, showing that they do not offer any insurmountable obstacle to the approach proposed here. Finally, Section 5.6 presents concluding remarks.

5.1. The stranding analysis

Following Kayne (1994) and Bianchi (1999), I assume that RRCs are generated by *head-raising*. As mentioned in Chapter 2 (Section 2.2.2), the main idea underlying

this proposal is that the head NP (the antecedent) of an RRC originates at the relativization site inside the subordinate clause and then rises to the left edge. The relative clause itself is generated as the complement of the so-called external determiner, with which the head NP associates after rising. A relative pronoun or operator is then to be analyzed as a *relative determiner* originally belonging to the internal head NP; see the representation in (114).

- (114) [DP D [CP [DP_{rel} NP [D_{rel} t_{NP}]] C [IP t_{DP}]]]
 e.g. *this* *book* *which* *I read*

From Kayne (1994), I adopted the idea that RRC extraposition is the result of VP-internal stranding. Under this approach, the antecedent is base-generated inside the RRC and undergoes leftward movement, stranding the RRC *in situ*, as schematically represented in (115):

- (115) Encontrei [uma pessoa]_i ontem [t_i que estava à tua procura]
met:1SG a person yesterday that was at.the your search

The key assumption of this proposal is the following: extraposed RRCs in CEP always involve the A'-movement of the antecedent, either via movement to the left periphery (when the antecedent is in a pre-verbal position) or via short scrambling¹⁴ (when the antecedent is in a post-verbal position).

I examine these two possibilities in greater detail in Sections 5.2 and 5.3, respectively. As the reader will notice, Section 5.3 requires more detail (and space) than Section 5.2. This is because the idea that the antecedent of an extraposed RRC undergoes movement to the left periphery is quite uncontroversial. By contrast, the idea that it undergoes short scrambling deserves a closer inspection and requires more complex explanatory devices. The question concerning the constituency of the dislocated constituent is reserved until Section 5.4.

¹⁴ In this dissertation, the term *short scrambling* refers to the scrambling to VP. See fn. 13 for additional details.

5.2. RRC extraposition and movement to the left periphery

In Section 4.1, I demonstrated that extraposed RRCs can take a *wh*-constituent, a preposed emphatic/evaluative phrase and a preposed focus as an antecedent. I repeat an example of each case here to illustrate the pattern.

[*wh*-constituent]

- (116) Quantas pessoas apareceram que não foram convidadas?
how.many people showed.up that not were invited
 ‘How many people showed up who were not invited?’

[emphatic/evaluative phrase]

- (117) Muito whisky o João bebeu que estava fora do prazo!
a.lot.of whisky the J. drank that was out of.the expiry.date
 ‘João drank a lot of whisky that was expired!’

[preposed focus]

- (118) Poucas pessoas conheço que fazem interpolação, mas todas elas
few people know:1SG that make interpolation but all they
produzem coisas deste tipo.
produce things of.this type
 ‘I know few people who make interpolation, but all of them produce things like this.’

In the literature on CEP, *wh*-constituents, emphatic/evaluative phrases and preposed foci have been argued to undergo leftward movement.

Wh-constituents are argued to move to the left periphery of the sentence, for instance, in Ambar (1992), Ambar et al. (1998), Ambar and Veloso (2001). Assuming a split CP system, Ambar and Veloso (2001) propose that there is a *wh*-projection in the left periphery of the sentence, which has strong N and V features. The raising of the *wh*-constituent is triggered by the need to check the N features, whereas the raising of the verb is triggered by the need to check the V features. Under this approach, a sentence like (116) (excluding the extraposed RRC) would be as in (119):

- (119) (TOP) [_{whP} quantas pessoas_i [_{wh'} apareceram_V [_{FP} [_{F'} t_V [_{IP} t_i t_V]]]]]]
how.many people showed.up

Emphatic/evaluative phrases are also taken to undergo leftward movement (see, e.g., Ambar 1999, Raposo 1995). In the syntactic representation of the sentence, Ambar

(1999) proposes that there is a projection called *EvaluativeP* sitting above IP but below CP, where Evaluative-like elements are licensed; see (120):

(120) [CP ... [EvaluativeP ... [TopicFocusP ... [IP ...]]]]

Under this approach, the features of E(valuative)P must be checked against evaluative features of lexical items. This explains why emphatic/evaluative phrases like *muito whisky* ‘a lot of whisky’ in (117) raise to [Spec, EvaluativeP].

Finally, Martins (in prep.) argues, in line with Hernanz and Bruccart (1987), Rizzi (1997), Cinque (1999) and related cartographic work, that preposed foci derive from movement. Under this analysis, *poucas pessoas* ‘few people’ in (118) is base-generated in a VP-internal position (as the complement of V) and undergoes movement to the left periphery.

It is worth noting that the exact landing site of the preposed constituents is not crucial here. The ideas outlined above are equally compatible with the existence of a functional projection in the CP domain dedicated to preposed foci or with analyses advocating a non-split CP domain.

For the current purposes, what is crucial is that *wh*-constituents, emphatic/evaluative phrases and preposed foci are base-generated not in the left periphery but instead in a VP-internal position. In other words, the relevant conclusion is that these constituents undergo leftward movement.

Turning now to the contexts of RRC extraposition, I submit that the RRC and its antecedent (in this case, a *wh*-constituent, an emphatic/evaluative phrase or a preposed focus) are base-generated within the RRC along the lines of the raising analysis. Then, these constituents undergo leftward movement, stranding the RRC *in situ*, as sketched in (121):

(121) a. ... *wh*-constituents_i ... [t_i RRC]
 b. ... emphatic/evaluative phrases_i ... [t_i RRC]
 c. ... preposed foci_i ... [t_i RRC]

I provide further details of the analysis in Section 5.4. For now, let us see how RRC extraposition from post-verbal positions can be accounted for under the stranding approach to extraposition.

5.3. RRC extraposition and short scrambling

This section aims to demonstrate that RRC extraposition from post-verbal positions can be accounted for in terms of short scrambling. Below, I begin by arguing that subjects in [Spec, VP], just like objects and subjects of unaccusative verbs, can scramble in CEP. I provide three arguments in favor of this hypothesis: adverb positioning, semantic effects and the trigger for scrambling (Section 5.3.1). Then, on the basis of the first two arguments, I show that RRC extraposition also involves short scrambling (Section 5.3.2). The trigger for scrambling in sentences involving RRC extraposition is discussed in Section 5.3.3. Finally, Section 5.3.4 demonstrates how to derive the occurrence of different constituents in the intervening position.

5.3.1. Subject and object scrambling in CEP

Costa (1998, 2004a) reports that CEP has a scrambling rule that allows objects to move from their base-position and adjoin to the VP. He also claims that the position of the scrambled object is indicated by its position relative to monosyllabic adverbs, such as *bem* ‘well’, which mark the left edge of the VP. The idea is that objects to the right of monosyllabic adverbs are in their base-position, whereas objects to the left of these adverbs are scrambled, as sketched in (122).¹⁵ This is illustrated in (123), taken from Costa (2004a: 40). In (123)a, the adverb-object order indicates that the object is not scrambled, whereas in (123)b, the object-adverb order indicates that the object is scrambled:

- (122) a. [IP V [VP Adv [VP t_v *non-scrambled constituent*]]]
 b. [IP V [VP *scrambled constituent* [VP Adv [VP t_v t_{obj}]]]]

- (123) a. O Paulo fala **bem** francês. [non-scrambled object]
 the P. speaks well French
 ‘Paulo speaks French well.’

¹⁵ In the present analysis, I assume (in line with Costa 1996, 1998, 2004a) that verbs move out of VP in CEP. Costa rejects Pollock’s (1989, 1994) analysis for French, according to which verbs may either stay inside VP or move up to Agr, depending on the occurrence of morphologically ambiguous forms (between a nominal and verbal interpretation). Such ambiguity simply does not arise with verbal forms, such as *falou* ‘spoke’ in (124), which is unambiguously a verbal form in the third person singular.

b. O Paulo fala francês **bem**. [scrambled object]
 the P. speaks French well

Costa also shows that objects are not the only constituents that may undergo scrambling. Indeed, subjects of unaccusatives can also scramble, as illustrated in (124), taken from Costa (2004a: 64). Here the adverb *depressa* ‘fast’ marks the left edge of the VP.

- (124) a. Chegou **depressa** o Paulo. [non-scrambled subject]
 arrived fast the P.
 ‘Paulo arrived fast’
 b. Chegou o Paulo **depressa**. [scrambled subject]
 arrived the P. fast

My claim is that the possibility of scrambling can be extended to subjects in [Spec, VP]. To my knowledge, this issue has not been previously addressed in the literature on CEP, but similar proposals have been discussed for other languages (e.g., Dutch/German and English).¹⁶ Hence, before proceeding with the argument, I will examine three cases that support this view.

Argument 1. *Distribution of adverbs*

A base-generated subject in [Spec, VP] may also surface in a post-verbal position, to the left of the monosyllabic adverb *bem* ‘well’, as illustrated in (125). If we maintain that (i) the monosyllabic adverb *bem* ‘well’ marks the left-edge of VP and (ii) the post-verbal subject is VP-internal (see Costa 1998, 2004a), then we must conclude that the subjects of unergative verbs can also scramble.

¹⁶ Broekhuis (2007) proposes that scrambling of objects and NP-movement of the subject in Dutch/German essentially involve the same operation, which he terms *subject/object shift*. Takano (1998) claims that English displays the short scrambling of accusative and nominative phrases.

- (125) A: a. Ninguém jogou nada.
nobody played nothing
 ‘No one played anything’
- B: b. Jogou o Sporting **bem** até aos últimos dez minutos.
played the S. well until the last ten minutes
 ‘Sporting played well until the last ten minutes. (Then Benfica reacted and scored two goals.)’

Argument 2. *Semantic effects*

When indefinite noun phrases are involved, the scrambled and non-scrambled orders can be semantically distinguished. More precisely, unscrambled indefinite objects may have a cardinal reading, whereas scrambled objects necessarily have a presuppositional reading (in the sense of Diesing 1992). For instance, see the contrast in (126). The unscrambled object in (126)a preferably has a cardinal, non-presuppositional reading. Under this interpretation, João can actually speak only one language. This contrasts with the scrambled order in (126)b. Here, the indefinite object can only have a presuppositional reading, which can be paraphrased as a partitive (‘one of the languages’).

- (126) a. O João fala **bem** uma língua.
the J. speaks well one language
 ‘João speaks one language well.’
- b. O João fala uma língua **bem**.
the J. speaks one language well
 ‘João speaks one language well (the other languages he speaks very badly).’

Now, consider the examples in (127) and (128), which involve the subject of an unaccusative verb and the subject of an unergative verb, respectively. As in the case of object scrambling, the subject may either precede or follow the adverb, but different semantic effects arise. For instance, in (127)a, the unscrambled subject preferably has a cardinal reading. Under this interpretation, the recipe requires only one kilo of potatoes. This contrasts with the scrambled order in (127)b, which necessarily involves a presuppositional reading. Under this interpretation, the recipe necessarily takes more than one kilo of potatoes. The same reasoning applies to (128).

Note that this is a welcome result; if scrambling is involved in (126)-(128), we expect that the same semantic effects will be obtained.¹⁷

[unaccusative verb]

(127) Context: recipe

a. Frite **bem** um quilo de batatas.
fry well one kilo of potatoes

‘Fry one kilo of potatoes well (=thoroughly).’

b. Frite um quilo de batatas **bem**.
fry one kilo of potatoes well

‘Fry one kilo of potatoes well (= thoroughly); (fry the other kilo of potatoes less).’

[unergative verb]

(128) a. Correram **bem** oito atletas.
ran well eight athletes

‘Eight athletes ran well.’

b. Correram oito atletas **bem**.
ran eight athletes well

‘Eight athletes ran well; (the other athletes did not run so well).’

Argument 3. *The trigger for scrambling*

It has been proposed in the literature that scrambling is movement to [Spec, AgrOP] driven by the requirement of accusative feature-checking (see De Hoop 1992, among others). Under this assumption, subject scrambling would be unexpected because the noun phrase in [Spec, VP] does not have an accusative feature to be checked by the complex V-Agr_o. Fortunately, this problem does not arise; Costa (1998, 2004a) shows that scrambling in CEP is not a case-driven movement.¹⁸ One of the arguments he provides in favor of this idea is precisely the possibility of subject scrambling (involving the subject of unaccusatives, as in (124)).

Alternatively, Costa (1998, 2004a) argues in favor of a prosodically/discourse-driven approach to scrambling, according to which scrambling

¹⁷ For similar semantic effects in object/subject shifts in German/Dutch, see Broekhuis (2007).

¹⁸ Also, Broekhuis (2007) does not assume case as the trigger for scrambling in Dutch/German (contra De Hoop 1992).

is used to create appropriate (information) focus configurations (e.g., to make the element bearing the sentence's nuclear stress escape it, see Reinhart 1995). Under this approach, the scrambling of subjects can be treated on par with that of objects. If scrambling is prosodically/discourse-driven, there is a priori no reason to block the scrambling of a constituent in [Spec, VP].

In summary, this section demonstrates that the possibility of short scrambling in CEP is not confined to objects and subjects of unaccusative verbs. On the basis of evidence from adverbs, semantic effects and the trigger of scrambling, it was argued that subjects in [Spec, VP] might also scramble. With these ideas in mind, let us examine how RRC extraposition from post-verbal positions can be accounted for in terms of short scrambling.

5.3.2. The involvement of short scrambling in RRC extraposition

The idea that RRC extraposition from post-verbal positions involves short scrambling is supported by two different arguments:

Argument 1. *Distribution of adverbs*

The antecedent of an extraposed-RRC may appear to the left of the monosyllabic adverb *bem* 'well', as illustrated in (129). Under the assumption that the monosyllabic adverb *bem* 'well' marks the left edge of the VP, the position of *uma candidatura* 'one application' indicates that this constituent has undergone short scrambling.

(129) A: a. Não analisaste com atenção nenhuma candidatura.

'You did not analyze any of the applications carefully.'

B: b. Analisei uma candidatura **bem** que foi proposta pela
analyzed:1SG one application well that was submitted by.the
 Universidade de Lisboa.
University of L.

'I analyzed one application that was submitted by the University of Lisbon well (= thoroughly); (the others I actually did not analyze very carefully).'

Argument 2. Semantic effects

When the antecedent of a non-extraposed RRC is indefinite, it may have a cardinal reading. However, when extraposition is involved, the antecedent necessarily has a presuppositional reading. This is illustrated in (130) below. The non-extraposed version in (130)a is compatible with the reading that there is only one homeless person in my neighborhood, whereas the extraposed version in (130)b necessarily presupposes that there is more than one homeless person in my neighborhood. The same reasoning applies to (131). The similar behavior of the antecedent of RRCs and scrambled indefinite constituents (see Section 5.3.1 above) suggests that in both cases, the indefinite noun phrase is scrambled.

- (130) a. Há no meu bairro um sem-abrigo que não pede dinheiro.
has in.the my neighborhood one homeless that not asks money
 ‘There is one homeless person in my neighborhood that does not ask for money.’
- b. Há um sem-abrigo no meu bairro que não pede dinheiro.
has a homeless in.the my neighborhood that not asks money
 ✓ Reading 1: (*presuppositional*) There is more than one homeless person in my neighborhood (but only one does not ask for money).
 * Reading 2: (*cardinal*) There is only one homeless person in my neighborhood (and he does not ask for money).
- (131) a. Apareceu no meu gabinete um aluno que precisava de ajuda.
showed.up in.the my office a student that needed of help
 ‘One student showed up in my office that needed help.’
- b. Apareceu um aluno no meu gabinete que precisava de ajuda.
showed.up a student in.the my office that needed of help
 ✓ Reading 1: (*presuppositional*) More than one student showed up in my office (but only one needed help).
 * Reading 2: (*cardinal*) Only one student showed up in my office (and he needed help).

These two tests consistently indicate that the antecedent of an extraposed RRC behaves like a scrambled object/subject with respect to adverb positioning and semantic effects. The trigger for scrambling in sentences involving RRC extraposition deserves more detailed consideration and is therefore discussed separately in the next section.

5.3.3. The trigger for short scrambling in RRC extraposition

Costa (1998, 2004a) claims that short scrambling is prosodically/discourse-driven in CEP. According to this view, scrambling creates appropriate focus configurations, removing unfocused material from the focus domain. In Section 5.3.3.1, I show that this approach works well for sentences where only one constituent is assigned narrow information focus¹⁹ but may be seen as problematic for other discourse contexts (e.g., sentences where more than one constituent is assigned narrow information focus and sentences displaying broad information focus). My proposal is that in these discourse contexts, scrambling may occur within a focus domain to assign more prominence in the discourse to the constituent that stays in the rightmost position. With this background in mind, Section 5.3.3.2 shows that a similar pattern is observed in RRC extraposition from post-verbal positions. Also in this case, the antecedent is contained within the focus domain and undergoes scrambling to assign more discourse prominence to the RRC that stays in the rightmost sentential position.

5.3.3.1 Short scrambling and information structure in CEP

As already mentioned, Costa (1998, 2004a) argues in favor of Reinhart's (1995) prosodic/discourse explanation of scrambling. The basic idea is that information structure has an impact on word order; the assignment of narrow information focus drives the constituent expressing new information to the rightmost position of the sentence, where it receives the sentence nuclear stress.²⁰ Scrambling is then used to create appropriate focus configurations by allowing some constituents to escape the position where sentence nuclear stress is assigned.

This approach accounts for sentences where only one constituent is assigned narrow information focus, as in (132). Here, the adverb is expected to occur in the rightmost position because it is the new information requested in the question. Hence, the object undergoes scrambling, being defocused, and the adverb receives the default stress.

¹⁹ For the opposition between broad and narrow focus, see Chapter 1 (Section 4.3).

²⁰ The *Nuclear Stress Rule* assigns prominence to the rightmost/lowest constituent of the sentence, as proposed in Zubizarreta (1998, 1999). See Chapter 1 (Section 4.3.1) for additional details.

- (132) A: a. Como é que o Paulo fala francês?
how is that the P. speaks French
 ‘How does Paulo speak French?’
- B: b. O Paulo fala francês **bem**. [scrambled object]
the P. speaks French well
 ‘Paulo speaks French well.’
- c. #O Paulo fala **bem** francês. [non-scrambled object]
the P. speaks well French
- (Costa 2004a: 68)

By contrast, if the object is questioned (as in (133)a), it must stay in the rightmost position and get the default stress. Therefore, the non-scrambled order in (133)c is derived.

- (133) A: a. O que é que o Paulo fala bem?
the what is that the P. speaks well
 ‘What does Paulo speak well?’
- B: b. #O Paulo fala francês **bem**. [scrambled object]
the P. speaks French well
- c. O Paulo fala **bem** francês. [non-scrambled object]
the P. speaks well French
- (Costa 2004a: 68)

This approach straightforwardly applies to sentences in which only one constituent is assigned narrow information focus. However, it is problematic when more than one constituent is contained in the focus domain. A typical discourse context allowing for the occurrence of double-focus sentences is given in (134)-(135) (see Costa 2004a: 80, for similar tests). Under the discourse set by (134)-(135), the constituents in boldface are assigned narrow information focus.

(134) Context: João took an exam that did not go well. His colleagues are commenting on what might have gone wrong.

A: a. O João não leu nada com atenção.
the J. not read nothing with attention

‘João did not read anything with care.’

B: b. O João leu **Os Maias** bem.
the J. read O. M. well

c. #O João leu bem **Os Maias**.
the J. read well O. M.

‘João read *Os Maias* [a book title] well. (I am sure of that because I have studied with him).’

(135) Context: Animal race in the forest

A. a. Os animais da floresta chegaram todos à meta muito devagar.
the animals of the forest arrived all to the finishing line very slowly

‘The animals of the forest took too long to cross the finishing line.’

B. b. Chegou **o coelho/um coelho** depressa.
arrived the rabbit a rabbit fast

c. #Chegou depressa **o coelho/um coelho**.
arrived fast the rabbit a rabbit

‘The rabbit/a rabbit arrived fast. (But the others arrived too slowly).’

Costa (2004a: 86) assumes that if a sentence has more than one focus (as in the VSO contexts), the leftmost focused constituent bears heavy stress. Then, all constituents following the heavy stress are interpreted as focus. This leads to the representations in (136), where *Os Maias* [book title] and *o coelho/um coelho* ‘the rabbit/a rabbit’ bear heavy stress (as indicated by the capital letters), and the adverb to their right is interpreted as focus.

(136) a. O João leu **OS MAIAS** bem.
the J. read O. M. well

b. Chegou **O COELHO/UM COELHO** depressa.
arrived a rabbit the rabbit fast

An unexpected implication of this analysis is that a scrambled constituent appears in the focus domain. In (134)b-(135)b, the object and the adverb are assigned narrow information focus, but the object is scrambled anyway, as confirmed by its relative position with respect to the adverb. Initially, the idea that a constituent interpreted as focus may undergo scrambling is surprising, under the assumption that scrambling serves to remove unfocused material from the focus domain. However, this analysis

receives some typological support from the so-called *focus-scrambling* in Dutch, which involves contrastive focus on a scrambled constituent (see Costa 2004a: 69).²¹

Additional evidence for the correlation between scrambling and focus comes from sentences involving broad information focus, as in (137):

(137) Context: Maria was expected to have a risky childbirth because she was going to have triplets.

A: a. Como correu o parto da Maria?
how went the labor of.the M.

‘How did the childbirth go?’

B. b. Mais ou menos. Nasceram dois bebés bem.
more or less were.born two babies well

‘Well and not so well. The birth of the first two babies went well. (It was the birth of the third baby that was more complicated).’

Despite the fact that the focus extends to the entire sentence in (137)b (*Nasceram dois bebés bem*), the constituent *dois bebés* ‘two babies’ is scrambled, as can be confirmed by its occurrence to the left of the adverb *bem* ‘well’.

The question that then arises is why a constituent undergoes scrambling within a focus domain. I would like to submit that in the context of double-focus or broad information focus, the constituent in the rightmost position receives more discourse prominence than the other constituents. Thus, scrambling can be used to create specific discourse effects (namely, to place the most prominent constituent in the rightmost position within the clause-internal space).

Let us examine exactly how this works in a sentence like (137)b. The constituent *dois bebés* ‘two babies’ is contained in a sentence with broad information focus. Nevertheless, *dois bebés* ‘two babies’ conveys less prominent information than *bem* ‘well’. There are two reasons for why this occurs. First, it is expected that during childbirth a baby is born. Second, *bem* ‘well’ is a direct response to *como* ‘how’ in

²¹ This construction is exemplified in (i), taken from Costa (2004a: 69):

(i) Jan zei dat ik DE KRANT gisteren las, (en het boek vandaag).
J. said that I the newspaper yesterday read, and the book today

‘Jan said that it was the newspapers that I read yesterday (and not the books today).’

(137)a. Being less prominent, *dois bebés* ‘two babies’ undergoes short scrambling, leaving the adverb *bem* ‘well’ in the rightmost position.

Additional evidence supporting this approach comes from broad information focus sentences exhibiting the *adverb - scrambled object* order. As shown in (138), a sentence displaying the *bem* ‘well’ - *dois bebés* ‘two babies’ order cannot constitute an appropriate answer to the question: ‘How did the childbirth go?’

(138) Context: Maria was expected to have a risky childbirth, because she was going to have triplets.

A. a. Como correu o parto da Maria?
how went the labor of.the M.

B. b. #Mais ou menos. Nasceram bem dois bebés.
more or less were.born well two babies

This can be explained by assuming that a constituent in the rightmost position tends to convey non-discourse-dependent (or non-presuppositional) information. Such a requirement is not fulfilled in (138)b because a less prominent constituent, expressing the fact that two babies were born, appears in the rightmost sentential position.²²

In summary, the data considered thus far indicate that scrambling may take place within a focus domain to assign discourse prominence to the rightmost sentential constituent. This phenomenon can be observed in broad information focus sentences or in sentences where more than one constituent is assigned narrow information focus.

5.3.3.2 Short scrambling and information structure in RRC extraposition

Several studies have suggested that there is a relationship between RC-extraposition and discourse/information structure. For instance:

This supports our contention that the extraposed relative clause is understood to be main assertion (...) and that the main clause (...) is taken to be background information.

(Ziv and Cole 1974: 775)

²² It should be noted that the constituent *dois bebés* ‘two babies’ in (138)b could only be interpreted as referring not to Maria’s babies but to other babies (out of many that were born, for instance, on the same day in the hospital). This is because the constituent in the rightmost position tends to convey non-discourse-dependent (or non-presuppositional) information.

The function of unextraposed restrictive relative clauses is identifying: the relative clause provides information by means of which the referent of the NP is identified or described. (...) The function of extraposed relative clauses is assertative. Extraposed relative clauses assertate some attribute of an individual who is already adequately identified.

(Ziv and Cole 1974: 773)

Extrapolation from NP is allowed only when the element extraposed to sentence-final position is interpreted as being more important than the rest of the sentence.

(Takami 1998: 27)

Extraposed REL-clauses *can* be used as a presentative device, especially when modifying REF-indefinite nouns.

(Givón 2001: 210)

This construction [extraposed RRC] is most likely when the informational content of the relative clause is greater than that of the material that would follow it in the matrix clause if it occupied the default position following the antecedent.

(Huddleston, Pullum, and Peterson 2002: 1066)

Although these views are not in complete agreement with one another, they clearly indicate that discourse and information structure play an important role in RC-extrapolation (for the same view on extrapolation from PPs, see Guéron 1980). In this section, I demonstrate that this idea is corroborated by the RRC extrapolation in CEP.

In CEP, extraposed RRCs may occur in two different contexts: (i) sentences displaying broad information focus; and (ii) sentences where more than one constituent is assigned narrow information focus. The first possibility is illustrated in (139)d-(140)a below, where the entire sentences (containing an extraposed RRC) express new information.

(139) Context: Two friends are talking about the conflict in the Middle East. They are concerned about the rising number of deaths.

A: a. Acho que o acordo de paz está novamente em risco.

‘I think that the peace deal is at risk again.’

B: b. A sério? Porque é que dizes isso?

‘Really? Why do you say so?’

A: c. Não sabes o que aconteceu?

‘Don’t you know what happened?’

d. Ontem explodiu uma bomba em Israel que causou mais 5 mortos.
yesterday exploded a bomb in I. that caused more 5 deaths

‘Yesterday a bomb exploded in Israel that caused more five deaths.’

(140) Context: One month before the final General Staff meeting, teachers have to inform students about their predicted grade for the end of semester. At the classroom, one teacher says to his students:

a. Há muitos alunos nesta turma que provavelmente vão chumbar.
has many students in.this class that probably go fail:INF

‘There are many students in this class that will probably fail.’

The possibility of having an extraposed RRC in sentences where more than one constituent is assigned narrow information focus is illustrated in (141)-(142). In this case, note that the antecedent, the intervening material, and the extraposed RRC express new information.

(141) Context: A hotel room was burgled. The cops arrived at the hotel and started questioning one of the employees.

A: a. Chegou alguém nos últimos dias que tivesse tido
arrived someone in.the last days that have:IMPERF.SUBJ had

um comportamento estranho?
a behavior strange

‘Have someone arrived in the last days who behaved in a strange way?’

B: b. Chegou um rapaz ontem que estava muito nervoso ...
arrived a boy yesterday that was very nervous.

‘A boy arrived yesterday who was very nervous (but I don’t think he is involved in the burglary)’.

(142) Context: Two friends are speaking about João, an art collector, who usually buys very expensive things.

A: a. O que é que o João comprou desta vez?
the what is that the J. bought of.this time

‘What did João buy this time?’

B: b. O João comprou uma escultura em Londres que vale mais de
the J. bought a sculpture in L. that is.worth more of
 100 000 dólares.
100,000 dollars

‘João bought a sculpture in London that is worth more than 100,000 dollars.’

What all of these constructions have in common is that the antecedent and the extraposed RRC are interpreted as the information focus. Hence, the generalization that captures the relationship between RRC extraposition from post-verbal positions and information focus in CEP can be formulated as in (143):

(143) *Relation between RRC extraposition from post-verbal positions and information structure (in CEP)*

The extraposed RRC and the antecedent must be contained in the same focus domain.

Contrary to some initial impressions, the fact that extraposition from post-verbal antecedents involves short scrambling (see Section 5.3) is not incompatible with the generalization in (143). As is clear from the discussion in Section 5.3.3.1, there is no conflict between focus and scrambling; scrambling may occur within a focus domain to assign discourse prominence to the rightmost sentential constituent.

Extending this insight to RRC extraposition, I would like to suggest that the scrambling of the antecedent has the effect of assigning discourse prominence to the RRC that is stranded in the rightmost sentential position. This hypothesis is corroborated by some of the ideas quoted above, namely the one that states that extraposition *is allowed only when the element extraposed to sentence-final position is interpreted as being more important than the rest of the sentence* (Takami 1998: 27).

From the considerations thus far, it follows that RRC extraposition cannot be defined as a purely syntactic phenomenon. The discourse-based approach proposed

here clearly indicates that if an RRC is extraposed, the information conveyed by the RRC is interpreted as being more prominent than the rest of the sentence.

Generalizing from this and similar data, it is possible to formulate a hypothesis regarding extraposition in general. It may be the case that the extraction of the first part of a duplex construction (in De Vries' 2002 terms) has a discourse effect that is independent of focus, i.e., that of assigning discourse prominence to the second part that is left behind. However, further research is necessary in this domain to warrant the validity of this suggestion.

5.3.4. Deriving the intervening material

Thus far, I have provided evidence for the idea that RRC extraposition in CEP may take a scrambled constituent as an antecedent. In this section, I submit that such a syntactic configuration is derived as follows: (i) the antecedent is generated together with the RRC; and (ii) the antecedent undergoes short scrambling and adjoins to the VP after raising, stranding the RRC *in situ*. This is sketched in (144)a; in example *a.*, the antecedent and the RRC are generated together in the subject position ([Spec, VP]),²³ and in example *b.*, these elements are generated in the complement position of V.

- (144) a. [_{IP} V [_{VP} S [_{VP} intervening material [_{VP} t_s RRC t_v]]]]
 b. [_{IP} V [_{VP} DO [_{VP} intervening material [_{VP} t_v t_{DO} RRC]]]]]

(144)a schematically represents an extraposed RRC with the subject of an unergative verb as an antecedent (corresponding to a sentence as (145) below). (144)b represents an extraposed RRC with a direct object (or the subject of an unaccusative verb) as an antecedent (corresponding to (146) and (147) below). Examples (145)-(147) are repeated here for ease of exposition.

²³ In this portion of the discussion I will abstract away from the assumption that there are two verb phrases in the clause, a vP and a VP (see Larson 1988, 1990), and for ease of exposition, I will represent the double VP-shell as a single VP-shell. The double VP-shell will be introduced only while discussing the syntax of double complement constructions.

- (145) Telefonou um rapaz ontem que queria informações sobre a tua casa.
phoned a boy yesterday that wanted details about the your house
 ‘A boy phoned yesterday who wanted details about your house.’
- (146) Encontrei um rapaz no cinema que perguntou por ti.
met:1SG a boy at.the cinema that asked for you
 ‘I met a boy at the cinema that asked for you.’
- (147) Chegou um senhor ontem que fez muitas perguntas sobre ti.
arrived a man yesterday that made many questions about you
 ‘A man arrived yesterday who asked many questions about you.’

Now, I must identify the elements that can (and cannot) occur as intervening material and demonstrate how the analysis proposed here can accommodate the various possibilities. Section 5.3.4.1 shows how to derive the occurrence of adverbs and PPs in the intervening position, and Section 5.3.4.2 explains why other constituents (e.g., the subject) cannot occur in the intervening position.

5.3.4.1 Deriving the occurrence of adverbs and PPs in the intervening position

When short scrambling is involved, only adverbs and PPs can intervene between the antecedent and the extraposed RRC. This is illustrated in (148)-(149) (repeated from (23) and (37) above) and in (150)-(151) below:²⁴

[subject]

- (148) Chegou um rapaz ontem que te quer conhecer.
arrived a boy yesterday that you:CL wants meet:INF
 ‘A boy arrived yesterday that wants to meet you.’
- (149) Ontem explodiu uma bomba em Israel que causou 5 mortos.
yesterday exploded a bomb in I. that caused 5 deaths
 ‘Yesterday a bomb exploded in Israel that caused 5 deaths.’

²⁴ Note that more than one adjunct may co-occur as intervening material, as in (41) above.

[object]

- (150) Encontrei uma rapariga ontem que perguntou por ti.
met:1SG a girl yesterday that asked for you

‘I met a girl yesterday that asked for you.’

- (151) Comprei uma boneca na feira de artesanato que é feita de
bought:1SG a doll at.the fair of craft that is made of
 pasta de papel.
paste of paper

‘I bought a doll at the craft fair that is made of paper paste.’

These intervening elements can either be modifiers (as in (148)-(150) above) or arguments of the verb (as in (152)-(153) below).

[indirect object]

- (152) Dei um livro à Maria que foi escrito por mim.
gave:1SG a book to.the M. that was written by me

‘I gave Maria a book that was written by me.’

[prepositional argument]

- (153) Deixei um recado em cima da mesa que é para a Rita.
left:1SG a message on top of.the table that is for the R.

‘I left a message on the top of the table that is for Rita.’

The derivation of contexts involving modifiers and arguments as intervening material is treated separately in A and B below.

A. Modifiers in the intervening position

Let us begin by examining the occurrence of adverbs in the intervening position. Currently, there is no consensus in the literature regarding the syntactic representation of adverbs. Broadly speaking, two major lines of research can be identified: the adjunction analyses and the functional Specifier analyses. The adjunction analyses claim that adverbs are adjoined to some projection (VP, IP, ...) (see, among others, Ernst 2002, Costa 1998, 2004a, 2004b). The functional Specifier analyses assume that adverbs occupy non-argumental Specifier positions and are licensed in a Spec-head configuration with respect to a head containing semantic features related to, e.g., mood, tense and aspect (see Cinque 1999, Alexiadou 1997). Partially related to this issue, the analyses available in the literature may also manifest divergence with

respect to the distribution of adverbs. Some linguists claim that adverbs are freely distributed within a sentence (see, e.g., Emonds 1976), whereas others point out that the distribution of adverbs is very restricted (see, e.g., Cinque 1999).

For reasons of overall coherence (namely, with respect to Costa's 1998, 2004a approach to short scrambling and to the tests used here to identify scrambled constituents), I assume a left-adjunction analysis of the adverbs that surface in the intervening position. Nevertheless, I leave the hypothesis open that the approach developed here may also be compatible with a functional Specifier analysis of adverbs.

Turning now to the analysis proper, consider (154) (taken from Costa 2004a: 6), which shows that an adverb in CEP may either precede or follow a verb.

- (154) a. O João **ontem** leu o livro.
 the J. yesterday read the book
 'João read the book yesterday.'
 b. O João leu **ontem** o livro.
 the J. read yesterday the book

Following Costa (2004a), I maintain the assumption that the adverb *ontem* 'yesterday' in (154) is left-adjoined to different projections. In (154)a, the adverb is left-adjoined to TP (see (155)a), and in (154)b, it is left-adjoined to VP (see (155)b).²⁵

- (155) a. [_{AgrSP} O João [_{TP} *ontem* [_{TP} leu [_{VP} t_v o livro]]]]
 b. [_{AgrSP} O João [_{TP} leu [_{VP} *ontem* [_{VP} t_v o livro]]]]

(Costa 2004a: 7)

I also assume, along with Costa 2004b, that adverbs only adjoin to the left. With this background in mind, I submit that if RRC extraposition involves an adverb in the intervening position, the adverb is left-adjoined to VP, as in (155)b. Then, the object/subject scrambles over the adverb, deriving the *antecedent-adverb-RRC* order represented in (156):

²⁵ Recall that CEP displays V-to-I movement, which derives the order Verb – Adverb/Object in (155)b (see also Section 5.3.1, fn. 15).

- (156) a. [_{IP} V [_{VP} DO [_{VP} adverb [_{VP} t_v t_{do} RRC]]]]
 b. [_{IP} V [_{VP} S [_{VP} adverb [_{VP} t_s RRC t_v]]]]

In more concrete terms, what this means is that the source structure of an extraposed RRC taking an object as antecedent is as depicted in (157). Then, if the antecedent of the RRC undergoes short scrambling (stranding the RRC *in situ*), the extraposed order in (158) is derived.

- (157) O João comprou ontem um portátil que custou 1000 euros.
the J. bought yesterday a laptop that cost 1000 Euros
 ‘Yesterday, João bought a laptop that cost 1,000 Euros’.

- (158) O João comprou um portátil ontem que custou 1000 euros.
the J. bought a laptop yesterday that cost 1000 Euros
 ‘João yesterday bought a laptop that cost 1,000 Euros’.

Let us consider now the occurrence of modifying PPs in the intervening position. There is an ongoing debate in the literature about the way in which modifying PPs integrate into the structure of the clause. Broadly speaking, the syntactic analysis of modifying PPs can be divided in three major groups: adjunction analyses, Larsonian analyses and Specifier analyses.

Adjunction analyses assume that modifying PPs are adjoined to VP. Two variants of this approach can be identified: modifying PPs can be taken to involve right-hand adjunction (Chomsky 1981) or left-hand adjunction (Barbiers 1995). Larsonian analyses neutralize the structural distinction between arguments and modifiers, claiming that modifying PPs are base-generated below the arguments of the verb as complements of V (see Larson 1988, 1990, Chomsky 1995: 333). Specifier analyses claim that modifying PPs (and arguments) are all merged in Spec positions in a strict order, with the verb in the innermost position; a different order of constituents may be derived by successively moving larger and larger constituents containing the VP into higher Specs (Cinque 2006).

The analyses proposed in the literature may also differ in the way that they account for the complements/modifying PPs order. Some approaches claim that there is a unique (and universal) order of merge between these constituents (Cinque 2006), whereas others claim that these constituents do not enter the derivation in a strict order (Jackendoff 1990).

In this study, I assume (in line with Barbiers 1995) that modifying PPs that surface in the intervening position are left-adjoined to the VP (just like intervening adverbs). Therefore, the derivation proceeds in the same way as described for adverbs: the antecedent raises leftward past the intervening PP and adjoins to VP, stranding the RRC *in situ*. This is presented in (159)a-b, where the extraposed RRCs have, respectively, an object and a subject as an antecedent.

- (159) a. [IP V [VP DO [VP modifying PP [VP t_v t_{do} RRC]]]]
 b. [IP V [VP S [VP modifying PP [VP t_s RRC t_v]]]]

Therefore, the source structure of an extraposed RRC with a subject as an antecedent corresponds to a sentence like (160), where the modifying PP is left-adjoined to VP, and the subject is in its base-position. Then, if the antecedent of the RRC undergoes short scrambling, stranding the RRC *in situ*, the extraposed order in (161) is derived.

- (160) Ontem explodiu em Israel uma bomba que causou 5 mortos.
yesterday exploded in I. a bomb that caused 5 deaths
 ‘Yesterday a bomb exploded in Israel that caused 5 deaths.’

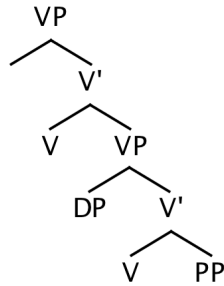
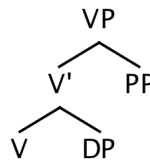
- (161) Ontem explodiu uma bomba em Israel que causou 5 mortos.
yesterday exploded a bomb in I. that caused 5 deaths

B. Complements in the intervening position

In double complement constructions, the PP may appear in the intervening position, as illustrated in (162) (repeated from (152)):

- (162) Dei um livro à Maria que foi escrito por mim.
gave:1SG a book to.the M. that was written by me
 ‘I gave Maria a book that was written by me.’

Although double complement constructions have received much attention in the generative literature, their exact status remains controversial (see, e.g., Kayne 1984, Larson 1988, Pesetsky 1995, Philips 1996). Indeed, one point of disagreement concerns the choice between the shell structure represented in (163)a and the layered structure represented in (163)b:

(163) a. *shell structure*b. *layered structure*

As noted in Philips (1996) and Costa (2004a), part of the debate results from the fact that the tests applied to these constructions yield contradictory results. For instance, (164) provides evidence for an analysis as in (163)a under the assumption that Licensing of Polarity Items requires c-command. In turn, (165)c provides evidence for a layered structure as in (163)b, because *give candy* is a constituent in (163)b but not in (163)a.

(164) a. John gave nothing to any of the children on his birthday.

b. *John gave anything to none of the children on his birthday.

(Costa 2004a: 144)

(165) John intended to give candy to children on his birthday.

a. ... and [give candy to children on his birthday] he did.

b. ... and [give candy to children] he did on his birthday.

c. ... and [give candy] he did to children on his birthday.

(Costa 2004a: 144)

Costa (2004a) additionally shows that binding facts suggest that the PP-DP order cannot be derived from the base DP-PP order through scrambling of the PP to the left of the DP. This is due to the fact that the PP can bind an anaphor contained in the DP (see (166)), which suggests that it occupies an A-position.

(166) A: a. A quem é que deste os livros?
 to whom is that gave:2SG the books

‘To whom did you give the books?’

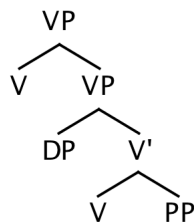
B: b. Dei [F A CADA AUTOR] o seu livro.
 gave:1SG to each author the his book

‘I gave his book to each author.’

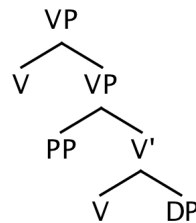
(Costa 2004a: 143)

Based in part on the facts mentioned above, Costa (2004a) (in line with Philips 1996) suggests that both the DP-PP and PP-DP orders can be base-generated in CEP. To keep the discussion simple, I will abstract away from the technical implementation of the analysis (see Costa 2004a for additional details), and I will simply refer to the two final structures;²⁶ see (167)a-(167)b:

(167) a. DP-PP order



b. PP-DP order



Crucially, Costa argues that the fact that both word orders are base-generated does not entail that they should be optional. According to his proposal, the structure in (167)b is only generated if necessary for satisfying binding requirements or any other constraint forcing the PP-DP order, such as heaviness.²⁷

²⁶ Importantly, the structures in (167)a and (167)b are apparently similar to the VP-shell structure represented in (163). However, (167)a and (167)b are derived by building a right-branching phrase marker from left to right (see Philips 1996).

²⁷ Without going into further detail, please note that under this approach, the constituency problem in (165)c is derived by the possibility of targeting a step of the V-DP-PP derivation in which V and DP form a VP, an option that is available under the right-branching structures (see Costa 2004a: 148).

As its point of departure, the analysis that I will propose here adopts Costa's (2004a) claim that the PP-DP order can be base-generated in CEP.²⁸ It also takes from Costa (2004a) the idea that heaviness factors may legitimize this configuration. With this in mind, let us briefly consider how a sentence like (168) (containing an extraposed RRC and a PP-complement in the intervening position) can be derived.

- (168) Dei um livro à Maria que foi escrito por mim.
 gave:ISG a book to.the M. that was written by me
 'I gave Maria a book that was written by me.'

First, I assume that (168) is derived from the PP-DP base-order (as in (169)). This is due to heaviness effects: the DP *um livro que foi escrito por mim* 'a book that was written by me' is heavier/longer than the PP *à Maria* 'to Maria' and therefore surfaces in the rightmost sentential position.²⁹

- (169) Dei à Maria um livro que foi escrito por mim.
 gave:ISG to.the M. a book that was written by me
 'I gave Maria a book that was written by me.'

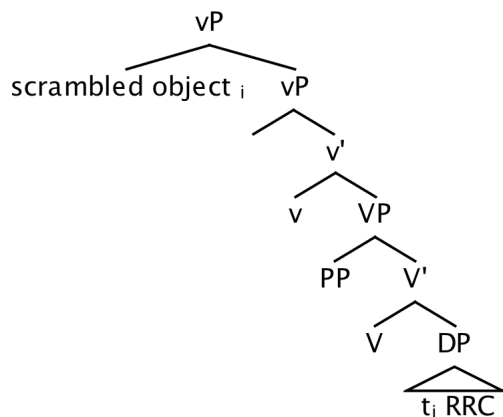
Then, *um livro*, 'a book', can be adjacent to the RRC (as in (169)) or may undergo short scrambling, assigning discourse prominence to the RRC that is stranded in the rightmost sentential position. In the later case, *um livro* 'a book' moves leftward past the position of the intervening PP and adjoins to VP, stranding the RRC *in situ*. This derives the pattern of RRC extraposition displayed in (168) above.

The idea that there are two verb phrases in a clause (the so-called double VP-shell approach proposed by Larson 1988, 1990) provides two possible landing sites for the scrambled object: left-hand adjunction to the higher vP or to the lower VP, as sketched in (170)a and (170)b, respectively:

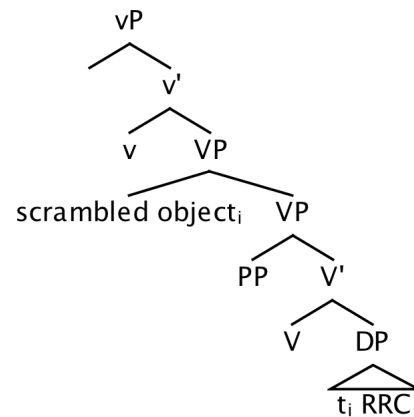
²⁸ Note that the present approach is also compatible with an analysis that postulates the DP-PP base-order. In this case, the PP-DP order would be derived from scrambling of the PP to the left of the DP, followed by scrambling of the antecedent to the left of the PP.

²⁹ Also note that, as mentioned above, there are different constraints that may lead to the PP-DP base-generation order. In (168), it results from heaviness, whereas in (166), it results from binding requirements.

(170) a.



b.



One possible way to identify the exact landing site of the scrambled object could be to examine its relative position with respect to a post-verbal subject in [Spec, vP]. However, as I discuss in Section 5.3.4.2 below, in sentences involving RRC extraposition, the subject and the object cannot independently co-occur in a post-verbal position. Therefore, this test must be discarded for the present purposes.

Another possibility is to assume that adverb placement can be used to identify the exact position of a constituent within a double VP-shell. Unfortunately, because Costa (1998, 2004a, 2004b) assumes a single VP-shell in his studies of adverbs in CEP, adverbs cannot be used as a reliable test for this specific purpose, at least until more research is developed in this domain.

Finally, let us consider the validity of another test: the so-called *Fronting/Preposing* (see, e.g., Costa 2004a: 49, 147) or *VP-topicalization* (see, e.g., Kato and Raposo 2007, Bastos 2001). This construction involves two instances of the same verb in a single sentence: an infinitival form in the preposed constituent and a finite form in the normal position of the verb in CEP (see (171), taken from Kato and Raposo 2007: 211):³⁰

³⁰ It is worth noting that there is no consensus in the literature as to the analysis of the construction in (172). Matos (1992: 195-196) claims that the preposed constituent is merged *in situ*, whereas Kato and Raposo (2007) suggest that it undergoes movement to the left periphery. In contrast, Matos (1992: 195-196) claims that the preposed constituent is a clausal constituent adjoined to the matrix clause, whereas Kato and Raposo (2007) claim that it is a topicalized VP. For a non-uniform approach to the phenomenon of VP-topicalization, see also Bastos (2001).

- (171) Visitar os amigos, a Maria visita todos os anos.
visit:INF the friends the M. visits every the years
 ‘Visit her friends, Maria does it every year.’

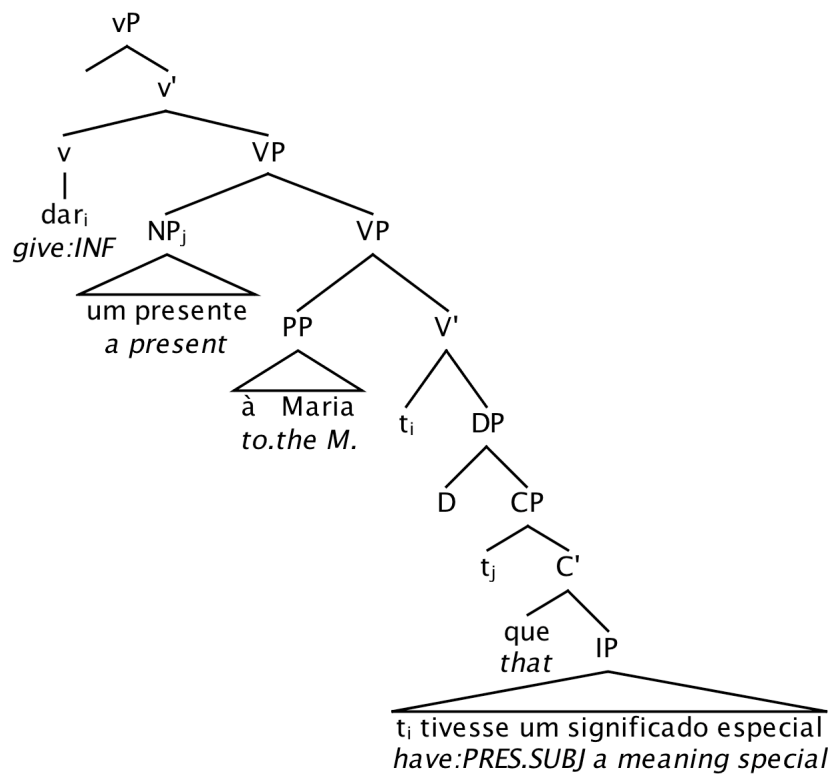
An extraposed RRC involving a double complement construction can surface in the preposed constituent, as illustrated in (172):

- (172) Eu queria dar um presente à Maria que tivesse um
I wanted give:INF a present to.the M. that have:PRES.SUBJ a
 significado especial e [dar um presente à Maria que
meaning special and give:INF a present to.the M. that
 tivesse um significado especial] eu dei.
have: PRES.SUBJ a meaning special I gave
 ‘I wanted to give a present that had a special meaning to Maria and give a present that had a special meaning to Maria I did it.’

Let us assume, along the lines of Kato and Raposo (2007), that this construction: (i) involves VP-topicalization and that (ii) the topicalized constituent contains a copy of the V (which moves to I) that is spelled out in its default infinitive form.³¹ Under these assumptions, the order of constituents within the topicalized constituent in (172) (i.e., the *verb - scrambled object* order) can only be derived if we assume that: (i) the verb is spelled in the light *v*; and (ii) the antecedent of the relative clause is left-adjoined to the lower VP, as sketched in (173). As can be easily concluded, if the antecedent of the RRC were adjoined to the higher *vP*, it would precede the verb *dar* ‘give’.

³¹ Kato and Raposo (2007) assume that the verb form that appears in the numeration is the infinitive, which after the addition of the inflection loses the final *r*. Therefore, when the verb is spelled out inside the VP, it surfaces in the default infinitive form, as no inflection was added to it at this point of the derivation. In contrast, when the verb is spelled out in I, it surfaces in a finite form because the addition of the inflection has already taken place.

(173)



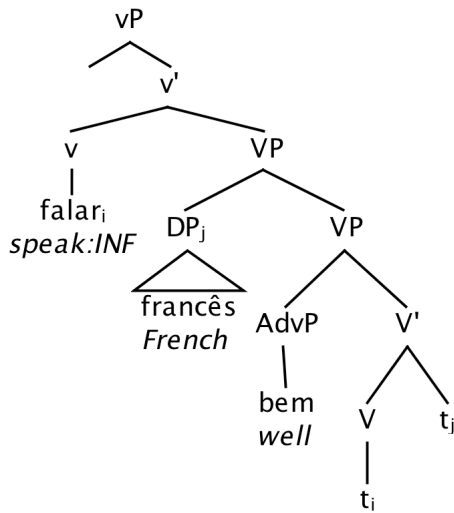
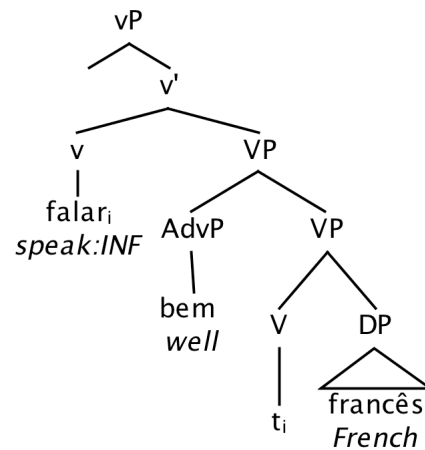
Note that the same line of reasoning applies to the instances of VP-topicalization that involve the monosyllabic adverb *bem* ‘well’ and a scrambled object, as in (174).

(174) a. Falar francês bem, o João fala. [scrambled object]
speak:INF French well the J. speaks

‘Speak French well, João does it.’

b. Falar bem francês, o João fala. [non-scrambled object]
speak:INF well French the J. speaks

The *verb - scrambled object - adverb* order in (174)a and the *verb-adverb* order in (174)b emerge from a configuration in which the verb is spelled out in the light *v*, and the scrambled object/monosyllabic adverb *bem* ‘well’ are left-adjoined to the lower VP, as sketched in (175)-(176). Note that if the adverb (and the scrambled object) were left-adjoined to vP, they would precede the verb *falar* ‘speak’.

(175) *verb-DO-adverb* order(176) *verb-adverb-DO* order

In summary, VP-topicalization suggests that scrambled constituents in CEP may be left-adjoined to the lower VP. However, note that other contexts independently reveal scrambled objects to be left-adjoined to the higher vP. This is the case of the sentences displaying the VOS order, as in (177). Under the assumption that the post-verbal subjects are in [Spec, vP] (see Costa 1998, 2004a), it is clear that the scrambled object in (177) is left-adjoined to the higher vP.

- (177) *Comeu a sopa o Paulo.*
ate the soup the P.
 ‘Paulo ate the soup.’

The same is true of the cases of RRC extraposition involving the subject of an unergative verb as an antecedent, as in (178) below, repeated from (40). In this case, the adverb *ontem* ‘yesterday’ is left-adjoined to the higher vP, and the subject in [Spec, vP] undergoes scrambling to a vP-adjoined position.

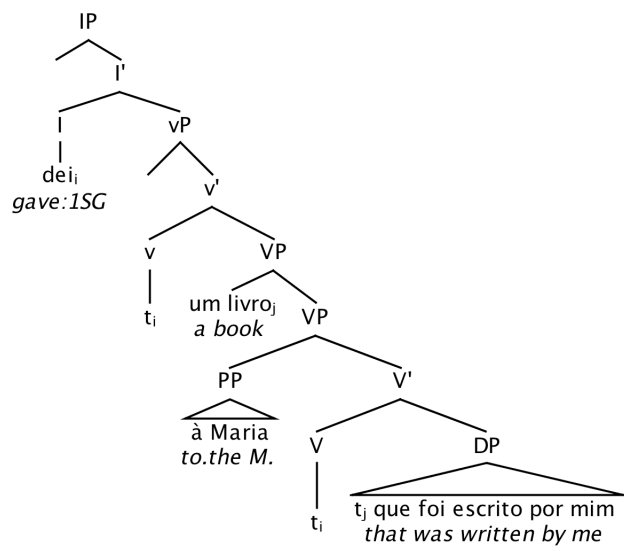
- (178) *Telefonou um rapaz ontem que queria informações sobre a tua casa.*
phoned a boy yesterday that wanted details about the your house
 ‘A boy phoned yesterday who wanted details about your house.’

Cumulatively, I conclude that scrambled constituents in CEP may be adjoined to the higher vP, e.g., in VOS contexts and in the contexts involving scrambling of the

subject of unergative verbs. However, they can also be left-adjoined to the lower VP, as is the case for scrambled objects in double object constructions.

Therefore, I propose that the (simplified) structure of an extraposed RRC with a PP complement as intervening material is as in (179):

(179)



5.3.4.2 Blocking the occurrence of other constituents in the intervening position

This section is devoted to demonstrating why the subject and the direct object cannot surface in the intervening position in CEP. Given that facts regarding the word order have an important bearing on the syntax of RRC extraposition, I first make a few remarks about the word order in CEP (in particular, with respect to subject inversion) and then demonstrate how the syntactic constraints that independently hold for CEP can explain the impossibility of the subject and the direct object surfacing in the intervening position.

In CEP, sentences with broad information focus exhibit a restriction on subject inversion that is related to the type of verb involved (see Martins, in prep.). Although subject inversion is possible with unaccusative, unergative and indirect transitive

verbs (see (180)b-d), it is impossible with direct transitive and ditransitive verbs (see (180)e-f).³²

- (180) [A]: a. O que aconteceu?
 the what happened
 ‘What happened?’
- [B]: b. Chegou uma carta anónima. [unaccusative verb]
 arrived a letter anonymous
 ‘An anonymous letter arrived.’
- c. Telefonou a Maria. [unergative verb]
 phoned the M.
 ‘Maria phoned.’
- d. Apareceram dois policiais em nossa casa. [indirect transitive verbs]
 showed.up two cops at our home
 ‘Two cops showed up at our home.’
- e. *Comprou o João uma casa. [direct transitive verb]
 bought the J. a house
 ‘João bought a house.’

³² A similar pattern has been observed in other constructions cross-linguistically (see Alexiadou and Anagnostopoulou 2001, for an overview). For instance, such a pattern is found in expletive constructions in French and English, which are well formed with intransitive verbs but not with direct transitive verbs (see (i) and (ii), respectively) and in stylistic inversion in French, which is also well formed with intransitive verbs but not with transitives (see (iii)).

- (i) a. Il est arrivé un homme.
 EXPL is arrived a man

‘There has arrived a man.’

- b. *Il a lu un élève le livre.
 EXPL has read a student:NOM the book:ACC

‘There has read a student the book.’

- (ii) a. There arrived a man.
 b. *There finished somebody the assignment.

(*ibidem* 2001: 195)

- (iii) a. Je me demande quand partira Marie.
 I wonder when leave:FUT M.

‘I wonder when Marie will leave.’

- b. *Je me demande quand achèteront les consommateurs les pommes.
 I wonder when buy:FUT the consumers:NOM the apples:ACC

‘I wonder when the consumers will buy the apples.’

(*ibidem* 2001: 195-196)

f. *Ofereceu o João um anel de noivado à Ana. [ditr. verb]
offered the J. a ring of engagement to.the A.

‘João offered an engagement ring to Ana.’

Under Costa’s (2004a) analysis of CEP, post-verbal subjects in simple declarative affirmative sentences arise in the following way: the verb moves up to I° and stops there. The subject does not precede it because it has never moved from its base-position.

On the basis of Costa’s analysis, the contrasts found in the paradigm (180) can be captured by the generalization in (181):³³

(181) *Restriction on subject inversion in CEP (I)*

The subject and the direct object cannot stay in a VP-internal position; one of them must vacate the VP.

However, the restriction in (181) does not hold for all syntactic and discourse contexts. Abstracting away from the sentences that involve V-to-C movement,³⁴ the subject and the direct object may co-occur inside the VP in sentences displaying narrow information focus. Two possible word orders may be found: (i) the VOS word order is found when the subject is assigned narrow focus (as in (182)); and (ii) the

³³ An explanation for this restriction is beyond the scope of this dissertation. Nevertheless, it is worth mentioning that an analysis such as Alexiadou and Anagnostopoulou’s (2001) may explain the restrictions under scrutiny. According to these authors, there is a general ban against having the subject and the direct object in a VP-internal position. This is explained by postulating that a head cannot have more than one unchecked Case feature in LF. Without going into the details of their analysis, the derivation of the ungrammatical V [_{VP} S O] generically proceeds as follows: (i) V raises overtly to T; (ii) after spell out, v raises to T forming a complex head (T^{max}); and (iii) T^{max} inherits the Case features of T (traditionally the nominative Case) and the Case features of v (traditionally the accusative Case), and as a consequence, the derivation crashes.

³⁴ According to Martins (in prep.), there are some factors that may contribute to making the VSO order available in broad information focus sentences, e.g., paratactic factual concessive constructions, which express the speaker’s disapproval of (or disappointment with) the unpredictability of an event or situation (see (i)).

(i) a. *Convidei eu a Maria para jantar e ela não apareceu.*
invited I the M. for dinner and she not appeared

‘I invited Mary for dinner but she didn’t come./Although I invited Mary for dinner, she didn’t come.’

(Martins, in prep.)

However, as A. Martins notes, these constructions seem to involve V-to-C movement.

VSO word order arises when both the subject and the direct object are assigned narrow information focus (as in (183)).

(182) *Subject is focused:*³⁵

[A]: a. Quem é que partiu a janela?
 who is that broke the window

‘Who broke the window?’

[B]: b. #Partiu o Paulo a janela.
 broke the P. the window

‘Paulo broke the window.’

c. Partiu (a janela) o Paulo.
 broke the window the P.

(Costa 2004a: 80)

(183) *Subject and direct object focused:*

[A]: a. Ninguém partiu nada.
 nobody broke nothing

‘Nobody broke anything’

[B]: b. Partiu o Paulo a janela.
 broke the P. the window

c. #Partiu a janela o Paulo.
 broke the window the P.

(Costa 2004a: 80)

Assuming (along with Costa 2004a) that in narrow information focus sentences, the post-verbal subject stays in its base-position,³⁶ the restriction in (181) can be reformulated as in (184):

(184) *Restriction on subject inversion in CEP (II)*

In sentences with broad information focus, the subject and the direct object cannot remain in a VP-internal position; one of them must vacate the VP.

³⁵ According to my judgment, sentence (182)c cannot occur with the direct object to the right of the verb, and only the subject is possible as an answer to (182)a. Nevertheless, assuming that other speakers may share Costa’s (2004a) judgments, I will pursue the argument as though the VOS order in CEP were possible in the context given in (182), leaving the investigation of this issue open for future research.

³⁶ According to Costa (2004a), the VOS order (in (182)c) is derived by short scrambling the object past the subject, whereas in the VSO order (in (183)b), the subject and the object remain in their base-position inside the VP.

With this in mind, let us return now to the syntax of RRC extraposition. The facts about CEP word order in inversion contexts predict the availability of RRC extraposition in sentences displaying narrow information focus, with the subject and the direct object as intervening material, as in (185):

- (185) a. [V DO_i S RRC_i] [subject in the intervening position]
 b. [V S_i DO RRC_i] [direct object in the intervening position]

In A. and B. below, I will show why this prediction is not borne out.

A. Subject in the intervening position

The analysis developed thus far predicts the occurrence of the subject in the intervening position in sentences displaying narrow information focus, when an extraposed RRC takes a direct object as an antecedent, as sketched in (186):

- (186) [_{IP} V [_{VP} DO [_{VP} S t_v [_{VP} t_v t_{DO} RRC]]]]

However, sentences involving the structure in (186) are ungrammatical in CEP, as illustrated in (187) below. In this case, the object *um bolo* ‘a cake’ should be allowed to undergo short scrambling (past the subject), stranding the RRC in the complement position of V, contrary to fact.

- (187) *Trouxe um bolo a Rita que tinha compota de morango.
brought a cake the R. that had jam of strawberry
 ‘Rita brought a cake that had strawberry jam.’

Therefore, an explanation is required for the contrast in (188):

- (188) a. [_{IP} V [_{VP} DO [_{VP} S t_v t_{DO}]]]
 b. * [_{IP} V [_{VP} DO [_{VP} S t_v t_{DO} RRC]]]

Recall from the discussion above (around (184)) that the VOS order in CEP arises in narrow information focus sentences, where only the subject is focused. The object is not interpreted as information focus because it is previously referred to in the discourse.

Therefore, in VOS sentences displaying RRC extraposition (as in (186)), the antecedent of the extraposed RRC (i.e., the direct object) is not also interpreted as an information focus. Such a configuration is therefore excluded by the generalization in (143), repeated here as (189), which dictates that the antecedent of the extraposed RRCs must be contained in the focus domain.

(189) *Relation between RRC extraposition from post-verbal positions and information structure (CEP)*

The extraposed RRC and the antecedent must be contained in the same focus domain.

B. *Direct object in the intervening position*

In CEP, sentences with narrow information focus can display VSO order (see, e.g., (183) above). However, an extraposed RRC taking a subject as an antecedent cannot surface with a direct object in the intervening position, according to the scheme in (190):

(190) *[V S DO _{t_S} RRC]

The impossibility of (190) is straightforwardly derived from the stranding analysis of RRC extraposition proposed here. Under a single VP-shell, the subject is base-generated in [Spec, VP] and the direct object in the complement position of V. Then, an RRC stranded in the subject position can never follow a direct object in the complement position of V.

5.3.5. Summary

Section 5.3 demonstrates that in CEP, RRC extraposition from post-verbal positions can be accounted for in terms of short scrambling. The main findings of this section can be summarized as follows:

- In CEP, subjects in [Spec, VP] (just like objects and subjects of unaccusative verbs) can undergo short scrambling (represented here as left-hand adjunction to VP).

- RRC extraposition from post-verbal positions involves short scrambling of the antecedent and stranding of the RRC in its base-position.
- Short scrambling does not necessarily remove unfocused material from the focus domain. It may take place within a focus domain to assign discourse prominence to the rightmost sentential constituent.
- In the context of RRC extraposition from post-verbal antecedents, the scrambling of the antecedent takes place within a focus domain. It has the effect of assigning discourse prominence to the RRC that is stranded in its base-position.
- The only constituents that may intervene between a post-verbal antecedent and its extraposed RRC are PPs and adverbs; these elements can be modifiers or arguments of the verb.
- The occurrence of these intervening elements can be accounted for by assuming that VP-modifiers (adverbs/PPs) are left-adjoined to the VP, whereas PP-complements are base-generated in the Specifier position of the lower VP. In both cases, scrambling moves the antecedent across the modifier/complement, stranding the RRC *in situ*.
- The impossibility of having a subject in the intervening position can be accounted for by a restriction on information structure that ensures that the extraposed RRC and the antecedent must be contained in the same focus domain.
- The impossibility of having a direct object in the intervening is excluded by the stranding analysis of RRC extraposition because an RRC stranded in the subject position can never follow a direct object in the complement position of V.

5.4. Deriving the properties of RRC extraposition in CEP

Having taken this excursus into the derivation of RRC extraposition (via movement to the left periphery or short scrambling), let us now return to the contrasting properties of RRC extraposition outlined in Section 4. There are three main properties of RRC extraposition that require an explanation:

A. *The definiteness effect*

B. *Restriction on extraposition from embedded positions*

C. *Restriction on extraposition from pre-verbal positions*

In the following, I show how the theoretical apparatus presented in the previous sections can account for these properties, which are discussed in turn.

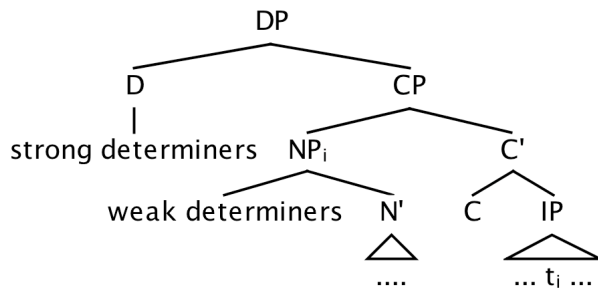
A. *The definiteness effect*

As mentioned in Section 4.1, extraposed RRCs in CEP can take weak noun phrases as their antecedent but not strong noun phrases. This property can be explained if we assume, following Bowers (1988), that strong and weak noun phrases differ in their structure. Strong quantifiers are of category D, whereas weak quantifiers are adjectives and attach within NP, as illustrated in (191):

- (191) a. [DP each [NP picture of manatees]]
 b. [NP [AP many] [N' pictures of manatees]]

Extending Bower's proposal to the raising analysis of relative clauses, I assume (in line with Kayne (1994) and Lee 2007) that strong determiners are located in the external determiner, whereas weak determiners are within NP. This explains in a straightforward manner why extraposed RRCs can take only weak noun phrases as antecedents. Weak noun phrases can be moved leftward as a constituent, whereas strong noun phrases cannot because there is no constituent that includes the strong determiner and the noun phrase but excludes the RRC, as shown in (192):

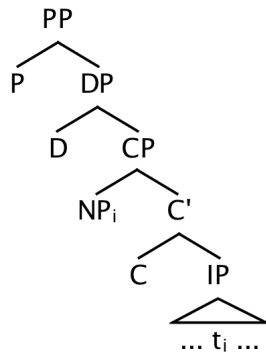
(192)



B. Restriction on extraposition from embedded positions

In CEP, RRC extraposition is not permitted if the antecedent is the object of a preposition. Again, this restriction is straightforwardly derived under the standard assumption that movement only applies to constituents. As sketched in (193), the preposition, the determiner and the noun phrase in [Spec, CP] do not form a constituent (excluding the RRC). As a result, they cannot undergo leftward movement, stranding the RRC *in situ*.

(193)



C. Restriction on extraposition from pre-verbal positions

In CEP, extraposed RRCs can take post-verbal subjects as antecedents but not pre-verbal subjects. Additionally, extraposed RRCs can take *wh*-constituents, emphatic/evaluative phrases and preposed foci as antecedents but not topics. Barbosa (2009) provides an explanation for this contrast in terms of a prosodic account of extraposition. Below, I first discuss Barbosa's (2009) proposal, showing that it is incompatible with the analysis adopted here. Then, I offer an alternative explanation for the phenomenon that rests upon the semantic interpretation of the antecedent.

As mentioned in Chapter 1 (Section 4.3.2, fn. 15), there are currently two competing proposals for the syntactic analysis of pre-verbal subjects in CEP. Indeed, Costa (2001, 2004a), Costa and Duarte (2002) claim that pre-verbal subjects A-move to [Spec, IP], whereas Barbosa (1995, 2000, 2009) claims that subjects are base-generated in a left-dislocated position (as adjuncts to CP/IP). The two hypotheses are sketched in (194)a and (194)b, respectively:

- (194) a. [IP S V [VP t_S t_V]]
 b. [IP/CP S [IP/CP V [VP *pro* t_V]]

One of Barbosa's arguments in favor of the left-dislocated position of subjects in CEP (and in Romance NSLs in general) is precisely the impossibility of extraposition from pre-verbal indefinite subjects. Assuming Truckenbrodt's (1995) prosodic approach to extraposition, Barbosa claims that relative clause extraposition is sensitive to Intonational Phrase (IntP) boundaries. More precisely, for extraposition to be possible, no IntP boundary may intervene between the antecedent and the rest of the clause. Because dislocated elements are (at least initially) mapped onto an IntP domain that is separated from the IntP domain onto which the rest of the clause is mapped, the impossibility of relative clause extraposition in CEP is straightforwardly derived (see (195), taken from Barbosa 2009).

- (195) a. Syntax:

[um homem que quer falar contigo]_k [IP *pro*_k apareceu]
a man that wants talk:INF with.you showed.up

'A man showed up that wants to talk to you.'

- b. Prosodic Structure:

[um homem que quer falar contigo]IntP apareceu]IntP
a man that wants talk:INF with.you showed up

In contrast, because the pre-verbal subject is in [Spec, IP] in Romance non-NLS (and English), no IntP boundary intervenes between the pre-verbal subject and the rest of the clause. Consequently, extraposition is allowed.

As for the cases in which non-referential QPs and focalized DPs appear in a pre-verbal position in CEP, Barbosa claims that these constituents are not left-dislocated but rather fronted by A-bar movement. In this case, no IntP boundary intervenes between the fronted constituent and the rest of the clause, and extraposition is allowed.

As can be easily concluded, Barbosa's account of RRC extraposition is not compatible with the stranding analysis of RRC extraposition proposed here because the subject is base-generated in a left-dislocated position. To be compatible with the analysis presented here, this account must be 'massaged' to provide for the base-generation of the subject in a VP-internal position.

Additionally, note that Barbosa's analysis makes the wrong prediction with respect to the availability of RRC extraposition in NSLs. Barbosa claims that there is a correlation between the possibility of extraposition from pre-verbal subject positions and the Null Subject Parameter. Specifically, she claims that NSLs do not allow extraposition from pre-verbal subjects, whereas non-NSLs allow for it. Again, this is simply not correct. Over the course of its history, Portuguese has always been an NSL, but in earlier periods of its history, it allowed extraposition from pre-verbal subjects, as illustrated in (196) (repeated from (97) above).³⁷

(196) se Algẽ A eles veer que diga que lĩj eu
 if someone to them come:FUT.SUBJ that says that to.him:CL I
 Alguna cousa diuía
 some thing owed

‘[And] if someone who says that I owed him something comes towards them...’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1275)

Alternatively, I would like to suggest that the explanation for the restriction on extraposition from pre-verbal subjects rests upon the semantic interpretation of the antecedent. More precisely, I claim that RRC extraposition in CEP obeys the *Interpretative Principle* given in (197):

³⁷ Interestingly, Fiéis and Lobo (2010) have shown that earlier stages of Portuguese are also problematic for Barbosa's hypothesis concerning the position of the subject in absolute gerund clauses. Barbosa claims that NSLs and non-NSLs contrast with respect to the possibility of having pre-verbal subjects in absolute gerund clauses: non-NSLs are subject initial (*Your brother having called...*), whereas NSLs are V/Aux initial (*Aparecendo a Maria...* lit. 'showing up Maria...'). Fiéis and Lobo (2010) show that this correlation is simply not correct. In earlier stages of its history, Portuguese is an NSL and allows for pre-verbal subjects in absolute gerund clauses, as illustrated in (i), taken from Fiéis and Lobo (2010):

(i) Joham Rrodriguez estando no logar, veo sobr'elle o concelho de Ledesma
 J. R. be:GER in.the place came over.him the ±group of L.
 'Joham Rodriguez being in the place, the group of Ledesma attacked him.'

One hypothesis that is worth exploring in future research is that the Null Subject Parameter does not necessarily correlate with specific subject positions. We may tentatively hypothesize that NSLs might display different positions for pre-verbal subjects and that this may be subject to cross-linguistic and diachronic variation. Under these assumptions, we can conjecture that earlier stages of Portuguese and CEP may differ in the structural position occupied by pre-verbal subjects. However, further research is necessary in this domain to warrant the validity of these suggestions.

(197) *Interpretative Principle*

The antecedent of an extraposed RRC must occur in a position non-ambiguously interpreted as non-topic (in Kuroda's 2005 sense).³⁸

The fact that the restriction on extraposition from pre-verbal positions is semantically motivated should not come as a surprise because several authors have already observed that word order in CEP reflects both information structure and the contrast between categorical andthetic judgments (in the sense of Kuroda 1965, 1972, 2005).³⁹

Let us now explore in detail how the Interpretative Principle in (197) explains the restriction on extraposition from pre-verbal subjects and topics.

C1. *Restriction on extraposition from pre-verbal subjects*

Assuming the distinction between categorical andthetic judgments originally proposed by Kuroda (1965), Martins (in prep.) provides evidence for the idea that [Spec, IP] is an ambiguous position in CEP. It can be filled by topic elements (i.e., the subject of predication in sentences expressing categorical judgments), but it can also be filled by non-topic elements (i.e., the subject of a sentence expressingthetic/descriptive judgments). For details and examples, see Chapter 1 (Section 4.3.2).

This explains why extraposed RRCs cannot take a pre-verbal subject as an antecedent. According to the Interpretative Principle in (197), the antecedent of an extraposed RRC must occur in a position non-ambiguously interpreted as non-topic. Given that [Spec, IP] does not satisfy this requirement, a constituent occurring in this position cannot be the antecedent of an extraposed RRC.

In contrast, as already shown in Chapter 1 (Section 4.3.2), post-verbal subjects occupy positions non-ambiguously interpreted as non-topic. Hence, a scrambled

³⁸ In this context, the term topic is not used as a syntactic concept (i.e., as referring to a constituent that is placed at the sentential left periphery) nor as a discourse-theoretical concept (i.e., as referring to a constituent that expresses old information in the organization of the discourse) but as a semantic concept. In this sense, it is understood as a constituent that expresses an *aboutness relation* (see Kuroda 2005). For more details, I refer the reader to Chapter 1 (Section 4.3.2).

³⁹ Based on the Brentano-Marty theory of judgments, Kuroda distinguishes two types of judgments: categorical/predicational vs.thetic/descriptive. A predicational/categorical judgment is a cognitive act of attributing a predicate to a subject, whereas athetic/descriptive judgment is grounded, in its basic form, on perception. For further details, see Chapter 1 (Section 4.3.2).

subject left-adjoined to VP satisfies the Interpretative Principle in (197) and, therefore, can be taken as the antecedent of an extraposed RRC.

C2. Restriction on extraposition from other pre-verbal constituents

The Interpretative Principle in (197) can also explain why RRC extraposition cannot take place from topics. Assuming a split-CP approach (see Rizzi 1997), according to which there are different functional projections especially dedicated to single discourse functions (e.g., Topic and Focus), the position occupied by a topic constituent is non-ambiguously interpreted as topic. Therefore, RRC extraposition is ruled out by the Interpretative Principle in (197).

Conversely, the position occupied by *wh*-constituents, emphatic/evaluative phrases and preposed foci is non-ambiguously interpreted as non-topic. Therefore, the possibility for extraposition from these constituents is straightforwardly derived.

5.4.1. Summary

The results of Section 5.4 are summarized in Table 6. A plus indicates that the stranding analysis can straightforwardly derive the restriction or derive it with reference to independent principles; a minus would indicate that it cannot.

Table 6. Evaluation of the stranding analysis of RRC extraposition in CEP

		Derived by the stranding analysis
<i>A. No extraposition from strong noun phrases</i>		+
<i>B. No extraposition from embedded positions</i>		+
<i>C. Pre-verbal constituents</i>	No extraposition from pre-verbal subjects	+
	Extraposition from <i>wh</i> -constituents	+
	Extraposition from emphatic/evaluative phrases	+
	Extraposition from preposed foci	+
	No extraposition from topics	+

I conclude that the stranding analysis accounts for the restrictions on RRC extraposition identified in Section 4.1. However, it is worth noting that this analysis has received much criticism in the literature (see, among others, Buring and Hartmann

1997, Koster 2000, De Vries 2002). In the next section, I review some of the arguments that have been put forth in the literature against the stranding analysis and show that they do not offer any insurmountable obstacle to the approach proposed here because they do not apply to CEP.

5.5. Problems with the stranding analysis

Some of the arguments that have been adduced in the literature against the stranding analysis are listed in A.-I. below and will be discussed in turn:

- A. *Extraposition from strong noun phrases*
- B. *Extraposition from PPs*
- C. *Ungrammaticality of the source structure*
- D. *Constraints on the surface position of extraposed RRCs*
- E. *Extraposition from subjects*
- F. *Extraposition from any constituent*
- G. *Mirror effects*
- H. *VP-topicalization*
- I. *Extraposition from split antecedents*

A. *Extraposition from strong noun phrases*

One of the most frequently adduced arguments against the stranding analysis is that it cannot derive extraposed RRCs taking a *definite article+head* as an antecedent. For instance, Koster (2000) shows that an extraposed RRC can take a *definite article+head* as an antecedent in Dutch, as illustrated in (198) (see also (66) above):

- (198) Hij heeft [de vrouw]_i gezien t_i die het boek geschreven heeft.
 he has the woman seen who the book written has

‘He has seen the woman who has written the book.’

(Koster 2000: 5)

Koster argues that sentences like (198) undermine the stranding analysis: *de* and *vrouw* do not form a constituent to the exclusion of the RRC and, as a result, cannot undergo leftward movement, stranding the RRC *in situ*. However, note that this does not constitute a problem for the analysis adopted here. CEP, unlike Dutch, does not

allow for extraposed RRCs with a *definite article+head* as an antecedent, as shown in Section 4.1. In actual fact, the unavailability of sentences like (198) in CEP supports an analysis of RRC extraposition in CEP in terms of stranding.

B. *Extraposition from PPs*

A similar obstacle for the stranding analysis regards extraposition from PPs. Koster (2000) points out that extraposition from NPs within PPs is entirely grammatical in Dutch. See (199) (see also (69) and (71) above):

- (199) Hij heeft met een vrouw gesproken die alles wist
he has with a woman talked who everything knew
 ‘He has talked with a woman who knew everything.’

Koster (2000) and De Vries (2002) claim that the stranding analysis cannot derive (199), i.e., *met een vrouw* ‘with a woman’ is not a constituent and consequently cannot be moved leftwards. Alternatively, assuming that the PP and the head can be generated separately, the movement of *een vrouw* ‘a woman’ to a position inside the PP would involve movement to a non c-commanding position. Again, this problem does not arise in CEP because RRC extraposition cannot take place from embedded positions, as shown in Section 4.1.

C. *Ungrammaticality of the source structure*

Another obstacle to the stranding analysis concerns the ungrammaticality of the source structure in languages like Dutch. Recall that under the stranding analysis, the antecedent and the RRC are base-generated together. According to Koster (2000), such an analysis does not even have initial plausibility because the presumed source structure in (200), displaying the SVO order, is ungrammatical in Dutch.

- (200) *Hij heeft gezien de vrouw die het boek geschreven heeft
he has seen the woman who the book written has
 ‘He has seen the woman who has written the book.’

(Koster 2000: 7)

Sentence (201) is ungrammatical because in an SOV language like Dutch, NP objects do not generally follow the verb. Note that in the extraposed version of (200), the forbidden sequence *V-NP would still be involved, as illustrated in (202):

- (201) Hij heeft [NP de vrouw]_i gezien [NP [NP t_i] [CP die het boek geschreven heeft]]
he has the woman seen who the book written has
 (Koster 2000: 7)

Of course, this problem does not arise in an SVO language like CEP. In this language, NP objects usually follow the verb. Therefore, sentences where the head plus its RRC are construed post-verbally are entirely grammatical, as illustrated in (202)a. Given that the sequence V-NP is not forbidden, an RRC is likely to be stranded in a post-verbal position, as illustrated in (202)b:

- (202) a. Encontrei ontem uma rapariga que perguntou por ti.
met yesterday a girl that asked for you
 ‘Yesterday I met a girl that asked for you.’
 b. Encontrei uma rapariga_i ontem t_i que perguntou por ti.
met a girl yesterday that asked for you

D. Constraints on the surface position of extraposed RRCs

Another problem concerns the clause-final position of extraposed RRCs. It has been argued in the literature that if the RRC extraposition is derived from stranding, it becomes a mystery as to why an extraposed RRC cannot surface in an intermediate position. Koster (2000) and De Vries (2002) demonstrate that if the antecedent is preposed in Dutch, the relative clause cannot be left behind at the normal object position; see (203), adapted from De Vries (2002: 254):

- (203) *Een man heb ik die een rode koffer draagt gesignaleerd.
a man have I who a red suitcase carries noticed
 ‘I have noticed a man who carries a red suitcase.’

De Vries (2002) argues that if extraposition were stranding, nothing should exclude the schematic derivation of (203) given in (204):

- (204) a. V [NP RC] →
 b. [NP RC]_i V t_i →
 c. NP Aux S [tNP RC]_i V t_i

Although it can be a problem for Dutch, such an objection is not applicable to CEP. In CEP, if the antecedent is focalized, the extraposed RRC can be left behind at the normal object position, as illustrated in (205):

- (205) Nada mais disse que valesse a pena até ao final
nothing more said:3SG that be.worthwhile:IMPERF.SUBJ until to.the end
 da conferência.
of.the conference
 ‘He did not say anything else that was worthwhile until the end of the conference.’

The same is true of contexts involving a *wh*-constituent as the antecedent of an RRC. As shown in (206), an extraposed RRC can be left behind at a non-final position:

- (206) Quantas pessoas apareceram que não foram convidadas naquela
how.may people showed.up that not were invited in.that
 festa horrível que organizámos em minha casa!
party horrible that organized:1PL at my house
 ‘How many people showed up that were not invited in that horrible party that we organized at my house!’

In sentences (205)-(206), there is a clear pause between the extraposed RRC and the constituent in the clause-final position. This pause, which appears to be crucial to the acceptability of these sentences, seems to suggest that an extraposed RRC may surface in a non-final position if the constituent following the RRC is mapped into an independent intonational phrase. This guarantees that the extraposed RRC receives prosodic stress and is interpreted with prosodic and discourse prominence.

Although various aspects of the interaction between extraposition and prosody remain open for future research, it is clear from the discussion above that the (prosodic) constraints on the surface position of extraposed RRC do not undermine the syntactic analysis of RRC extraposition in CEP in terms of stranding.

E. *Extraposition from subjects*

Another problematic aspect discussed by Koster (2000) is the possibility of having extraposition from subjects in Dutch, as illustrated in (207):

- (207) Een vrouw heeft het boek geschreven die alles wist
 a woman has the book written who everything knew
 ‘A woman who knew everything has written the book.’

(Koster 2000: 8)

It is generally assumed that Dutch verbs are spelled out in V (apart from verb second of the finite verb in main clauses). Therefore, under the stranding analysis, the source structure of (207) would have the head plus its RRC to the right of the VP as a subject, which is not a legitimate base-position for the subject in any language.

However, this problem does not arise in CEP. Suppose that subjects are base-generated VP-internally, as suggested by Koopman and Sportiche (1991). Furthermore, consider that CEP displays V-to-I movement. Under these two assumptions, it is clear that an extraposed RRC taking a subject as an antecedent can be stranded in its base-position. As depicted in (208), the RRC can be stranded in [Spec, VP] (under a single VP-shell), preceding the trace of V (which moved to I):

- (208) [IP V [VP S [VP adverb [VP t_S RRC [v' t_V]]]]]

F. *Extraposition from any constituent*

De Vries (2002) shows that RRC extraposition in Dutch can take place from any constituent. He claims that the stranding analysis is implausible because it requires that all the material must vacate the VP or an even higher projection. Focusing on extraposition from non-objects, he claims that under this approach, everything would be generated within the VP, and the VP would always be emptied, e.g., as in (209).⁴⁰

- (209) S Aux ... DO V AdvP [VP [t_S RRC] t_V t_{DO}]

⁴⁰ Additionally note that the derivation in (209) would still be a problem because it leads to the wrong word order.

However, note that this does not constitute a problem for RRC extraposition in CEP. In this language, there is independent V-to-I movement. Therefore, the main verb always vacates the VP. Furthermore, note that pre-verbal subjects cannot take an extraposed RRC as an antecedent (see Section 4.1). Therefore, they do need to vacate the VP.

G. *Mirror effects*

De Vries (2002) shows that if two RRCs are extraposed in Dutch, a mirror effect emerges: an RRC extraposed from the object must precede an RRC extraposed from the subject. This is illustrated in (210), taken from De Vries (2002: 248):

- (210) a. Een zekere misdadiger heeft de kluis gekraakt die tweehonderd
a certain criminal has the safe cracked that two.hundred
 diamanten bevatte, die ook meneer X heeft vermoord.
diamonds contained, who also mister X has killed
 lit. 'A certain criminal has cracked the safe that contained two diamonds, who also has killed mister X.'
- b. *Een zekere misdadiger heeft de kluis gekraakt die ook meneer X
a certain criminal has the safe cracked that also mister X
 heeft vermoord, die tweehonderd diamanten bevatte.
has killed that two.hundred diamonds contained

He argues that this is a problem for the raising analysis, as we would expect crossing dependencies (as represented in (211)), which are contrary to fact.

- (211) S ... DO ... [_{ts} RRC_s] ... [_{t_{DO}} RRC_{DO}]

Unfortunately, this test does not yield conclusive results in CEP. A sentence that could virtually instantiate the mirror effect under discussion could be one involving an RRC extraposed from a *wh*-constituent and an RRC extraposed from the object. However, such sentences are excluded, independently of the relative order of the two extraposed RRCs, as shown in (212):

- (212) a. *[Quantas pessoas]_i requisitaram [livros]_j ontem t_j que são
how.many people checked.out books yesterday that are
 recomendados pelo Ministério t_i que são sócias da biblioteca?
recomended by.the ministry that are members of.the library
 ‘How many people checked out books yesterday that are recomendaded by the
 ministry that are library members?’
- b. *[Quantas pessoas]_i requisitaram [livros]_j ontem t_i que são sócias
how.many people checked.out books yesterday that are members
 biblioteca t_j que são recomendados pelo Ministério?
library that are recomendaded by.the ministry

Another possible candidate for such a test would be a sentence involving an RRC extraposed from a preposed focus and an RRC extraposed from the object. Again, the two possible orders excluded:

- (213) a. *[Poucas pessoas]_i conseguiram comprar [casas]_j no leilão de ontem
few people managed buy:INF houses at.the auction of yesterday
 t_j que fossem baratas t_i que fossem satisfeitas.
that be:IMPERF.SUBJ.3PL cheap that be:IMPERF.SUBJ.3PL satisfied
 lit. ‘Few people managed to buy houses at the auction yesterday that be cheap
 that be happy.’
- b. *[Poucas pessoas]_i conseguiram comprar [casas]_j no leilão de ontem t_i
few people managed buy:INF houses at.the auction of yesterday
 que fossem satisfeitas t_j que fossem baratas.
that be:IMPERF.SUBJ.3PL satisfied that be:IMPERF.SUBJ.3PL cheap

Importantly, sentences (212)-(213) become grammatical if we eliminate one of the extraposed RRCs. The explanation behind the ungrammaticality of (212)b-(213)b is unclear to me. Nevertheless, I tentatively propose that two independent factors may contribute to the ungrammaticality of these sentences. First, it may be due to the effect of processing factors; as De Vries (2002: 248) notes, sentences with two extraposed RRCs are extremely hard to comprehend. Second, it may be explained by the same restriction that prevents the occurrence of the subject and the object in a post-verbal position. As stated in Section 5.3.4.2, this idea is not new. In CEP, there is a restriction against having a subject and an object in a post-verbal position in sentences displaying broad information focus. Note that in (212)-(213), after the extraction of the antecedents, the subject and the object positions are still filled with an RRC. Therefore, it is likely that the restriction against the occurrence of a subject and an object in a post-verbal position becomes active in these contexts as well.

H. *VP-topicalization*

Koster (2000) and De Vries (2002) claim that if an extraposed RRC is stranded within the VP, the verb and the extraposed RRC should be topicalized together (VP-topicalization). However, this is simply not allowed in Dutch, as is illustrated in (214):

- (214) **[gezien die een rode jas draagt] heb ik de man.*
seen who a red cote wears have I the man
 ‘I have seen the man who wears a red coat.’

(De Vries 2002: 256)

Let us determine if this test can be applied to CEP, considering first the RRC extraposition from post-verbal antecedents. The test of VP-topicalization was already discussed in Section 5.3.4.1. There, it was shown that there is no consensus in the literature concerning the syntactic analysis of sentences like (215) (repeated from (171) for ease of exposition):

- (215) *Visitar os amigos, a Maria visita todos os anos.*
visit:INF the friends the M. visits every the years
 ‘Visit her friends, Maria does it every year.’

Let us assume here (along the lines of Kato and Raposo (2007)) that this construction involves VP-topicalization (but see fn. 30 for a brief discussion). Under the analysis proposed here, a post-verbal antecedent and its RRC are within the VP. Therefore, both elements are expected to surface in the topicalized constituent in a construction like (215). This prediction is borne out, as illustrated in (216):

- (216) *Encontrar uma pessoa na escola que esteja interessada em*
find:INF a person in.the school that be:PRES.SUBJ interested in
ir para Angola, não acredito que encontres.
go:INF to A. not believe:1SG that find:2SG
 ‘Find a person in the school that is interested in going to Angola, I do not believe you will do it.’

Consider now the sentences involving extraposition from a pre-verbal constituent. In this case, there are three options in CEP: the antecedent may be a preposed focus, an

emphatic/evaluative phrase or a *wh*-constituent (see Section 4.1). However, when these elements are extracted from the VP, VP-topicalization is simply not allowed, as shown in (217)-(219):

- (217) a. Nada de jeito ele viu na sua recente ida a Paris.
nothingworthwhile he saw in.the his recent visit to P.
 ‘He did not see anything worthwhile in his recent visit to Paris.’
 (Raposo 1995: 456)
- b.*Ver na sua recente ida a Paris, nada de jeito ele viu.
see:INF in.the his recent visit to P. nothing worthwhile he saw
- (218) a. Muito whisky bebi ontem à noite!
a.lot.of whisky drank:1SG yesterday at.the night
 ‘I drank a lot of whisky last night!’
- b.*Beber ontem à noite, muito whisky bebi!
drink:INF yesterday at.the night a.lot.of whisky drank:1SG
- (219) a. Quantas pessoas conheceste em Inglaterra?
how.many people met:2SG in England
 ‘How many people did you meet in England?’
- b.*Conhecer em Inglaterra, quantas pessoas conheceste?
meet:INF in England how.many people met:2SG

Consequently, it comes as no surprise that extraposed RRCs taking a preposed focus, an emphatic/evaluative phrase or a *wh*-constituent as an antecedent are not allowed in a topicalized VP; see (220)-(222):

- (220) Poucas pessoas conheço que vão ao ginásio.
few people know:1SG that go to.the gym
 ‘I know few people who go to the gym.’
- *Conhecer que vão ao ginásio, poucas pessoas conheço.
know:INF that go to.the gym few people know:1SG
- (221) a. Muito whisky bebi ontem que estava fora do prazo!
a.lot.of whisky drank:1SG yesterday that was out of.the expiry.date
 ‘I drank a lot of whisky yesterday that was expired!’
- b.*Beber ontem que estava fora do prazo, muito whisky
drink:INF yesterday that was out of.the expiry.date a.lot.of whisky
 bebi!
drank

- (222) a. Quantas pessoas conheces que vão ao ginásio?
how.many people know:2SG that go to.the gym
 ‘How many people do you know that go to the gym?’
- b. *Conhecer que vão ao ginásio, quantas pessoas conheces?
know:INF that go to.the gym how.many people know:2SG

It seems fair to conclude that the impossibility of having an extraposed RRC within a topicalized constituent in (220)-(222) does not undermine the stranding analysis. The ungrammaticality of examples (217)-(219) demonstrates that this can be independently explained by the fact that VP-topicalization is incompatible with the extraction of a preposed focus, an emphatic/evaluative phrase or a *wh*-constituent.

I. *Extrapolation from split antecedents*

De Vries (2002) claims that English and Dutch allow for split antecedents, as illustrated in (223). In this example, the relative pronoun triggers plural agreement on the verb in the relative clause (which shows that (223) is not simply a *Right Node Raising* construction).

- (223) Ik heb een vrouwi_i gezien en jij hebt een man_j bespied
I have a woman seen and you have a man spied.on
 die_{i+j} beide een rode jas droegen.
who both a red coat wore:PL
 ‘I saw a woman and you have spied on a man who wore a red coat.’
 (De Vries 2002: 264)

According to De Vries (2002), the stranding analysis cannot derive sentences like (223) because the head and its relative clause are always generated together. Hence, the plural relative pronoun and verb in (223) cannot be derived.

Fortunately, this problem does not even arise in CEP because RRCs with a split antecedent are completely excluded:

- (224) *Eu comprei um computador ontem e o meu marido ofereceu-me
I bought a computer yesterday and the my husband offered.me:CL
 uma impressora hoje que estavam em promoção no centro comercial.
a printer today that were at discount at.the center shopping
 ‘I bought a computer yesterday and my husband offered me a printer today that
 were at a discount at the shopping center.’

5.5.1. Summary

In this section, I examined nine problems that have been adduced in the literature against the stranding analysis. Because most of the problems were identified in the literature on Dutch (especially by Koster 2000 and De Vries 2002), it was possible to systematically compare the behavior of Dutch and CEP with respect to the same phenomena, and the results are summarized in Table 7. Here, the stranding theory is evaluated in the following way: a plus indicates that the stranding analysis can derive the property straightforwardly or with reference to independent principles, a minus indicates that it cannot, and an asterisk indicates that the property does not hold for a specific language.

Table 7. Global evaluation of the stranding analysis of RRC extraposition in CEP/Dutch

	Dutch	CEP
A. <i>Extraposition from strong noun phrases</i>	-	*
B. <i>Extraposition from PPs</i>	-	*
C. <i>Ungrammaticality of the source structure</i>	-	*
D. <i>Constraints on the surface position of extraposed RRCs</i>	-	+
E. <i>Extraposition from subjects</i>	-	+
F. <i>Extraposition from any constituent</i>	-	*
G. <i>Mirror effects</i>	-	*
H. <i>VP-topicalization</i>	-	+
I. <i>Extraposition from split antecedents</i>	-	*

I conclude that the stranding analysis can account for the properties of RRC extraposition in CEP but not in Dutch. The differences in RRC extraposition exhibited in both languages reinforce the conclusion that I drew on the basis of the empirical data discussed in Section 4.2, i.e., that RRC extraposition is not a uniform phenomenon, being subject to cross-linguistic variation.

5.6. Conclusion

This section focused on the syntactic nature of RRC extraposition in CEP. I showed that an analysis in terms of stranding can account for the properties of RRC extraposition in CEP. Specifically, I proposed that RRC extraposition in CEP involves A'-movement of the antecedent, either via short scrambling (when the antecedent is in a post-verbal position) or via movement to the left periphery (when the antecedent is in a pre-verbal position).

I attempted to keep the technical details of the analysis to a minimum. However, because the theory proposed here has an important impact on different domains of the clause structure (e.g., on the VP and the CP domain), its implementation required some technical discussion, especially with regard to short scrambling. The fact that word order in CEP is constrained by discourse/semantic/prosodic effects also added somewhat complex explanatory devices to the picture.

Nevertheless, it is worth noting that the analysis behind this theoretical apparatus actually amounts to a simple idea: RRC extraposition in CEP results from leftward movement of the antecedent and stranding of the RRC. Note further that the complex restrictions/principles that seem to interfere with this phenomenon have been independently proposed in the literature to account for other phenomena, e.g., for the different word-order patterns found in CEP.

Finally, it is worth noting that I confined myself to the presentation of the core ideas necessary for the implementation of the proposal. There are still some issues that deserve further, detailed inquiries, which I leave open for future research.

6. The analysis of RRC extraposition in earlier stages of Portuguese

In this section, I submit that RRC extraposition in earlier stages of Portuguese involves the same syntactic structure as coordination. Section 6.1 establishes the basic tenets of this analysis, introducing Koster's (2000) and De Vries' (2002) approaches to extraposition. Section 6.2 depicts how the *specifying coordination plus ellipsis* analysis proposed by De Vries (2002) can account for the properties of RRC extraposition outlined in Section 4.3. Finally, in Section 6.3, I address some problems of this analysis, providing solutions capable of overcoming some of its drawbacks.

6.1. The specifying coordination analysis

6.1.1. Koster (2000)

Koster (2000) proposes that in general, phrase structure takes two forms: primary phrase structure and parallel structure. Syntactically, both forms display the same configuration, consisting of a Specifier, a head and a complement. However, they are licensed in different ways. As Koster puts it:

Primary phrase structure has a functional part and a lexical part embedded in it. All lexical elements must be licensed in some functional position to their left, a consequence of universal head-initial structure (Kayne 1994). The elements of parallel structure are not directly licensed in this way, but at the most indirectly, by linking them to elements of the primary phrase structure.

(Koster 2000: 16)

Coordination has been seen as a form of parallel structure. However, Koster claims that parallel structure should be conceived as a broader phenomenon, encompassing coordination, extraposition, specifications found in equatives and possibly other phenomena such as appositions and right dislocations.

Assuming Munn's (1993) and Kayne's (1994) analysis of coordination (see also Johannessen 1998), Koster claims that parallel structure is syntactically represented as in (225). The primary phrase structure element is in the Specifier position, and the parallel conjunct is in the complement position of a Boolean head.

(225) [XP₁ [**Boolean head** XP₂]]

Although the parallel construal has a uniform syntactic configuration in (225), it encompasses structures with different semantics, depending on the nature of the Boolean head involved. In standard coordination, the Boolean head corresponds to coordinators such as *and* and *or*. In extraposition (and in equatives), the parallel construction involves an empty head (as in the asyndetic coordination in the traditional grammar). Koster (2000) represents this empty head as a colon head (':') and claims that it functions as an abstract Boolean operator, leading to the addition of properties, i.e., to the introduction of a specifying addition.

To support the idea that standard coordination and extraposition involve a similar syntactic representation, Koster (2000) demonstrates that they behave alike with respect to a number of properties.

First, in standard coordination two conjuncts may be non-adjacent in Dutch, as illustrated in (226)a:

- (226) a. *Zij heeft Marie gezien en mij*
she has M. seen and me
 ‘She saw Mary and me.’
- b. *Zij heeft Marie en mij gezien*
she has M. and me seen

(Koster 2000: 16)

Example (226)a cannot be derived from (226)b through rightward movement because it would constitute a violation of Ross’ Coordinate Structure Constraint. Alternatively, Koster proposes that both sentences involve a parallel construal and that the difference between them may be attributed to the properties of pied-piping.⁴¹ In this view, the coordinated phrases in (226) involve the abstract representation given in (227). According to Koster, the first XP checks the features of *and* (or [*and* XP]). This mechanism expresses the fact that an XP of a given type in the complement position typically requires an XP of the same type in the Specifier position.

- (227) [XP [*and* XP]]

If the Specifier position is filled by the checking phrase only, the adjacency between the two conjuncts is derived, as in (226)b. In this example, both conjuncts are noun phrases, and the noun phrase in the Specifier position (*Jan*) checks the features of *en* (or [*en Marie*]); see (228):

- (228) *Hij heeft [[NP Jan] [en [NP Marie]]] gezien*
he has J. and M. seen
 ‘He saw John and Mary.’

(Koster 2000: 18)

⁴¹ Note that in this paper, Koster (2000) proposes an extension of the concept of pied-piping that is not standardly assumed in the literature on the topic, given that this phenomenon is traditionally associated with movement.

However, as in the cases of standard pied-piping, the checking phrase can be contained in a larger constituent (e.g., a VP or AgrOP). This happens in sentences involving non-adjacency of conjuncts, as in (226)a. In this case, the checking phrase (*Jan*) and the elements to be checked (*en* or *en Marie*) are the same. The difference is that in (226)a, the Specifier position is filled both by the checking phrase and by a larger constituent containing it, as depicted in (229).

- (229) Hij heeft [_{AgrOP} [_{AgrOP} [_{NP} *Jan*]] [_{AgrO} [_{VP} gezien]]] [_{NP} *en Marie*]]
he has J. seen and M.
 (Koster 2000: 18)

As in the standard cases of pied-piping, the extension of the checking phrase has a limit: it cannot go beyond clausal boundaries. This explains the Right Roof Constraint on this construction. As illustrated by the ungrammaticality of (230), the checking phrase (*Jan*) cannot be contained in a subject clause CP that does not contain the elements to be checked (*en Marie*).

- (230) *_[CP dat hij Jan gezien heeft] is duidelijk en Marie
that he J. seen has is clear and M.
 ‘It is clear that he saw Jan and Marie.’
 (Koster 2000: 18)

Koster also demonstrates that all forms of parallel construal have the properties of Ross’ Coordinate Structure Constraint. Among other things, this entails that the first conjunct cannot be moved without the second, as is illustrated in (231):

- (231) *_{Jan heb ik} [_{t en Marie}] gezien
J. have I and M. seen
 ‘I saw Jan and Marie.’
 (Koster 2000: 19)

In light of these facts,⁴² Koster proposes that extraposition does not have the properties of movement but those of parallel construal. Let us examine in detail how this approach works in the case of RRC extraposition.

According to Koster, relative clauses are analyzed in terms of parallel construal. This is schematically represented in (232), where the relative clause is taken to provide a further specification of the head placed in the Specifier of the colon.⁴³

- (232) [NP [NP een vrouw] [: [CP die alles wist]]]
 a woman who everything knew
 ‘a woman who knew everything’

(Koster 2000: 22)

The extraposition of relative clauses is then derived from the property of pied-piping. If only the head occurs in the Specifier position, there is adjacency between the head and the relative clause (as in (233)a). If the Specifier is filled by a larger constituent containing the head, the extraposed order is derived (as in (233)b).

- (233) a. Ik heb [NP [NP een vrouw] [: [CP die alles wist]]] gezien
 I have a woman who everything knew seen
 ‘I saw a woman who knew everything.’
 b. Ik heb [[AgrOP [NP een vrouw]] gezien] [: [CP die alles wist]]
 I have a woman seen who everything knew

(Koster 2000: 23)

In (233)b, the checking head (*een vrouw* ‘a woman’) is included in AgrOP, but more inclusive phrases can occupy the Specifier position of a parallel construal. For instance, if the antecedent of an extraposed RRC is in [Spec, IP], the entire IP surfaces in the Specifier position, but if the antecedent is a topic, the minimal CP surfaces in this position, as illustrated in (234):

⁴² Koster (2000) also discusses evidence from specifications found in equatives, which I will not address here.

⁴³ As for the contrast between RRCs and ARCs, Koster (2000) claims that the colon indicates set interaction in the case of RRCs and set union in the case of ARCs. He also suggests that RRCs and ARCs can be distinguished by the level of attachment of the specifying conjunct (NP or DP).

- (234) a. [_{IP} [_{IP} [Een vrouw] heeft hem gezien] [: [die alles wist]]]
a woman has him seen who everything knew
 ‘He saw a woman who knew everything.’
 b. [_{TopP} [_{TopP} [Een vrouw] heeft hij t gezien] [: [die alles wist]]]
a woman has he seen who everything knew
 (Koster 2000: 23)

As in the cases of standard coordination discussed above, the extension of the checking phrase has a limit: it cannot go beyond the minimal CP containing the relative clause. This explains the ungrammaticality of (235):

- (235) * [_{CP} Dat hij een vrouw gezien heeft] is duidelijk die alles wist
that he a woman seen has is clear who everything knew
 ‘It is clear that he saw a woman who knew everything.’
 (Koster 2000: 23)

Moreover, the Coordinate Structure Constraint also applies to relative clause-extrapolation. According to Koster (2000), this is confirmed by the impossibility of having the first part of the construction moved away from the relative clause in sentences like (236):

- (236) *Een vrouwi heeft hij [ti die alles wist] gezien
a woman has he who everything knew seen
 ‘He saw a woman who knew everything.’
 (Koster 2000: 23)

Koster’s (2000) account is conceptually attractive because it unifies a variety of apparently unrelated constructions under the label of *parallel construal*. However, it faces substantial empirical and theoretical difficulties, which I briefly note below.

From an empirical point of view, the biggest problem is that Koster’s approach overgenerates in a number of ways. If the constituent that surfaces in the Specifier position may belong to any category (within the minimal CP domain), extrapolation should take place from any constituent, and this is simply not true. As discussed in Section 4, in some languages, there are important restrictions on RRC extrapolation. For instance, in CEP, RRC extrapolation cannot take place from strong noun phrases or from embedded positions. As can be easily concluded, Koster’s (2000) analysis leaves these restrictions unexplained.

Second, the syntactic structure proposed by Koster allows the Specifier and the complement positions of a parallel construal to be filled by unequal categories. For example, an extraposed RRC taking an object as an antecedent would have an AgrOP in the Specifier position and an RRC in the complement position. As De Vries (2002, 2009) notes, the problem is that in this case, the constituents are not of the same category nor functionally equivalent, which is not allowed in standard coordination (e.g., *He looks great and at me).

Another empirical problem with this analysis concerns the assumption that the categorial status of the constituent in the Specifier position depends on the structural position of the antecedent. This assumption works in a language like Dutch, where it is generally assumed that the verb is spelled out in V (apart from verb second in the finite verb main clauses) but raises some problems for languages displaying V-to-I movement. As a point of departure, let us examine a sentence containing an RRC extraposed from an object, as in (237) (repeated from (150)) from CEP:

- (237) Encontrei uma rapariga ontem que perguntou por ti.
met:1SG a girl yesterday that asked for you
 ‘I met a girl yesterday that asked about you.’

Under the assumption that the checking phrase in the first conjunct extends until the structural position occupied by the antecedent, in a sentence like (237), the Specifier position would be filled by the VP, and the complement position would be filled by the RRC, as depicted in (238):⁴⁴

- (238) [_{IP} encontrei_i [[_{VP} uma rapariga_j [_{VP} ontem t_i t_j]] [: que perguntou por ti]]]
met:1SG a girl yesterday that asked for you

This structure is problematic because V-to-I movement violates the Coordinate Structure Constraint, which prevents movement from one conjunct in a coordinate structure unless movement also occurs from the other conjunct (Ross 1967).

⁴⁴ In accordance with the analysis proposed in Section 5.3, I assume that in the structural representation given in (238), the direct object is scrambled (i.e., left-adjoined to VP). Note, however, that this is not crucial here: the same line of reasoning would hold if the object were in its base-position.

Therefore, if RRC extraposition were derived by the parallel structure proposed by Koster, a sentence like (237) would be ungrammatical, contrary to fact.⁴⁵

From a theoretical point of view, Koster's proposal is also problematic because it is not compatible with the raising analysis of relative clauses (because it assumes that the head and the RRC are generated separately). Therefore, among other things, it cannot account for the reconstruction effects discussed in Chapter 2, which suggest that the head of the RRC is generated in an RRC-internal position, as witnessed in (239):

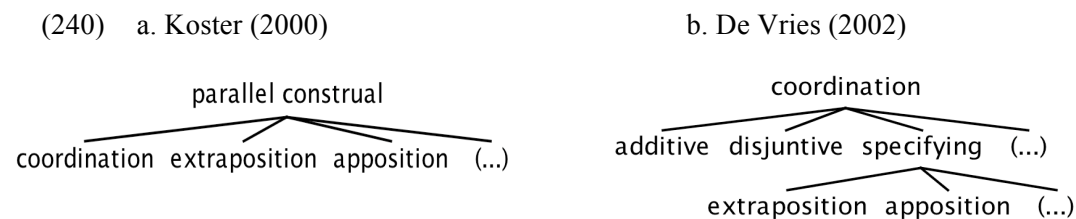
(239) Bill liked the [stories about himself_i] which John_i told.

Capitalizing on Koster's (2000) proposal, De Vries (2002) proffers a different account to extraposition, which overcomes some of the drawbacks of Koster's proposal. The next section is devoted to the presentation and discussion of this analysis.

6.1.2. De Vries (2002)

Building on ideas by Koster (2000), De Vries (2002) proposes the *specifying coordination plus ellipsis analysis* of extraposition. There are at least three major differences between the two proposals.

First, De Vries (2002) explicitly analyzes extraposition as coordination. Whereas Koster (2000) resorts to the concept of parallel construal and claims that coordination and extraposition are particular subcases of parallel construal, De Vries (2002) assumes the concept of coordination as the encompassing notion. The conceptual divergence between the two proposals is depicted in (240)a-b:



⁴⁵ Here, I abstract away from the movement of the subject to [Spec, IP], which would also violate the Coordinate Structure Constraint. See De Vries (2002, 2009) for similar problems in Dutch.

De Vries (2002) claims that coordination is a syntactic construction with varying semantics. Aside from the traditional types of coordination (such as additive and disjunctive), there is another type called specifying coordination that is involved, e.g., in extraposition, apposition and other constructions (such as parenthesis and left- and right-dislocation; see De Vries 2009 for a general overview). In all of these constructions, the second conjunct provides an alternative description, an example, or a property of the first conjunct.

The second difference between the two proposals concerns the categorial status of the constituents that occupy the Specifier and the complement position. Like Koster, De Vries assumes that RRC extraposition is obtained according to the scheme in (241): the antecedent is generated within the Specifier position, and the extraposed RRC is generated within the complement position of an abstract head.⁴⁶

(241) [_{CoP} [... antecedent ...] [Co [... RRC ...]]]

However, unlike Koster, De Vries proposes that the constituents that occupy the Specifier and the complement positions are of the same category. In this framework, the first conjunct may range from VP to CP, depending on the position of the antecedent. The second conjunct has the same categorial status as the first conjunct; it repeats the material contained in the first conjunct, adding the extraposed RRC in its canonical position. Then, the repeated material is phonologically deleted. This is illustrated in (242), where the antecedent of the extraposed RRC is a direct object. Here, both conjuncts are represented as involving the AgrOP-level of projection (under the assumption that in Dutch the object moves to [Spec, AgrOP], for reasons of case).

⁴⁶ The structure in (241) involves an abstract coordinator that is semantically specialized: it constitutes an asymmetric relationship of specification between the two conjuncts. Koster (2000) symbolically represents this relator using a colon; De Vries (2002) employs an ampersand plus a colon ‘&:’. Here, I will simply use the more general denotation Co for the coordinative head.

- (242) [_{CP} Ik heb ... [_{CoP} [_{AgrOP-1} de man gezien] [_{Co}
I have the man seen
 [_{AgrOP-2} [_{DP} ~~de man~~ die zijn tas verloor] ~~gezien~~]]]]
the man who his bag lost seen
 (De Vries 2002: 241)

De Vries (2002, 2009) proposes that the deletion used in the specifying coordination plus ellipsis analysis has three important characteristics:

- The deletion may involve nonconstituents and discontinuous material.
- The deletion of all material that is repeated is obligatory.
- The deletion is directed forward (i.e., left-to-right).

These characteristics can be observed in the structural representation in (242). The fact that the deletion may involve nonconstituents is illustrated by the deletion of *de*, ‘the’, and *man*, ‘man’, which do not form a constituent. The requirement that all of the repeated material must be deleted is confirmed by the deletion of *de*, ‘the’, *man*, ‘man’, and *gezien*, ‘seen’, which are the elements repeated from the first conjunct. Finally, the demand on forward deletion can be demonstrated by the fact that deletion only targets constituents in the second conjunct. For more details, see De Vries (2002, 2009) and G. de Vries (1992).

The third difference between the two proposals concerns the (non)-autonomous syntactic status of extraposition. According to Koster, extraposed and non-extraposed orders involve the same grammatical configuration. As already shown in Section 6.1.1, standard coordination and relative clauses involving adjacency (between conjuncts and between the head and the relative clause, respectively) are analyzed in terms of parallel construal, consisting of a Specifier, head and complement configuration. In this view, extraposition does not involve a different grammatical configuration, being simply derived from the property of pied-piping (i.e., the possibility of having a larger constituent in the Specifier position containing the checking phrase). In contrast, under the specifying coordination plus ellipsis account, the extraposed and the non-extraposed configurations involve a different derivational story. In this case, the specifying coordination configuration is present in sentences involving extraposition but not in sentences involving the normal (i.e., non-extraposed) order.

Now, after this brief comparison between the two proposals, let us determine if De Vries' approach is capable of overcoming the drawbacks of Koster's analysis mentioned above. First, consider the violation of the Coordinate Structure Constraint caused by V-to-I movement. Under the specifying coordination plus ellipsis analysis, this problem does not arise because there is a representation of the verb inside the second conjunct as well. Therefore, the verb is moved in an across-the-board fashion⁴⁷, as shown in the schematic representation in (243):

$$(243) \quad [IP \ V \ [CoP \ [VP \ DO \ [VP \ adverb \ t_v \ t_{DO}]] \ [Co \ [VP \ \cancel{DO} \ RRC \ [VP \ adverb \ t_v \ t_{DO}]]]]]$$

Another advantage of De Vries' approach is that it eliminates unbalanced coordination. Recall that under Koster's approach, the Specifier and the complement positions of the parallel construal can be filled by unequal categories. Under De Vries' analysis, this problem does not arise because both conjuncts are of the same category (e.g., a VP, as in (243)).

Finally, from a theoretical point of view, De Vries' analysis has the advantage of being compatible with the raising analysis of RRCs because the head is syntactically present in the second conjunct.

However, there is one non-trivial problem that remains unsolved in this proposal. As in Koster's analysis, the specifying coordination plus ellipsis analysis overgenerates in a number of ways: it predicts that RRC extraposition from any constituent should be allowed. However, though this may be true of languages like Dutch, it is simply not correct for languages like CEP, where RRC extraposition cannot take place, e.g., from strong noun phrases and embedded positions (see Section 4.1).

This may be a serious problem for uniform/universalist approaches to RRC extraposition. However, it can be taken as an advantage for the dual approach to RRC extraposition advocated here because it offers a straightforward explanation for the fact that different stages of the same language (and different languages) may differ on

⁴⁷ As is well known, across-the-board extraction is not subject to the Coordinate Structure Constraint. Rules apply in an across-the-board fashion if they affect all conjuncts in a coordinate structure at the same time. This is what happens in (243): V-to-I movement extracts the V out of both conjuncts.

the properties of RRC extraposition. More precisely, the fact that the specifying coordination plus ellipsis analysis cannot explain the restrictions on RRC extraposition found in CEP is a welcome result. It corroborates the hypothesis that RRC extraposition in CEP is derived from stranding, whereas RRC extraposition in earlier stages of Portuguese is derived from specifying coordination plus ellipsis.

To provide further support for this claim, I show how the specifying coordination plus ellipsis analysis can account for the properties of RRC extraposition in earlier stages of Portuguese in the next section. See Section 7 for parametrization, language change and differences.

6.2. Deriving the properties of RRC extraposition in earlier stages of Port.

In Section 4, I showed that CEP and earlier stages of Portuguese behave differently with respect to the following properties:

- A. *The definiteness effect*
- B. *Restriction on extraposition from embedded positions*
- C. *Restriction on extraposition from pre-verbal positions*

I propose that the different restrictions to RRC extraposition found in the diachronic (and cross-linguistic) dimension can be explained under a dual approach to RRC extraposition. The rationale behind this proposal is that RRC extraposition is not a unitary phenomenon; it may involve stranding or specifying coordination plus ellipsis. Languages and different stages of the same language differ with respect to the type of extraposition they display. Considering (in particular) the case of Portuguese, the hypothesis is that RRC extraposition is generated by stranding in CEP and by specifying coordination plus ellipsis in earlier stages of Portuguese.

In Section 5.4, I demonstrated that the properties of RRC extraposition in CEP can be accounted for in terms of stranding. Now, I show how the contrasting properties of RRC extraposition in earlier stages of Portuguese can be derived from the specifying coordination plus ellipsis analysis.

A. *The definiteness effect*

In earlier stages of Portuguese, extraposed RRCs can take strong noun phrases as their antecedent. This property can be straightforwardly derived under the specifying coordination plus ellipsis analysis because there is no movement relationship between the visible antecedent and the extraposed RRC. As illustrated in (244) (repeated from (90)), the strong noun phrase *aquella dia* ‘that day’ in the first conjunct is a constituent: it is detached from the relative clause and base-generated in the first conjunct of the coordinate structure. In contrast, the strong noun phrase *aquella dia* ‘that day’ in the second conjunct is not a constituent (because there is no constituent that includes the determiner and the noun and excludes the RRC). However, this is not a problem because it is the DP (containing the antecedent and the RRC) that undergoes leftward movement. Given that deletion may target nonconstituents, the repeated material in the second conjunct is deleted, and RRC extraposition is derived.

- (244) [mas [CoP [IP [DP *aquella dia*]_i sem falha aveo t_i]
but that day without fail came
 [Co [IP [DP ~~*aquella dia*~~ que foram i todos]_i ~~sem falha aveo t_i]]]]
that day that went there all without fail came
 ‘but the day everyone went there came without fail.’~~

B. *Restriction on extraposition from embedded positions*

In earlier stages of Portuguese, extraposed RRCs can take the object of prepositions as their antecedent. The same reasoning as above applies to this case: as illustrated in (245), the PP *de mui poucos* ‘of very few’ in the first conjunct is a constituent because it is detached from the relative clause and base-generated in the first conjunct of the coordinate structure. In contrast, *de mui poucos* ‘of very few’ in the second conjunct is not a constituent. However, this is not a problem because it is the PP (containing the RRC) that undergoes leftward movement. Then, the repeated material in the second conjunct is deleted, leading to RRC extraposition.

- (245) [_{CP} que [_{CoP} [_{IP} [_{PP} de mui poucos]_i sabemos t_i]
 that of very few know:1PL
 [Co [_{IP} [_{PP} ~~de mui poucos~~ que bebesses vinho]_i sabemos t_i]]]]
 of very few that drink:IMPERF.SUBJ wine know:1PL

C. *Restriction on extraposition from pre-verbal positions*

C1. *Restriction on extraposition from pre-verbal subjects*

In earlier stages of Portuguese, extraposed RRCs may take pre-verbal subjects as an antecedent. This can be derived by resorting to IP level coordination; see (246):

- (246) [_{CoP} [_{IP} S V DO] Co [_{IP} S RRC ~~V DO~~]]

From a comparative perspective, the fact that CEP does not allow RRC extraposition from pre-verbal subjects is surprising. As shown in Section 5.4, the explanation for the pattern of ungrammaticality in CEP depends upon the Interpretative Principle in (247). RRC extraposition from [Spec, IP] is not allowed because such position is ambiguously filled by topic and non-topic elements.

(247) Interpretative Principle

The antecedent of an extraposed RRC must occur in a position non-ambiguously interpreted as non-topic (in Kuroda's 2005 sense).

Apparently, nothing prevents RRC extraposition in earlier stages of Portuguese from being subject to the same semantic restrictions as CEP. However, as shown in Section 4.3, there is strong empirical evidence suggesting that earlier stages of Portuguese (and other languages) allow for it.

Somewhat tentatively, I would like to suggest that CEP and earlier stages of Portuguese may resort to different strategies to resolve the ambiguity expressed in (247). Whereas in CEP the ambiguity associated with [Spec, IP] is resolved syntactically and prosodically (through subject inversion), in earlier stages of Portuguese, it may be resolved only prosodically. In this case, a constituent in [Spec,

IP] can be unambiguously interpreted as non-topic if it is prosodically marked by a pitch accent.⁴⁸

As elaborated below, this may suggest that there is a language split as far as the codification of semantic information is concerned. Some languages codify the topic/non-topic status of the subject prosodically and syntactically (as may be the case of CEP), whereas other languages (and different stages of the same language) may codify it only prosodically (as seems to be the case for earlier stages of Portuguese).

C2. Restriction on extraposition from other pre-verbal constituents

The behavior of RRC extraposition from other pre-verbal positions is summarized in Table 8.

**Table 8. Extraposition from other pre-verbal positions
(earlier stages of Portuguese)**

		Earlier stages of Portuguese
<i>Extraposition from other pre-verbal positions</i>	<i>wh</i> -constituents	+
	emphatic/evaluative phrases	+
	preposed foci	+
	topics	-

Extrapolated RRCs taking a *wh*-constituent, an emphatic/evaluative phrase or a preposed focus as an antecedent can be derived without further ado by resorting to coordination of a dedicated functional projection of the left periphery (see (248)). For ease of representation, the functional projections dedicated to each discourse value in the left periphery (e.g., FocusP, EvaluativeP) are represented by FP.

$$(248) \quad [\text{CoP} [\text{FP } \textit{wh-const./emphat.phrase/prepo.focus} \text{ S V }] \\ \text{Co} [\text{FP } \textit{wh-const./emphat.phrase/prepo.focus} \text{ RRC S-V }]]$$

⁴⁸ Here, I assume that the kind of prosodic prominence that serves to mark focused constituents is the pitch accent (see, e.g., Avesani and Vayra 2003).

Comparing (248) and (246), it becomes clear that in (248) the coordinate structure involves a higher level of projection than in the case of RRC extraposition from the subject.

The impossibility of extraposition from topics follows from the Interpretative Principle in (247). Although extraposition generated by specifying coordination plus ellipsis is not syntactically constrained (because the second conjunct can be freely attached at any structure level (within the minimal CP domain)) the semantic principle in (247) prevents extraposed RRCs from taking topics as an antecedent. As mentioned in Section 5.4, this is because under a split-CP approach, the position occupied by topicalized constituents is non-ambiguously interpreted as topic.

C3. New contexts of extraposition

In earlier stages of Portuguese, RRC extraposition can occur from scrambled constituents in [Spec, IP] (see Section 4.3). This can be accounted for by resorting to coordination at the IP level, as schematically represented in (249):

(249) [_{CoP} [_{IP} *scrambled constituent* S V] Co [_{IP} ~~*scrambled constituent*~~ RRC S-V-]]

The reason why RRC extraposition from scrambled constituents in [Spec, IP] is not available anymore in CEP is independently explained by the loss of IP scrambling in the history of Portuguese. According to Martins (2002), the loss of IP scrambling is a result of a change in the properties of the AgrS functional head. AgrS ceased to allow multiple Specifiers, i.e., it lost the option for being associated with an Attract-all-F EPP feature. Therefore, IP scrambling disappeared because a structural position for scrambled objects ceased to be available in the IP space. In this view, it is easy to see why RRC extraposition cannot be derived from IP scrambling in CEP; a structural position for scrambled constituents is not available anymore in the IP space.

6.2.1. Summary

This section presents my analysis of the properties of RRC extraposition in earlier stages of Portuguese in terms of the specifying coordination plus ellipsis analysis proposed by De Vries (2002). The results are summarized in Table 9. The specifying coordination plus ellipsis analysis is evaluated in the following way: a plus indicates

that the analysis can derive the property straightforwardly or with reference to independent principles; a minus indicates that it cannot.

Table 9. Evaluation of the specifying coordination plus ellipsis analysis of RRC extraposition in earlier stages of Portuguese

	Derived by the stranding analysis	
<i>A. Extraposition from strong noun phrases</i>	+	
<i>B. Extraposition from embedded positions</i>	+	
<i>C. Pre-verbal constituents</i>	Extraposition from pre-verbal subjects	+
	Extraposition from <i>wh</i> -constituents	+
	Extraposition from emphatic/evaluative phrases	+
	Extraposition from preposed foci	+
	No extraposition from topics	+

On the basis of these results, I conclude that the specifying coordination plus ellipsis analysis accounts for the properties of RRC extraposition identified in Section 4.3. It goes without saying that if RRC extraposition in earlier stages of Portuguese was generated by stranding, not all of these properties would be derived.

Despite the success of the specifying coordination plus ellipsis analysis in deriving the properties of RRC extraposition in earlier stages of Portuguese, it is worth noting that this analysis also uncovers some problems, e.g., with respect to scope relations. In the next section, I discuss one scope relation that can be documented in historical Portuguese: the licensing of subjunctive mood in extraposed RRCs.

6.3. Problems with the specifying coordination plus ellipsis analysis

The specifying coordination plus ellipsis analysis faces some problems in explaining the scope relations that can be established between the matrix and the extraposed RRC. Given the limitations of historical inquiry, I will confine this section to the discussion of one scope relation that is documented in the written sources: the licensing of subjective mood in extraposed RRCs.⁴⁹

⁴⁹ I thank Jairo Nunes (p.c.) for drawing my attention to these facts.

The choice of mood in RRCs is not determined lexically because it occurs in the context of verbal complementation. Subjunctive RRCs are typically licensed in a set of intensional environments created, for example, by strong intensional predicates, negation, future tense, interrogatives, conditionals or imperatives (see Quer 1998: 105; Quer 2001: 91). Moreover, it is standardly assumed that some of these contexts, such as intensional predicates or negation, only license subjunctive RRCs in their complement or c-command domain (see Quer 1998: 105). Therefore, it can be assumed that in an RRC like (250), the subjunctive mood is licensed by the c-commanding negative marker *não* ‘not’:

- (250) Não abro anexos que possam ter vírus.
not open attachments that might:PRES.SUBJ have:INF virus
 ‘I do not open attachments that might have a virus.’

Crucially, subjunctive mood is also licensed in extraposed RRCs. See (251), from CEP:

- (251) Não apareceu uma única pessoa ontem que tivesse o
not showed.up a single person yesterday that had:IMPERF.SUBJ the
perfil adequado.
profile appropriate
 ‘Not even a single person showed up yesterday that had the appropriate profile.’

This can be easily accounted for under the assumption that extraposed RRCs in CEP are derived from stranding; a subjunctive RRC is licensed because an RRC stranded in a VP-internal position is in the c-command domain of the negation. However, the situation is not as straightforward in the case of extraposed subjunctive RRCs derived from specifying coordination plus ellipsis. In this case, extraposed RRCs may also surface in the subjunctive, as illustrated in (252), taken from a 15th-century Portuguese manuscript:

- (252) ca nom ha cousa no mundo que tanto deseje
since not has thing in.the.world that so.much want:PRES.SUBJ.ISG
como honra de cavallaria
as honor of cavalry
 ‘since there is nothing in the word that I want so much as the honor of cavalry.’
 (Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)

However, it is not uncontroversial that in (252), the extraposed RRC in the second conjunct is in the scope of the negation. Recall that, according to De Vries (2002), the categorial status of conjuncts depends on the position of the antecedent. Under this assumption, a sentence like (252) involves coordination at the VP-level of projection, as depicted in the simplified structure given in (253):

- (253) [_{CP} *ca* *nom* *ha* [_{CoP} [_{VP} *cousa* *no* *mundo*] [_{Co}
since not has thing in.the world
 [_{VP} ~~*cousa*~~ *que* *tanto* *deseje* *como* *a* *honra* *de* *cavallaria* ~~*no*~~ ~~*mundo*~~]]]]
thing that so.much want:SUBJ as the honor of cavalry in.the world

In this environment, the negative marker *nom* ‘not’ is not inside the conjuncts, and the subjunctive RRC surfaces in the second conjunct. Then, the problem arises as to how the subjunctive mood can be licensed in the second conjunct.⁵⁰

De Vries (2005) claims that in a configuration like (254), the second conjunct YP, as opposed to the first XP, is *invisible* for the higher context in terms of c-command.⁵¹ This entails that XP but not YP is c-commanded by RP, i.e., by some phrase higher up in the syntactic context of the coordination phrase.⁵²

⁵⁰ A similar problem arises in (253) for the interpretation of the word *cousa* ‘thing’. *Cousa* is a contextually negative word, i.e., a word that receives a negative meaning from a negative word in the relevant context (see Martins 2008). Again, it is not clear how the negative meaning of *cousa* arises in the coordinate structure in (253).

⁵¹ In the schematic representation in (254), I abstract away from the *three-dimensional* or *behindance* approach to coordination proposed by De Vries (2004, 2005). De Vries (2005) explains the invisibility of the second conjunct by means of the relation *behindance*. Resorting to a spacial metaphor, he claims that conjuncts are behind each other in a three-dimensional structure. The basic idea is that there are two types of merge: one based on dominance (d-Merge) and the other based on behindance (b-Merge). The latter type blocks c-command relations from the higher context. Applying this to the structure in (254), he claims that Co and YP are ‘behind’ (i.e., b-included in) Co’. Therefore, they are not c-commanded by XP, according to the definition of c-command in (i).

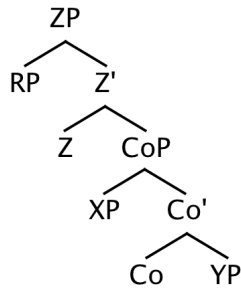
(i) *C-command*

If Merge (A,B) then A c-commands B and all the constituents d-included in B.

⁵² According to De Vries (2005), the invisibility of the second conjunct also explains the lack of c-command between conjuncts. However, note that there is no consensus in the literature on Portuguese regarding this issue. Matos (1995) proposes that the fact that an element inside the first conjunct may c-command an element inside the second conjunct of a coordinate structure explains the ungrammaticality of sentences like (i). This claim relies on the assumption that if CoP and the first conjunct in its Specifier are of the same category (IP in (i)), the computational system treats them as being non-distinct, i.e., as segments of the same category.

(to be continued)

(254)



If we assume that second conjuncts are invisible for the higher context, we predict that in a structure like (253), the negation does not c-command the RRC in the second conjunct. Therefore, the subjunctive mood should not be licensed, and sentences like (252) should not be allowed, contrary to fact. Needless to say, no problem would arise if second conjuncts were not taken to be invisible for the higher context. In this case, the subjunctive mood in (253) would be licensed by a c-commanding negative marker located above CoP.

However, maintaining the idea that second conjuncts are invisible for the higher context, there are at least two possible ways to circumvent the problem. The first one is to assume that when RRC extraposition is involved, conjuncts always involve the

-
- (i) * Ela_i viu esse livro na livraria mas a Maria_i decidiu não o comprar.
she saw that book in.the bookstore but the M. decided not it buy:INF
 ‘She saw that book in a bookstore, but Maria decided not to buy it.’

(Matos 1995: 311)

Chaves (2007), however, claims that this correlation is not correct because in some contexts, the allegedly impossible binding is allowed:

- (ii) Ela_i prometeu que chegava a horas e eu acredito sempre nas
she promised that arrive on time and I believe always in.the
 promessas da Ana_i.
promises of.the A.

‘She promised she would arrive on time and I always believe Ana’s promises.’

(Chaves 2007: 44)

In fact, the grammaticality of (ii) seems to suggest that (i) is not ungrammatical due to a syntactic binding principle but perhaps because of a discourse principle operating across sentences. (Notice that co-reference of the type illustrated in (i) is usually not possible across sentences as well (e.g., She came home. #Mary was tired)).

minimal CP projection (or IP, if CP is not projected) containing the antecedent.⁵³ In this case, the negation is contained within the second conjunct, and the RRC is in its c-command domain, as sketched in (255):

- (255) [_{CoP} [_{CP} ca nom ha cousa no mundo] [Co [_{CP} ~~ea~~ ~~nom~~ ~~ha~~ ~~cousa~~
 since not has thing in.the world since not has thing
 que tanto deseje como a honra de cavallaria ~~no mundo~~]]]
 that so.much want as the honor of cavalry in.the world

Another hypothesis is to assume that the invisibility⁵⁴ of the second conjunct depends on the particular coordinative head involved. According to De Vries (2009), there are two different types of specifying coordinators: (i) non-restrictive (parenthetical)

⁵³ The idea that conjuncts are generally root CPs is proposed by Wilder (1994) for normal coordination. Under this approach, the apparent coordination of small conjuncts is derived by ellipsis (*i.e.*, by the deletion of PF material). The author rejects the idea that coordination can be applied at any level of the syntactic structure and claims that this is simply an effect of ellipsis in non-initial conjuncts.

⁵⁴ De Vries (2007) defines syntactic *invisibility* as the inability to maintain c-command-based relations with elements of the host. According to this proposal, paratactic constituents (involved in coordination, apposition and parenthesis) are invisible for the higher context. This amounts to saying that they are structurally integrated with the host but not in a way that can be detected by c-command relations. De Vries (2007) investigates the relationships that are based on the notion of c-command (e.g., binding, movement, and polarity items) and claims that the invisibility of paratactic constituents leads to a number of predictions that are indeed borne out (see (i)). For additional details and exemplification, see De Vries (2007).

(i) *Predictions by invisibility*

- a. *No movement*: there cannot be movement from paratactic constituents into the host.
- b. *No idiom chunks*: no idiom can be split across a paratactic boundary.
- c. *No Q-binding*: a pronoun in paratactic constituents cannot be bound by a quantified expression in the host.
- d. *No A-binding*: a reflexive in paratactic constituents cannot be bound by an antecedent in the host.
- e. *No Condition B effects*: a pronoun in paratactic constituents does not cause Condition B effects with respect to a coreferent expression in the host.
- f. *No Condition C effects*: an R-expression in paratactic constituents does not cause Condition C effects with respect to a coreferent expression in the host.
- g. *No NPIs*: no negative polarity item in paratactic constituents can be licensed by an operator in the host.
- h. *No PPI effects*: no positive polarity item in paratactic constituents can be disqualified by an operator in the host.
- i. *No dependent Force*: the illocutionary Force of paratactic constituents is independent of the host's.
- j. *No dependent Mood*: the Mood of paratactic constituents is independent of the host's.

(De Vries 2007: 208)

specifying coordinators, which are involved, e.g., in appositions, parentheticals and ARCs; and (ii) restrictive specifying coordinators, which are involved in extraposition. Under De Vries' account, the second conjunct is invisible in terms of c-command in both cases. As mentioned above, this leaves the question of why extraposed RRCs can surface in the subjunctive mood unexplained. To account for the scope relations necessary for subjunctive mood licensing, we may alternatively suggest that the invisibility of the second conjunct depends on the particular coordinative head involved. Non-restrictive (parenthetical) specifying coordinators trigger the invisibility of the second conjunct, whereas restrictive specifying coordinators do not⁵⁵. This would explain why parenthetical material is invisible for c-command relations (see De Vries 2005, 2007, 2009), whereas extraposed material is not.

However, further research is necessary in this domain to warrant the validity of these suggestions. In particular, it is necessary to explore in contemporary languages that generate RRC extraposition by specifying coordination plus ellipsis other scope relations between the extraposed material and the higher context, which unfortunately cannot be tested in earlier stages of Portuguese.

6.4. Conclusion

This section discusses the syntactic nature of RRC extraposition in earlier stages of Portuguese. Based on De Vries (2002), I have claimed that RRC extraposition in earlier stages of Portuguese involves a special type of coordination, called specifying coordination. Under this approach, a constituent containing the visible antecedent is related by coordination to a constituent containing the extraposed RRC. Repeated material is phonologically deleted. Schematically, RRC extraposition appears as follows:

(256) ... [CoP [XP1 antecedent YP] [Co [XP2 [~~antecedent~~ RRC] YP]]]

⁵⁵ Under the three-dimensional approach to coordination proposed by De Vries (2005) (see fn. 51), this could be implemented as follows: the non-restrictive (parenthetical) specifying coordinator triggers behindance (i.e., b-Merge), whereas the restrictive specifying coordinator triggers dominance (i.e., d-Merge).

RRC extraposition generated by (256) is an extremely flexible operation, which allows, among other things, extraposition from any constituent. This explains why earlier stages of Portuguese, as opposed to CEP, allow extraposition from strong noun phrases, embedded positions and pre-verbal subjects. As mentioned in Section 5, the constrained nature of RRC extraposition in CEP can be partially derived from the restrictions on movement inherent to the stranding analysis.

As a final point, let me mention that the study of RRC extraposition in earlier stages of Portuguese faces obvious difficulties, given the limited nature of the written sources and the impossibility of manipulating data. This fact was particularly evident in the impossibility of testing different scope relations between the matrix and the extraposed RRC. For this reason, the cross-linguistic comparison developed in Section 4.2 was crucial in showing that earlier stages of Portuguese pattern like Germanic languages in the properties of RRC extraposition. Therefore, comparative research can provide an interesting and fruitful method to overcome the limitations of historical data. In particular, studying the behavior of RRC extraposition in Germanic languages can provide the means to understand more about the syntax of RRC extraposition in earlier stages of Portuguese than if we were limited to written sources.

7. Comparative remarks

This section is devoted to comparative remarks on the syntax of RRC extraposition. Here, I tentatively offer an integrated account of the facts of variation and change considered in this chapter. More concretely, I show that the dual approach adopted here provides an important tool to explain the synchronic and diachronic variation found within the same language and across languages. I begin by presenting how the dual approach to RRC extraposition can explain the change that took place in the history of Portuguese (Section 7.1). Then, I demonstrate how the dual approach can help us to understand the variation found in the cross-linguistic dimension (Section 7.2).

7.1. Diachronic change in RRC extraposition

In this section, I argue that the dual approach to RRC extraposition proposed in this chapter provides a viable basis for understanding why extraposed RRCs ceased to be generated by the specifying coordination plus ellipsis analysis in the history of Portuguese. According to the analysis outlined above, the RRC extraposition in earlier stages of Portuguese is generated by specifying coordination plus ellipsis, whereas RRC extraposition in CEP is generated by stranding. As shown in Sections 5 and 6, this explains the different behavior of extraposed RRCs with respect to: (i) extraposition from strong noun phrases, (ii) extraposition from embedded positions, and (iii) extraposition from pre-verbal positions.

This diachronic development entails that the starting point has a structure like (257)a, where the visible antecedent occurs in the first conjunct of a coordinate structure. Then, at a certain period in the history of Portuguese, RRC extraposition began to be generated by a different structure (see (257)b), where the visible antecedent is internally generated.

- (257) a. specifying coordination plus ellipsis structure
 ... [CoP [XP₁ antecedent YP] [Co [XP₂ [~~antecedent~~ RRC] YP]]]
 ↓
 b. stranding structure
 ... [antecedent_i YP [t_i RRC]]

Keeping in mind Lightfoot's (1991, 1999) insights into the relation between language change and language acquisition, this implies that positive evidence triggering the specifying coordination structure ceased to be available to the learners. I submit that such evidence is found in the context of RRC extraposition involving a strong noun phrase or the object of a preposition as an antecedent. When exposed to these contexts, children in earlier stages of Portuguese knew that the antecedent was externally (and not internally) generated because no movement chain could be established between the visible antecedent and the RRC-internal position. As shown in Section 6.2 above, this is because strong noun phrases and PPs do not form a constituent (excluding the RRC).

Capitalizing on what we know about the history of Portuguese, I would like to suggest that the relevant contexts of extraposition were robustly attested in the

primary linguistic data until the 16th century. After this period, their frequency significantly decreased due to an independent change that took place in the history of Portuguese: the loss of IP scrambling. Martins (2002) reports that earlier stages of Portuguese displayed medial scrambling, which consisted of the movement of various types of constituents (e.g., DPs, PPs, APs, and AdvPs) to multiple Specifier positions selected by the functional head IP. As I pointed out in Section 4.3, IP scrambling is among the contexts that generate RRC extraposition in earlier stages of Portuguese, as illustrated in (258)-(259), repeated here for ease of exposition:

- (258) que llj eu Alguna cousa diuía que nõ seia escripto
that to.him:CL I some thing owed that not be:PRES.SUBJ written
 en Estamãda
in this will

‘(And if there arrives someone who says) that I owed him something which is not written in this will ...’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1275)

- (259) E pera totalas cousas e cada hũa delas ffaser que uerdadeyro
and to all.the things and each one of.them make:INF that real
 e lĩdemo procurador pode e deue ffaser
and legitimate proxy can and should make:INF

‘And to make all the things and each one of them that a real and legitimate proxy can and should make.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1317)

Based on these assumptions, I tentatively claim that with the loss of IP scrambling after the 16th century, there was a decrease in the frequency of extraposition contexts in general. As a result, Portuguese began displaying short scrambling only, which consisted of the movement of noun phrases (either subjects or direct objects) to a VP-adjoined position (see Section 5.3.1). In this environment, the linear distance between the antecedent and the extraposed RRC decreased, and more importantly, PPs ceased to occur in a scrambled position.⁵⁶ Given the loss of an important trigger of the specifying coordination analysis, children converged on a new grammar. As schematically represented in (257), in the new grammar, the visible antecedent is

⁵⁶ Costa (2004a) proposes that PPs cannot undergo short scrambling in CEP. I will assume that this is correct, though I have no explanation for why the scrambling of PPs ceased to be an option in CEP.

analyzed as being generated in an RRC-internal position, and the RRC extraposition is analyzed as involving the rising of the antecedent (dispensing with the coordinate part of the structure).

The hypothesis that RRC extraposition ceased to be generated by coordination can be independently confirmed by another change that apparently took place within the history of Portuguese involving the loss of extraposition of conjuncts. As in Dutch (see (260), repeated from (226)a above), earlier stages of Portuguese allowed for extraposition of the second conjunct of a coordinate phrase, as illustrated in (262)-(267):

- (260) Zij heeft Marie gezien en mij
she has M. seen and me
 ‘She saw Mary and me.’

(Koster 2000: 16)

- (261) e sse os #ijj anos o dyto canpo nõ chantardes ou a
and if the three years the mentioned land not plant:FUT.SUBJ.2PL or the
dyta ujnha
mentined vineyard

‘and if you do not plant the aforementioned land or the aforementioned vineyard during the three years...’

(Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1294)

- (262) E por séér mays firme esta carta seelamos dos nossos séélos e
and by be:INF more firm this letter stamp:1PL of.the our stamps and
outra tal.
other such

‘And, to be irrevocable, we stamp this letter and a duplicate of it.’

(Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1273)

- (263) que he setuada na egreia de sã Johane da praça ã que
that is located in.the church of S. J. d. P. in that
o dicto diego afomso Jaz êterrado e seus filhos
the mentioned D. A. lies buried and his sons

‘that is located in the church of S. Johane da Praça, in which the aforementioned Diego Afomso and his sons lie buried.’

(A. Costa 2004a: 415, year 1425)

- (264) E estes bẽes E quinhom acõteço aa dita lionor
and these belongings and part happened to mentioned L.
uaasquez madrre da dicta viollante em seu drecto e meatade
V. mother of.the mentioned V. in her right and half
 da dita terça.
of.the mentioned third
 ‘and these belongings and part and half of the aforementioned third part went to the
 aforementioned Lionor Vaasquez, mother of the aforementioned Viollante, by her
 own right.’
 (A. Costa 2004b, year 1425)
- (265) E quando el vio Lançarot ir e a donzella
and when he saw L. go:INF and the damsel
 ‘And when he saw Lançarot and the damsel coming...’
 (Piel and Nunes 1988; *Demanda do Santo Graal*; 15th century-manuscript)
- (266) Tamanho o ódio foi e a má vontade
such the hate was and the bad will
 ‘Such was the hate and the malice’
 (Ramos 1982; L. de Camões, *Os Lusíadas*, Book I, 71, 16th century)
- (267) Contra hũa dama, ó peitos carniceiros,
against a dame oh chests ±bloodthirsty
 Feros vos amostrais e cavalleiros?
ferocious you:CL show:2PL and knights
 ‘Out on ye, ruffians; swords against a dame!
 Oh, unexampled butchers, knighthood’s shame!’
 (L. de Camões, *Os Lusíadas*, Book III, 130, 16th century, cited in Dias 1933/1970:
 328)

Indeed, the same is true of Latin, as illustrated in (268)-(273):

- (268) quae frigus defendant et solem
which cold keep.off and sun
 ‘To keep off the cold and the sun.’
 (Cato 48.2, cited in Devine and Stephens 2006: 586)
- (269) Postea salem addito et cumini paululum
next salt add:IMP and cumin:GEN bit
 ‘Next add salt and a bit of cumin’
 (Cato 156.5, cited in Devine and Stephens 2006: 586)

- (270) Deos placatos pietas efficiet et sanctitas
gods appeased devotion make:FUT and holiness
 ‘Religious devotion and holiness will appease the gods.’
 (Cicero, *De Officiis* 2.11, cited in Devine and Stephens 2006: 587)
- (271) Aqua restabat et terra
water remained and earth
 ‘There remained water and earth.’
 (Cicero, *De Natura Deorum* 2.66, cited in Devine and Stephens 2006: 589)
- (272) socius populi Romani atque amicus
ally people:GEN Roman:GEN and friend
 ‘Ally and friend of the Roman people’
 (Cicero, *Pro Lege Manilia* 12, cited in Devine and Stephens 2006: 589)
- (273) frigido loco et sicco
cool place and dry
 ‘In a cool, dry place’
 (Columella, *De Re Rustica* 12.47.1, cited in Devin and Stephen 2006: 422)

In contrast, extraposition of the conjunction and the second conjunct of a coordinate phrase is completely excluded in CEP, as illustrated in (274)b and (275)b:

- (274) a. O Pedro e a Maria chegaram.
the P. and the M. arrived
 ‘Pedro and Maria arrived.’
 b. *O Pedro chegou e a Maria.
the P. arrived and the M.
- (Colaço 2006: 79)
- (275) a. O medo e a ansiedade espalharam-se.
the fear and the anxiety spread
 ‘The fear and the anxiety spread.’
 b. *O medo espalhou-se e a ansiedade.
the fear spread and the anxiety
- (276) a. Eu vi o João e a Maria no cinema.
I saw the J. and the M. at.the cinema
 ‘I saw João and Maria at the cinema.’
 b. *Eu vi o João no cinema e a Maria.
I saw the J. at.the cinema and the M.

Depending on the analysis of conjunct extraposition adopted, one may advance two different explanations for the contrast between CEP and the earlier stages of Portuguese described above. If we assume that conjunct extraposition involves specifying coordination, the schematic representation in (277) is derived. In this case, two distinct syntactic configurations of coordination are involved: the specifying coordination generates extraposition, and regular coordination conjoins *Jaap* to *Joop* within the second conjunct.

- (277) Ik heb [_{CoP} [_{AgrOP1} [_{DP} Joop] gezien]]
 Co [_{AgrOP2} [_{CoP} [_{DP1} ~~Joop~~] Co [_{DP2} Jaap]]] ~~gezien~~
 ‘I have Joop seen and Jaap.’

However, as De Vries (2002: 279) himself notes, a simpler structure is also available, involving only one structure of coordination (the regular coordination), as sketched in (278). In this case, conjunct extraposition is derived not from specifying coordination but rather from the possibility of attaching the second conjunct (of regular coordination) to a larger constituent containing the first conjunct:

- (278) Ik heb [_{CoP} [_{AgrOP1} [_{DP} Joop] gezien]]
 Co [_{AgrOP2} [_{DP} Jaap] ~~gezien~~]

Under the first hypothesis sketched in (277), the contrast between CEP and earlier stages of Portuguese with respect to conjunct extraposition can be explained by the loss of the specifying coordination (plus ellipsis) structure.⁵⁷ More precisely, if we assume that variation is limited to the lexicon (see Chomsky 1995) and that what is taken as representing a syntactic change should instead be taken as representing a change in the lexical items, then the change observed in the history of Portuguese may be explained by the fact that the abstract restrictive specifying coordinator that is involved in extraposition in earlier stages of Portuguese is no longer available in the lexicon.

⁵⁷ Note that this hypothesis provides further evidence for the analysis of conjunct extraposition and RRC extraposition in terms of specifying coordination. The fact that both structures undergo diachronic change seems to suggest that they are both derived in the same way (i.e., through specifying coordination).

Alternatively, under the hypothesis sketched in (278), the different possibilities of conjunct extraposition may be derived from the asymmetries between the ellipsis types available in different stages of Portuguese. Because regular coordination can still be applied at different levels of the syntactic structure in CEP, one may conjecture that earlier stages of Portuguese and CEP differ with respect to the types and properties of ellipsis they allow. Indeed, whereas earlier stages of Portuguese allow for a broad range of ellipsis types (including the ones that derive conjunct extraposition), CEP is severely constrained in them.⁵⁸

In general, we may conclude that extraposition in Portuguese is subject to different constraints across time. Empirical evidence for this claim comes from conjunct and RRC extraposition in the diachrony of Portuguese. Focusing in particular on RRC extraposition, I have argued that RRC extraposition ceased to involve the specifying coordination plus ellipsis analysis and began to be generated by stranding. This change was due to the decrease in the displacement operations caused by the loss of IP scrambling, which led to a decrease in utterances manifesting the specifying coordination plus ellipsis structure. Because the available triggering experiences changed in a critical way, children converged on a new grammar, beginning to generate RRC extraposition by stranding.⁵⁹

7.2. Cross-linguistic contrasts

Throughout this chapter, cross-linguistic evidence was presented that points to a non-unitary approach to RRC extraposition. Special attention to this topic was given in Section 4.2, where it was demonstrated that RRC extraposition in various languages has different properties. Although the overview offered in Section 4.2 has several limitations in terms of cross-linguistic coverage, it provides sufficient

⁵⁸ Further research on ellipsis phenomena in earlier stages of Portuguese is, however, necessary to warrant the validity of this hypothesis.

⁵⁹ Another possibility that is worth exploring in future research is that both the specifying coordination and the stranding structure could generate RRC extraposition in earlier stages of Portuguese. The stranding would be available for those cases where it leads to the same overt results as the specifying coordination, whereas the specifying coordination would take over in the ‘difficult’ cases (*i.e.*, in the cases that cannot be derived from stranding). Under this approach, the explanation for the change given above is still maintained: the constrained nature of RRC extraposition in CEP would be explained by the fact that RRC extraposition started to be exclusively generated by stranding (because the specifying coordination structure or the specifying coordinator ceased to be an option).

evidence to conclude that CEP contrasts with other languages (e.g., English and Dutch) with respect to the properties of RRC extraposition. The main findings of Section 4.2 are summarized in Table 10, repeated here for ease of exposition.

Table 10. A comparative approach to RRC extraposition (CEP vs. English and Dutch)

	CEP	English	Dutch
A. <i>Extraposition from strong noun phrases</i>	-	+	+
B. <i>Extraposition from embedded positions</i>	-	+	+
	subjects	+	+
	<i>wh</i> -constituents	+	+
C. <i>Extraposition from pre-verbal constituents</i>	emphatic/evaluative phrases	+	+
	preposed foci	+	+
	topics	-	-

To account for these contrasts, I tentatively suggest that there are two different types of structure that can generate RRC extraposition: the specifying coordination plus ellipsis structure and the stranding structure. These structures are not instantiated in all languages because it is possible to identify at least two types of languages:

- Type I. Languages that generate RRC extraposition by stranding, e.g., CEP (and possibly Italian, Spanish and French).
- Type II. Languages that generate RRC extraposition by specifying coordination plus ellipsis, e.g., English and Dutch.

The dual approach proposed here provides useful insight to understand the synchronic and diachronic variation reported throughout this chapter. In particular, it explains why RRC extraposition in earlier stages of Portuguese exhibits the same syntactic properties as RRC extraposition in Dutch. I submit that this is due to the fact that both are generated by specifying coordination (plus ellipsis).

Another interesting conclusion drawn in Section 4.2 is that Romance languages do not behave in a uniform way with respect to RRC extraposition. The main findings of this comparison are summarized in Table 11, repeated here for ease of exposition.

Table 11. A comparative approach to RRC extraposition in Romance languages

	CEP	Italian	Spanish	French
A. <i>Extraposition from strong noun phrases</i>	-	-	-	-
B. <i>Extraposition from pre-verbal subjects</i>	-	-	-	+

French exhibits a peculiar behavior: it contrasts with other Romance languages in allowing extraposition from a pre-verbal position, but it also contrasts with some Germanic languages (like English and Dutch) in not allowing extraposition from strong noun phrases.

As mentioned above in the previous section, I believe that these facts can be accounted for if we assume that languages may differ in the way they resolve the ambiguity of a constituent in [Spec, IP], expressed in (279):

(279) *Interpretative Principle*

The antecedent of an extraposed RRC must occur in a position non-ambiguously interpreted as non-topic (in Kuroda's 2005 sense).

Whereas in CEP (and possibly in Spanish and Italian), the ambiguity associated with [Spec, IP] is resolved syntactically and prosodically (through subject inversion), in French and in earlier stages of Portuguese, it may be resolved only prosodically. In this case, a constituent in [Spec, IP] can be unambiguously interpreted as non-topic if it is prosodically marked by pitch accent. Ultimately, this amounts to saying that the cross-linguistic variation in RRC extraposition from pre-verbal subjects is determined by how the different languages mark the topic/non-topic status of the subject.

Of course, further comparison between languages in this domain is necessary to understand if these hypotheses are correct.

8. Conclusion

As I announced in the introduction, the main goal of this chapter was to contribute to a better understanding of the syntax of RRC extraposition. This was achieved by discussing new empirical evidence from CEP and earlier stages of Portuguese, which was systematically compared to data from other languages (with special emphasis on English and Dutch).

From a descriptive point of view, three contrasting properties of RRC extraposition were identified:

- The definiteness effect.
- Extraposition from embedded positions.
- Extraposition from pre-verbal positions.

I provided empirical evidence suggesting that earlier stages of the same language and different languages may manifest divergence with respect to the properties of RRC extraposition that they display. More precisely, I have shown that:

- Earlier stages of Portuguese contrast sharply with CEP with respect to the properties of RRC extraposition.
- RRC extraposition in earlier stages of Portuguese is, to a large extent, Germanic-like, unlike CEP.

Exploring the theoretical impact of these findings, I submitted that the variation found in the syntax of RRC extraposition is not compatible with a uniform approach to the phenomenon. Therefore, I argued for a dual approach to RRC extraposition, whereby RRC extraposition may involve two different structures:

- Specifying coordination plus ellipsis (De Vries 2002).
- VP-internal stranding (Kayne 1994).

In the diachronic dimension, I proposed that:

- In earlier stages of Portuguese, RRC extraposition is generated by specifying coordination plus ellipsis.
- In CEP, RRC extraposition is generated by VP-internal stranding.

In the cross-linguistic dimension, I suggested that there are at least two types of languages:

- Type I. Languages that generate RRC extraposition by stranding, e.g., CEP (and possibly Italian, Spanish and French).

- Type II. Languages that generate RRC extraposition by specifying coordination plus ellipsis, e.g., English and Dutch.

Note that the present analysis, though non-unitary, is restrictive because the theoretical apparatus used to explain the variation in RRC extraposition is limited to structures and operations that are independently available in the grammar (i.e., coordination, movement, and stranding). Taking this view, RRC extraposition is not conceived as a construction type but rather as the result of structures operating in a theoretically legitimate way. Therefore, the prediction is that one or another of these structures may manifest alone in particular languages, being able to derive superficially identical structures.

Ultimately, this approach amounts to stating that the different analyses proposed in the literature for RRC extraposition may be seen not as competing analyses for a single construction but rather as complementary analyses for two distinct structures. The approach adopted here suggests that the concept of extraposition may be descriptively useful (in unifying a variety of apparently related constructions) but lacks explanatory force because it does not unequivocally correspond to a construction type.

Chapter 4

Variation and Change in the Syntax of Appositive Relative Clauses

Contents

1. Introduction.....	229
1.1. Goals	229
1.2. Chapter outline.....	230
2. The syntax of ARCs.....	231
2.1. The competing analyses.....	231
2.1.1. Orphanage analyses	231
2.1.2. Constituency analyses.....	233
2.2. Unitary vs. non-unitary analyses.....	235
2.2.1. Cinque (1982)	236
2.2.2. Smits (1988).....	240
2.2.3. Bianchi (1999)	241
2.2.4. Cinque (2008)	243
3. Towards a dual approach to ARCs: evidence from Portuguese	246
3.1. <i>O qual</i> -ARC: an overview of its basic use and properties.....	247
3.2. Two types of <i>o qual</i> -ARCs: diachronic evidence.....	254
3.2.1. Some properties in contrast.....	255
3.2.1.1 Internal head.....	255
3.2.1.2 Extraposition	267
3.2.1.3 Pied-piping.....	286
3.2.2. Additional properties in contrast.....	303
3.2.2.1 Clausal antecedent	303
3.2.2.2 Split antecedents	307
3.2.2.3 Coordination of the <i>wh</i> -pronoun with another DP.....	310
3.2.2.4 Illocutionary force.....	312
3.2.2.5 Coordinator	315

3.3. Summary	317
4. An analysis of the two types of ARCs	319
4.1. <i>o qual</i> -ARCs in earlier stages of Portuguese	319
4.1.1. The specifying coordination analysis.....	320
4.1.2. The expansion of the specifying coordination analysis	325
4.1.3. Summary	334
4.2. <i>O qual</i> -ARC in CEP.....	335
4.2.1. The raising analysis.....	335
5. Deriving the contrasting properties of the two types of <i>o qual</i> -ARCs	338
5.1. Internal head.....	338
5.2. Extraposition	340
5.3. Pied-piping.....	353
5.4. Clausal antecedent	361
5.5. Split antecedents	361
5.6. Coordination of the <i>wh</i> -pronoun with another DP.....	364
5.7. Illocutionary force.....	365
5.8. Coordinator	369
5.9. Summary	369
6. Some comparative remarks.....	370
6.1. Diachronic change in the syntax of <i>o qual</i> -ARCs	370
6.2. Synchronic variation in the syntax of <i>o qual</i> -ARCs	377
6.3. Cross-linguistic contrasts	383
7. Conclusion	385

1. Introduction

In the literature on the syntax of appositive relative clauses (ARCs), considerable attention has been given to the idea that ARCs do not constitute a unified type of construction (e.g., Cinque 1982, 2008, Smits 1988). These analyses contrast with the traditional view, according to which the different ARCs found in all languages can be derived from the same syntactic structure.

In this chapter, I will go against the traditional view and claim that there is no unified account of ARCs across languages. This claim is supported by the study of a syntactic change that took place within the history of Portuguese involving ARCs introduced by the complex relative pronoun *o qual* lit. ‘the which’ (henceforth *o qual*-ARCs). The investigation of this micro-variation allows us to conclude that the syntactic properties of *o qual*-ARCs have changed significantly over time and that this fact can only be explained by a non-unified approach to the phenomenon.

That this conclusion can be reached in the diachronic dimension is particularly telling given the highly constrained nature of this variation. Given that the variation under analysis is found within the same language (Portuguese), with the same ‘syntactic construction’ (ARC), introduced by the same relativizer (*o qual*), it is possible to control important variables that may interfere with the results obtained in other studies (involving, for instance, the comparison of languages quite distant historically and typologically).

1.1. Goals

The major goal of this chapter is to contribute to a better understanding of the syntax of ARCs, bringing into discussion new empirical evidence from earlier stages of Portuguese and Contemporary European Portuguese (CEP).

From a descriptive point of view, the specific goals of this chapter are as follows:

- To establish clear syntactic properties to distinguish *o qual*-ARCs in CEP from *o qual*-ARCs in earlier stages of Portuguese.

- To correlate the variation documented in the diachronic dimension with the one found in the cross-linguistic dimension.

From a theoretical point of view, the specific goals of this chapter are as follows:

- To demonstrate that ARCs do not constitute a unitary phenomenon.
- To show that *o qual*-ARCs in CEP differ in structure and derivation from the *o qual*-ARCs in earlier stages of Portuguese.
- To argue for a dual approach to *o qual*-ARCs, according to which *o qual*-ARCs in CEP involve the head raising analysis (Kayne 1994, Bianchi 1999) and *o qual*-ARCs in earlier stages of Portuguese involve the specifying coordination analysis (De Vries 2006b).
- To demonstrate that the dual approach to ARCs provides a good basis for understanding the variation found within a language and across languages, both in the synchronic and diachronic dimensions.

1.2. Chapter outline

This chapter is organized as follows. Section 2 provides background, giving an overview of existing analyses of ARCs. In this section, two main lines of research are addressed: unitary approaches and non-unitary approaches to ARCs.

Section 3 introduces evidence for a dual approach to ARCs based on diachronic evidence from Portuguese. As an introduction to the topic, Section 3.1 gives an overview of the properties of *o qual*-ARCs in CEP and in earlier stages of Portuguese. Section 3.2 sets out the syntactic properties that differentiate the two constructions.

To derive these properties, Section 4 outlines two different syntactic analyses: the head raising analysis (Kayne 1994, Bianchi 1999) and the specifying coordination analysis (De Vries 2006b). Section 5 then shows how these two analyses explain the contrasting properties of *o qual*-ARCs in CEP and in earlier stages of Portuguese.

Finally, Section 6 offers some comparative remarks. Section 6.1 shows how the dual approach to *o qual*-ARCs explains the change that took place in the history of Portuguese. Sections 6.2 and 6.3 show how it can explain synchronic variation within the same language and across languages. The conclusions are presented in Section 7.

2. The syntax of ARCs

2.1. The competing analyses

Syntactic analyses of ARCs can be grouped together under the following bipartite classification¹ (see De Vries 2006b and Arnold 2007, for an overview):

- *Orphanage analyses*
 - Radical orphanage analyses (e.g., Safir 1986, Fabb 1990, Espinal 1991).
 - Non-radical orphanage analyses (e.g., Ross 1967, Emonds 1979, Demirdache 1991).
- *Constituency analyses*
 - Adjunction analyses (e.g., Jackendoff 1977, Perzanowski 1980).
 - Head raising analyses (e.g., Vergnaud 1974, Kayne 1994, Bianchi 1999).
 - Coordination analyses (e.g., Koster 1995, 2000, De Vries 2002, 2006b).

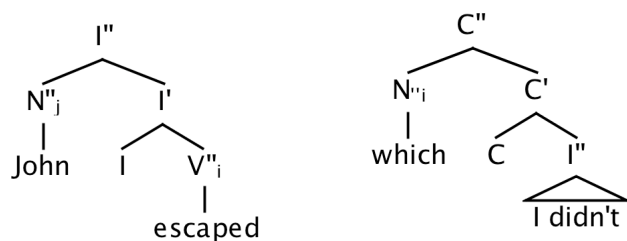
2.1.1. Orphanage analyses

The central plank of the orphanage analyses is that the ARC and the antecedent are generated separately. Two variants of this approach can be identified: radical orphanage analyses and non-radical orphanage analyses.

Radical orphanage analyses propose that there is no syntactic link between the ARC and the sentence containing the antecedent at any level of syntactic representation. For instance, Fabb (1990), one of the proponents of this approach, claims that the ARC does not enter any syntactic relation with the matrix (such as modification, specification, or theta-assignment, etc.). This fact is illustrated in (1), taken from Fabb (1990: 61).

¹ This classification corresponds to what Emonds (1979) calls the Main Clause Hypothesis and the Subordinate Clause Hypothesis.

(1)



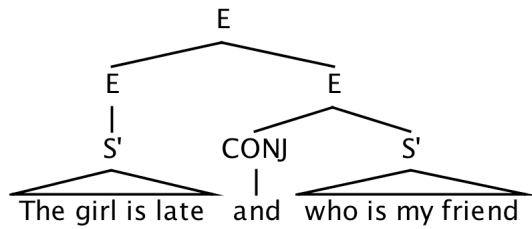
As can be observed in (1), there is no syntactic link between the relative CP and the sentence *John escaped*. Other than some pragmatic notion of *aboutness*, the only relation established is between the antecedent and the relative pronoun. According to Fabb, this relationship involves the sharing of the same referential index (see index $_i$ in (1)), a condition that is satisfied at the level of discourse structure rather than in the syntax. The adjacency between the ARC and the antecedent is then derived only at the discourse level.

In contrast, non-radical orphanage analyses propose that the antecedent and the ARCs are generated separately in the syntax; the ARC, however, is part of the syntactic structure of the matrix clause at some syntactic level. Emonds (1979), one of the proponents of this approach, suggests that ARCs are derived from underlying conjoined clauses. The adjacency between the antecedent and the ARC is derived from extraposition (interpreted as rightward movement) of the intervening material. The derivation of a sentence such as *The girl, who is my friend, is late* can then be represented as follows:²

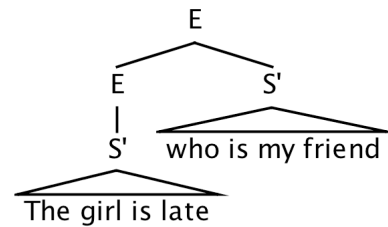
² For ease of exposition, I present the representation of Emonds analysis given in Demirdache (1991: 104). In this representation, the symbol E(xpression) stands for the highest category in a sentence, which cannot be subordinated.

(2)

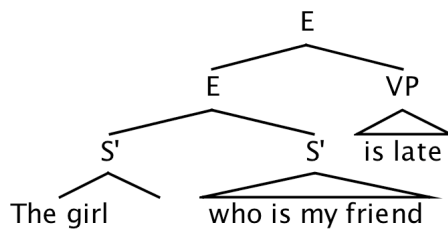
a.



b.



c.

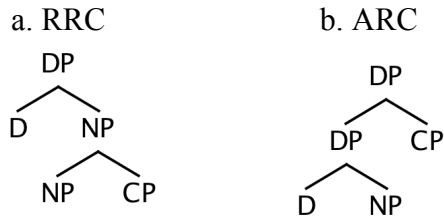


In this analysis, the derivation of an ARC can be summarized as follows: (i) at D-Structure, two main clauses are conjoined ((2)a); (ii) then, the conjunction *and* is deleted, and the relative is directly attached to E ((2)b); (iii) finally, the constituent that intervenes between the antecedent and the ARC (a VP in (2)) is moved to the right, right-adjointing to the main clause ((2)c).

2.1.2. Constituency analyses

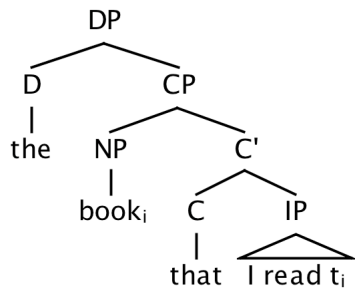
At the opposite extreme, constituency analyses claim that the antecedent and the ARC form a constituent. In this approach, the standard analysis is that the ARC is adjoined to the antecedent. The development of the *adjunction analysis* of relative clauses goes back to Ross (1967), Chomsky (1977) and Jackendoff (1977). This analysis entails a similar derivation for restrictive relative clauses (RRCs) and ARCs in the sense that both are subordinate clauses adjoined to the antecedent. The difference concerns the level of attachment: in present-day syntax RRCs may be viewed as adjoined to the NP-level, whereas ARCs (with nominal antecedents) are attached to the DP-level (see, e.g., Demirdache 1991: 109). The contrast is sketched out in (3):

(3)



Another structure that qualifies as constituency analysis is the so-called *head raising analysis* (henceforth *raising analysis*). As mentioned in Chapter 1, the raising analysis of relative clauses was originally proposed by Schachter (1973) and Vergnaud (1974, 1985). It has been revived and moderated by Kayne (1994) and Bianchi (1999) within a general antisymmetric framework of syntax. The main idea is that the head NP (the antecedent) of a relative clause originates at the relativization site inside the subordinate clause and then rises to the left edge. The relative clause itself is generated as the complement of the so-called external determiner, with which the head NP may associate after raising. See the (simplified) representation in (4):

(4)

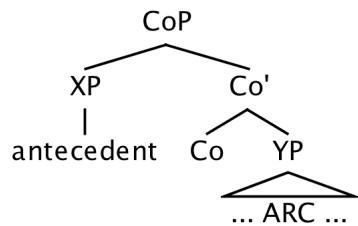


Kayne (1994) and Bianchi (1999) propose that both RRCs and ARCs are derived via the head-raising structure given in (4). To account for the scope-related contrasts between both constructions, they hypothesize that ARCs involve covert remnant movement (at LF) of the relative IP to the Specifier position of the external determiner D, where it is no longer in the scope of either D or the head NP (see (5)).

(5) [_{DP} [_{IP} I read t_{DP}]] [_{D'} this [_{CP} book which t_{IP}]]] (LF)

Another type of analysis that can be included under the umbrella of the constituency analyses is the coordination analysis of ARCs. This approach is proposed by Koster (1995, 2000) and De Vries (2002, 2006b) with different details of implementation. The basic assumption is that the ARC is coordinated to the antecedent, as shown in (6):

(6)



The scheme in (6) is implemented in different ways: for Koster (1995, 2000), YP=CP; consequently, ARCs usually involve unbalanced coordination (because XP can correspond to DP, AP, PP, VP, IP); for De Vries (2002, 2006b), YP=DP and, consequently, ARCs with nominal antecedents always involve balanced coordination. Interestingly, De Vries' analysis of ARCs also involves raising, as an abstract D in the second conjunct takes the relative clause as complement; this corresponds to a raising-style configuration of a full relative construction (a DP containing a relative clause), as shown in (7):

(7) [_{CoP} DP Co [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} t_k]]]]
 e.g. *Jack* , \emptyset \emptyset *who* \emptyset *lives in Paris*

The interesting aspect of this analysis is that an ARC is in fact an RRC in apposition to the overt antecedent. I will come back to this analysis in more detail in Section 4.1.1.

2.2. Unitary vs. non-unitary analyses

ARCs are traditionally regarded as a unitary type of construction. Under this view, the properties of ARCs found in all languages can be derived the same syntactic structure. There are, however, early indications in the literature that one universal analysis of ARC across and within languages is untenable (cf. Cinque 1982, Smits 1988, Bianchi

1999, and, more recently, Cinque 2008). In the next sections, I will present the most relevant aspects of these approaches, focusing on the evidence provided by each author for a non-unitary approach to ARCs.

2.2.1. Cinque (1982)

Cinque (1982), when considering the properties of ARCs in Italian, clearly states that a single syntactic structure cannot account for the properties of ARCs found in this language. In particular, he is concerned with the contrast between ARCs introduced by *che/cui* and by *il quale*. According to Cinque, these two types of ARCs exhibit contrasting properties, and, consequently, must be derived from two different structures.

Among the syntactic properties that differentiate the two types of ARCs, Cinque (1982) highlights the contrasts regarding the restrictions on the relativized positions and the restrictions on the availability of pied-piping.

As for the restrictions on the relativized position, Cinque shows that when either a subject or an object is relativized, a *wh*-pronoun appears in *il quale*-ARCs but not in *che-cui*-ARCs. In the latter, only the form *che* is found, which is identical to the ordinary complementizer of subordinate clauses. Examples (8)-(9) illustrate this restriction: the *a*. examples involve relativization of the subject, whereas the *b*. examples involve relativization of the direct object.

(8) a. Giorgio, *che/*cui* ti vuole, è là.

‘Giorgio, who wants you, is there.’

b. Giorgio, *che/*cui* stimi, l’ha fatto.

‘Giorgio, who you esteem, has done it.’

(9) a. Giorgio, *il quale* ti vuole, è là.

b. Giorgio, ?*il quale* stimi, l’ha fatto.

(Cinque 1982: 248-249)

As for the restrictions on pied-piping, Cinque notes that in *che/cui*-ARCs, no pied-piping is allowed except for that of PPs, whereas in *il quale*-ARCs, pied-piping of different types of phrases is available. Examples (10)-(11) illustrate this contrast: *a*.

examples illustrate pied-piping of DP, *b.* examples illustrate pied-piping of complex PP, and *c.* examples illustrate pied-piping of an infinitival clause.

- (10) a. *Giorgio, la figlia di cui fuma, è contrario.
lit. 'Giorgio, the daughter of whom smokes, is against it.'
- b. *Giorgio, alla figlia di cui hai scritto, è in collera.
lit. 'Giorgio, to the daughter of whom you have written, is angry.'
- c. *Giorgio, fuggire da cui non osava, è morto.
lit. 'Giorgio, fly from whom he did not dare, has died.'
- (11) a. Giorgio, la figlia del qual fuma, è contrario.
b. Giorgio, alla figlia del quale hai scritto, è in collera.
c. Giorgio, fuggire dal quale non osava, è morto.

Based on these empirical contrasts, Cinque (1982) hypothesizes that in Italian, there are two separate paradigms of ARCs, one belonging to the 'core' grammar of Italian and other being peripheral to it. Under Cinque's proposal, the core grammar of Italian employs the structure $[_{NP} NP \bar{S}]$ for relativization. This structure derives both RRCs and ARCs introduced by the relativizers *che/cui*. In addition to this structure, a more peripheral option is available for *il quale*-ARCs, in which the relative clause is regarded as a juxtaposed clause with the structure $NP \dots, \bar{S}, \dots$. Note that, under the terms proposed in Section 2.1, these structures qualify, respectively, as a constituency analysis and an orphanage analysis.

Because the use of *il-quale* introducing ARCs is felt to be slightly more formal in style than the use of the relativizers *che/cui*, Cinque argues that stylistic markedness can be interpreted as a manifestation of the use of a more peripheral structure allowed by the grammar; that is, the structure: $NP \dots, \bar{S}, \dots$

In this view, the restrictions on the relativized positions and on the availability of pied-piping can be derived as clarified below.

A. Restriction on subject/object relativization

In Italian, when a subject or an object is relativized, a *wh*-pronoun surfaces in *il quale*-ARCs but not in *che-cui*-ARCs. According to Cinque (1982), the explanation

for this contrast rests upon the principle of “obligatory deletion of *wh*-phrases up to recoverability.” In simple terms, *wh*-phrases are obligatorily deleted in *che/cui*-ARCs, but not in *il quale*-ARCs; in the latter, the deletion of the *wh*-phrase is not allowed.

In more formal terms, a *wh*-phrase can be deleted (i) if it is non-distinct from the head and (ii) if it is c-commanded by the head. In *che/cui*-ARCs, the *wh*-phrase is non-distinct from the head and is c-commanded by the head (in the configuration [_{NP} NP \bar{S}]); therefore, the *wh*-phrase is deleted (and the complementizer is expanded to *che*).

In contrast, in *il quale*-ARCs, the *wh*-phrase is not c-commanded by the head (in the configuration NP ..., \bar{S} , ...). Recall that in this configuration, the relative clause is juxtaposed and, consequently, is syntactically invisible for c-command relations. This explains why the *wh*-phrase cannot be deleted.

B. Restriction on the availability of pied-piping

Cinque (1982) proposes an ambiguous lexical characterization of some *wh*-pronouns. The basic idea is that all relative pronouns belong to the inventory of lexical (bound) anaphors of language, but only a few can be further used as non-anaphoric elements. This is the case of *il quale* but, crucially, is not the case of *cui*.

Anaphoric pronouns can enter the structure [_{NP} NP \bar{S}] because in this configuration, they are c-commanded by the head (as required by the principle A of the Binding Theory). They cannot, however, enter the structure NP ... \bar{S} , ... because in this case, the c-command requirement is not satisfied.

Non-anaphoric pronouns behave differently in this respect. They can enter the structure NP ... \bar{S} , ... simply because they are not limited by principle A of the Binding Theory.

The restrictions on the availability of pied-piping follow from the (non)-anaphoric nature of *wh*-pronoun. Assuming that S and NP are the only governing categories, the restrictions found in *che/cui*-ARCs can be represented as in (12). The minimal governing categories of *che/cui* are emphasized for ease of exposition.

- (12) a. ... $\boxed{NP_i}$ $[S [COMP [PP P [NP_i wh] [-WH] S]]]$... [pied-piping of PP]
 b. ... $[NP NP_i [S [COMP \boxed{NP} \bar{N} [PP P [NP_i wh]]] -WH] S]$... [pied-piping of an NP]
 c. ... $[NP NP_i [S [COMP \boxed{S} \dots [NP_i wh]]] -WH] S]$... [pied-piping of a clause]
 (Cinque 1982: 255)

Pied-piping of PPs is allowed because the *wh*-anaphor is bound within its minimal governing category (the emphasized NP). In turn, pied-piping of noun phrases and clauses is not allowed because of the violation of principle A of the Binding Theory. Note that in (12)b and (12)c, the *wh*-anaphor is free in its minimal governing category (the emphasized NP in (12)b and the emphasized S in (12)c).

The unconstrained availability of pied-piping observed in *il quale*-ARCs is explained by the non-anaphoric nature of *il-quale* pronouns. Unlimited by principle A of the Binding Theory, *il-quale* pronouns can refer back to their antecedent independently of the category of the pied-piped constituent. The relation between non-anaphoric *il quale* pronouns and the antecedent can be analogized to the relation between a demonstrative pronoun and its antecedent, the nature of the relation being one of discourse grammar rather than one of sentence grammar.

In sum, Cinque (1982) shows that Italian ARCs involve two different ARCs: one with the structure $[NP NP \bar{S}]$ (introduced by the relativizers *che/cui*) and another with the structure $NP \dots, \bar{S}, \dots$ (introduced by the relativizer *il quale*). Adding a cross-linguistic dimension to the discussion, he claims that this non-unitary approach does not universally hold. For instance, whereas French and Italian display two constructions, English only displays the structure $NP \dots, \bar{S}, \dots$. In his view, this explains some of the differences between English ARCs and Italian/French ARCs, namely, that only *wh*-pronouns are allowed to introduce ARCs in English (*Mary, who/*that/ *Ø you met yesterday*) and the possibility of generalized pied-piping. Because Cinque elaborates on the cross-linguistic variation in his 2008 paper, more details on the topic will be postponed until Section 2.2.4.

2.2.2. Smits (1988)

Smits (1988) claims that a single syntactic analysis cannot account for the heterogeneous types of ARCs found within a language and across languages. Alternatively, two different analyses are proposed: a constituency analysis for ARCs (in this case, an adjunction analysis) and an orphanage analysis of ARCs in which the antecedent and the ARC are two completely independent parts of the sentence containing them (see (13)).

(13) ..._[XP antecedent] ... _[RC rel. ... R ...]...

Two arguments are provided for the existence of the structure in (13). One is that ARCs may have split antecedents,³ as illustrated in (14) (from Demirdache 1991: 166).

(14) *A man* entered the room and *a woman* went out, who were quite similar.

A constituency analysis is untenable because it would require the derivation of (14) to start with ARCs adjoined to each noun phrase, with one of the ARCs deleted later. The problem is that such an analysis would not be able to explain the presence of a plural verb form (*were*) found in the alleged visible ARC. An orphanage analysis, in contrast, can successfully derive ARCs with split antecedents. In this case, there is no direct structural link between the antecedent and the ARCs; therefore, nothing prevents ARCs from taking non-conjoined noun phrases as antecedent.

A second argument in favor of an orphanage type of ARCs concerns the existence of *pseudo-relatives* and *apparent extraposed ARCs*. These types are illustrated, respectively, in (15), from French, and (16), from Dutch.

(15) Marie est là, qui pleure comme une Madeleine.
M. is there who cries like a M.

‘Marie is there, and she is crying her heart out.’

(Smits 1988: 181)

³ The term *split antecedent* is used for an antecedent that consists of more than one non-conjoined noun phrases as in *John suggested to Mary that they should leave*; here, *they* takes as an antecedent the split antecedent *John* and *Mary*.

(16) Ik wilde mijn zuster opzoeken, die echter niet thuis was.
I want my sister visit who however not at.home was

‘We wanted to visit my sister, who wasn’t at home, however.’

(Smits 1988: 185)

Pseudo-relative clauses look like extraposed relatives; however, they are interpretationally different because they express an event in progress (whereby the apparent antecedent – *Marie* in (15) – participates). Apparent extraposed ARCs also look like extraposed relatives, but they have a specific type of meaning (such as continuative, resultative, or contrastive). Therefore, instead of referring to a noun phrase, they modify the whole state of affairs that is expressed in the preceding clause. For instance, the ARC in (16) expresses the result of the action described in the main clause with the meaning “but the action described in the predicate was in vain” (see Smits 1988: 186).

Given the interpretation associated with pseudo-relative and apparently extraposed ARCs, Smits considers that these constructions are closer to adverbial clauses than to true relative clauses, and, consequently, should be analyzed as involving the orphanage structure given in (13). In this type of structure, the ARC is not limited by locality constraints and may freely occur in the rightmost position of the sentence.

In sum, Smits (1988) argues for the existence of two different types of ARCs: one derived from a constituency structure and another derived from an orphanage structure. Although the constituency structure appears to be the most frequent type found across languages, there is also empirical evidence in favor of the orphanage structure. Two types of evidence are discussed: (i) ARCs with split antecedents and (ii) apparently extraposed ARCs and pseudo-relatives.

2.2.3. Bianchi (1999)

The limitations of unitary approaches are also addressed in Bianchi (1999: 151ff), at least to some extent, by suggesting that not all types of ARCs can be derived from the same syntactic structure (in this case, the raising analysis).

One serious problem Bianchi faces concerns the analysis of ARCs with non-nominal antecedents, as in (17):

- (17) a. Mary is courageous, which I will never be.
 b. John is in the garden, which is where I should be.
 c. Mary has resigned, which John hasn't.
 d. John was late, which was unfortunate.

(Bianchi 1999: 151)

The raising analysis cannot derive such examples in any plausible way. Essentially, the antecedent has to be selected by the relative determiner D_{rel} within the relative clause; thus, it has to be a nominal projection.

Another property Bianchi identifies as problematic for the raising analysis of ARCs concerns the so-called *relatif de liaison* or *connecting relative*. These terms refer to relative pronouns that apparently introduce a main clause. Such an impression is given by the fact that they are separated from the antecedent by a full stop or other heavy punctuation (e.g., a colon or a semicolon). Example (18), from Latin, is a case in point.

- (18) id oppidum Lentulus Spinther X cohortibus tenebat; qui Caesaris
this town L. S. with.ten cohorts held who of.Caesar
 adventu cognito profugit ex oppido
arrival known fled from town

'Lentulus Spinther held this town with ten cohorts; who, when he was informed of the arrival of Caesar, left the town.'

(Caesar, *De Bello Civili*, I, 15, cited in Ramat 2005)

According to Bianchi (1999: 152), this construction cannot be derived from the raising analysis because the head would be separated from the relative clause by a sentence boundary.

Even while recognizing these problems, Bianchi (1999) holds to a unitary approach to ARCs. To solve this paradox, she argues that sentences such as (17) and (18) may not be relative constructions at all; the relative pronoun can be taken as an anaphoric pronoun and the purported ARC may be either coordinated to the main clause or parenthetical.

2.2.4. Cinque (2008)

Cinque (2008) again takes up the non-unitary approach put forth in his 1982 paper, adding more empirical evidence and proposing a theoretical apparatus that reflects the recent developments in syntactic theory.

Assuming a cross-linguistic perspective, Cinque (2008) claims that there are two different types of ARCs: the *integrated* and the *non-integrated* constructions (which roughly correspond to the constituency and orphanage constructions proposed in this chapter). Some languages display both constructions (e.g., Italian and French), while others display only one. In the latter case, two options are available: some languages display only the integrated construction (e.g., Northern Italian dialects and, possibly, Chinese), whereas others display exclusively the non-integrated construction (e.g., English and Romanian).⁴

In Italian, *il quale*-ARCs belong to the non-integrated type, whereas *che/cui*-ARCs belong to the integrated type.⁵ Evidence for this distinction comes from a number of syntactic properties that differentiate *che/cui*-ARCs from *il quale*-ARCs.⁶ These properties are summarized in Table 1. Here a plus ‘+’ means possible and a minus ‘-’ means impossible.⁷

⁴ Note that other languages apparently lack ARCs. As Cinque (2007) notes, these languages have to resort to coordination (like Gungbe and Bunun) or to the apposition of generic nouns (such as ‘person’) followed by a restrictive clause (as in Mixtecan).

⁵ In Cinque (1982) and Cinque (2008), it is assumed that there are two different types of ARCs. However, Cinque (2008) does not retain the idea (suggested in Cinque 1982) that the two constructions have different ‘statuses’ in the grammar of Italian, one belonging to the core grammar and the other being peripheral to it. In Cinque (2008) it is simply assumed that these constructions may coexist in the same language.

⁶ Some of these properties were discussed in Cinque (1982), but in less detail.

⁷ There is only one exception. With respect to the property *Subjects and objects represented by a wh-pronoun*, a minus means ‘no’ and a plus means ‘yes’. As for the remaining properties, the contrast between + possible/ - impossible holds.

Table 1. Properties distinguishing *che/cui*-ARCs from *il quale*-ARCs

	<i>che/cui</i> -ARCs	<i>il quale</i> -ARCs
Subjects and objects represented by a <i>wh</i> -pronoun	-	+
Generalized pied-piping	-	+
Non-declarative illocutionary force	-	+
Non-adjacency	-	+
Split antecedents	-	+
Retention of the 'internal' head	-	+
Non-identity of the 'external' and 'internal' heads	-	+
Non-nominal antecedents	-	+
Relative clause preposing	-	+
Parasitic gaps	+	-
Temporal DPs as antecedent	+	-
<i>wh</i> -pronoun coordinated with another DP	-	+

From a cross-linguistic perspective, Cinque (2008) shows that the behavior of ARCs with respect to the properties displayed in Table 1 depends on the type of ARCs available in each language. Consider, for instance, the case of English, a language that only displays the non-integrated type. ARCs in English pattern with Italian *il-quale*-ARCs in that they:

- Obligatory retain *wh*-pronouns when the subject or the object is relativized (*Mary, who/*that/ *Ø you met yesterday*).
- Allow pied-piping of phrases other than PPs (*Mary, to hire whom would be a great opportunity, ...*).
- May have non-declarative illocutionary force (*Your father, by whom will we ever be forgiven for what we have done?, would have never behaved like that*).
- May have an additional internal head (*John was almost at the end of his financial resources, which fact led him to look for a cheaper house*).

Conversely, in Northern Italian dialects, which only display the integrated type, ARCs pattern with Italian *che/cui*-ARCs; as a result, they behave like Italian *che/cui*-ARCs as far as the properties in Table 1 are concerned.

As for the syntactic analyses of these two types of ARCs, Cinque (2008) proposes that integrated ARCs involve a syntactic structure similar to RRCs. Following the ideas put forth in his recent work (2003, 2009), he claims that relative clauses are generally merged in a pre-nominal position and that the post-nominal

order found in languages such as English and Portuguese is ultimately derived from the leftward movement of the head past the relative clause.⁸

The syntactic derivation that Cinque (2008) proposes for integrated ARCs is given in (19). Note that, first, the relative IP is merged above the determiner/demonstrative;⁹ this yields the order in (19)a, with the relative clause in a pre-nominal position. The procedure for deriving the post-nominal order is complex, as it involves a *matching* and a *raising* variant. The basic idea is that after IP raising to a higher licensing position, the complementizer $C_{(1)}$ is merged and attracts the *wh*-pronoun/internal head. Then, there are two possible ways for the derivation to proceed. Under the matching variant, the complementizer $C_{(2)}$ is merged, which attracts the external head, and the internal head is deleted. Under the raising variant, the external head is not raised but rather deleted *in situ* under identity with the raised internal head.

- (19) a. $[IP_{rel} [DP \text{ Dem } [Num [A \text{ NP}]]]]$ (merge of $C_{(0)}$ and attraction of IP) \rightarrow
 b. $IP_{relj} C_{(0)} [t_j [DP \text{ Dem } [Num [A \text{ NP}]]]]$ (merge of $C_{(1)}$ and attraction of the *wh*-pronoun/ 'internal Head') \rightarrow
 c. $wh_i - [C_{(1)} [IP_{rel} t_i]_j C_{(0)} [t_j [DP \text{ Dem } [Num [A \text{ NP}]]]]]$ (merge of $C_{(2)}$ and attraction of the 'external Head') \rightarrow
 d. $[DP \text{ Dem } [Num [A \text{ NP}]]]_k C_{(2)} wh_i - [C_{(1)} [IP_{rel} t_i]_j C_{(0)} t_j [t_k]]$
 quei dieci bei gattini che io amo
 'those ten nice kittens, which I love'

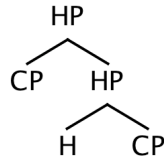
The account Cinque (2008) provides for non-integrated ARCs is more tentative. Assuming that Kayne's LCA may extend to Discourse Grammar, linear precedence in a discourse is taken to reflect asymmetric c-command. This fact is captured by assuming that a linearly preceding main sentence is placed in the Specifier of a

⁸ With this proposal, Cinque (2003, 2009) aims to derive the different types of relative clauses found in the languages of the world (post-nominal, pre-nominal, internally headed, headless, correlative) from the same structure. To implement this idea, he proposes that pre-nominal relatives (found in "rigid" OV languages) reflect the structure of merge, and that postnominal (and the other types of) relatives are derived from this basic structure.

⁹ RRCs differ from ARCs in this respect; whereas in the former, the relative IP is merged in the Specifier above the Specifiers that host attributive adjectives and numerals and below the projection that hosts determiners and demonstratives, in contrast, in the latter, the relative IP is merged in a Specifier of a nominal projection dominating DP (to be outside of the scope of the determiner or demonstrative).

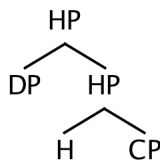
(empty) head, which, in turn, takes the following main sentence as its complement, as in (20):

(20)



However, discourse fragments do not always involve concatenation at the CP level. A DP may precede a CP, as in: *A pink shirt? I will never wear any such thing in my life!* (Cinque 2008: 118). Hence, a representation like (21) is also available:

(21)



According to Cinque, similar representations are involved in non-integrated ARCs. In this case, the complement position hosts the ARC and the Specifier hosts the relevant discourse unit containing the antecedent. If non-integrated ARCs take an antecedent across the discourse, the Specifier hosts the preceding sentence (or discourse fragment), as in (20). If there is adjacency between a nominal antecedent and the non-integrated ARC, the Specifier position hosts a DP, as in (21).

In sum, Cinque (2008) claims that there are two types of ARCs: an integrated and a non-integrated type. Languages may display both, one, or neither of the syntactic types. According to the author, the traditional view of the ARCs as a homogeneous class results from the fact that earlier studies focused on English, a language that only displays the non-integrated type.

3. Towards a dual approach to ARCs: evidence from Portuguese

This section provides empirical evidence from Portuguese to support the idea that ARCs are not a unitary class. The thrust of the argument is that ARCs introduced by *o*

qual have undergone a change from one syntactic type to another throughout the history of Portuguese.

This section is organized as follows. Section 3.1 starts with an outline of the properties of *o qual*-ARCs in Portuguese. Section 3.2 sets out eight syntactic properties that differentiate contemporary *o qual*-ARC from *o qual*-ARCs in earlier stages of Portuguese. Whenever possible, a relation is established between the variation in the diachronic and synchronic dimensions both within a language and across languages.

3.1. *O qual*-ARC: an overview of its basic use and properties

The relative pronoun *o qual* (and its counterparts in other Romance languages: Italian *il quale*, Engadine *il quèl*, French *lequel*, Catalan *el qual*, Spanish *el cual*, Rumanian *care*) is etymologically derived from the Latin form QUĀLIS (see, among others, Posner 1996: 306).

In Latin, QUĀLIS was used as a *wh*-element in interrogative clauses (see (22)) and exclamatives (see (23)). It could also occur in correlative structures of the type QUĀLIS ... TALIS (see (24)) (cf. Ernout and Thomas 1972: 156).

(22) *qualis ista philosophia est?*
what this philosophy is

‘What kind of philosophy is this?’

(Cicero, *De Finibus* 2, 27, cited in Ernout and Thomas 1972: 156)

(23) *hei mihi, qualis erat!*
oh my what was:3SG

‘Oh my! How sad he looked!’

(Vergilius, *Aeneis*. 2, 274, cited in Ernout and Thomas 1972: 156)

(24) *qualis pater, talis filius*
like father like son

‘Like father, like son.’

(Stone 2005: 198)

However, there is no evidence for the use of QUĀLIS as a relative pronoun in Classical Latin, nor in Vulgar Latin (see Middleton 1999: 121-122). For this reason, there is a

debate in the literature regarding the emergence of this relative pronoun. The basic question is whether it is a Romance innovation or if it is the result of a process that started in Latin. For further details about this debate, see Kunstmann (1991), Middleton (1999), and Ramat (2005).

In CEP, the form *qual*, derived from QUĀLIS, can still be used in interrogatives (see (25)), exclamatives (see (26)), comparatives (see (27)), and as a member of correlative connectives (see (28)).

(25) a. Quais livros compraste?
what books bought:2SG

‘What books did you buy?’

b. Quais compraste?
what bought:2SG

‘Which ones did you buy?’

c. Qual deles tinha coragem para começar?
which of.them had courage to start:INF

‘Which of them had the courage to start?’

(*a.* and *b.* examples are from Brito, Duarte and Matos 2003: 464; *c.* example is from Cunha and Cintra 1997: 355)

(26) Quais feitios, qual vida!
what tempers what life

‘Tempers life, my eye!’¹⁰

(Miguel Torga, *Contos da Montanha*, cited in Cunha and Cintra 1997: 356)

(27) Nadava qual peixe.
swam:3SG like fish

‘He swam like a fish.’

(Cuesta and Luz 1971/1980: 507)

(28) Este chapéu é tal qual o meu.
this hat is just like the mine

‘This hat is just like mine.’

(Cuesta and Luz 1971/1980: 507)

¹⁰ The sentence in (26) expresses metalinguistic negation, which occurs in English sentences such as *Like hell Al and Hilary are married; Al and Hilary are married my eye* (Drozd 2001: 55).

Crucially, in these contexts, *qual* cannot be preceded by a definite article. In contrast, when introducing relative clauses, *qual* is always preceded by the definite article *o* ‘the’; see (29).¹¹

- (29) Este governo atacou os direitos dos professores, *(os) quais
 this government attacked the rights of.the teachers the which
 ficaram bastante prejudicados com as medidas tomadas.
 stayed very affected with the measures taken
 ‘This government limited the rights of the teachers, who were very affected by the
 measures implemented.’

In this case, the definite article is incorporated in the relative pronoun because no element can break the adjacency between the definite article and the *wh*-element (**o* preposition/adverb/noun/adjective *qual*). The internal complexity of *o qual* explains why it is sometimes dubbed a *complex pronoun* (see Inada 2007).

Within the complex pronoun, the definite article is inflected for gender and number, whereas *qual* is inflected only for number (*o qual* ‘the:MASC.SG which:SG’; *os quais* ‘the:MASC.PL which:PL’; *a qual* ‘the:FEM.SG which.SG’, and *as quais* ‘the:FEM.PL which:PL’).

The main syntactic and semantic properties of *o qual*-ARCs in CEP have already been reported in the literature by Brito (1991) and Brito and Duarte (2003). I will briefly mention two relevant properties here for the sake of completeness:

- The pronoun *o qual* is compatible with human and non-human antecedents; see (30)a-b below:

¹¹ The same is true of other Romance languages: the counterparts of the relative *o qual* also incorporate a definite article. The only exception is the Rumanian *care*, which cannot be preceded by the article. It could have an article, however, in the 19th century: masc. *care-le*, fem. *care-a* (see Ramat 2005).

(30) a. Já entreguei o processo ao meu advogado, o qual dispõe
already gave:1SG the process to.the my lawyer the which has
 de um mês para contestar a decisão.
DE:PREP a month to contest:INF the decision

‘I have already given the process to my lawyer, who has a month to contest the decision.’

b. No passado dia 2 realizou-se a 5ª edição do Portugal
in.the last day 2 took.place.SE:CL the 5th edition of.the P.

Fashion, na qual participaram dez estilistas portugueses.
Fashion in.the which participated ten fashion.designers Portuguese

‘On the 2nd of this month, the 5th edition of the Portugal Fashion took place, in which ten Portuguese fashion designers participated.’

• *o qual*-ARCs can relativize:

- The subject (as in (30)a).
- The direct object (as in (31)).
- The object of a preposition ((32) illustrates the relativization of the indirect object and (30)b the relativization of the oblique complement).¹²

(31) Recebemos algumas candidaturas incompletas, as quais excluimos
received:1PL some applications incomplete the which excluded:1PL
 de imediato.
of immediate

‘We received some incomplete applications, which we excluded immediately.’

(32) O ministro foi alvo de muitas críticas, às quais respondeu com
the minister was target of many critiques to.the which answered:3SG with
 agressividade.
aggressiveness

‘The minister was subject to severe criticism, to which he replied in an aggressive manner.’

¹² In CEP, *o qual* may also introduce RRCs. However, in such syntactic environment, it cannot relativize the subject and the object. By way of illustration, see examples in (i), taken from Brito (1991: 156).

(i) a. *O homem o qual escreveu é meu amigo.
the man the which wrote is my friend

‘The man who wrote (it) is a friend of mine.’

b. *O homem o qual eu vi ontem é meu amigo.
the man the which I saw yesterday is my friend

‘The man whom I saw yesterday is a friend of mine.’

Finally, a word is in order regarding the use of *o qual*-ARCs in CEP. In general, relative clauses introduced by *o qual* are perceived by the speakers as a formal and ‘prestige’ construction. Its use is regarded as somewhat artificial and less natural than, for instance, the use of relative clauses introduced by *que* ‘that’ (the same form as the complementizer). Closely related to this is the fact that relative clauses introduced by *o qual* are more frequent in written than in spoken Portuguese.

To confirm this tendency, a corpus-based analysis of spoken and written Portuguese was performed. Two corpora of roughly the same size were selected: *C-ORAL-ROM* (containing 320,452 words, in the Portuguese section), for spoken Portuguese, and a subcorpus of CRPC (*jornal_annotado_RL*), containing texts from some Portuguese newspapers (with 336,151 words). The results clearly showed the tendency of *o qual* to occur in written texts; the spoken corpus yields 56 tokens of *o qual*, whereas the written corpus contains 207 tokens.¹³

The same tendency was observed for other contemporary Romance languages. Corominas and Pascual (1980: 257ff) show that contemporary spoken Spanish has totally abandoned *el cual*, and Fiorentino (1999: 92-93) notes the rarity of *il quale* in a spoken corpus of contemporary Italian (see Ramat 2005, and references therein).

Earlier stages of Portuguese pattern with CEP with respect to the properties of *o qual*-ARCs just mentioned.

Regarding the internal structure of *o qual* in ARCs, the same pattern is found: in the corpus edited by Martins (2001), the *wh*-element is always combined with the definite article, and no element can disrupt the article+*wh*-element sequence. In this corpus, the *wh*-element may occur without the definite article, but not in ARCs.¹⁴ Its occurrence is limited to RRCs (see (33)) and free relative clauses, with or without an additional internal head (see (34) and (35), respectively).¹⁵

¹³ Note that these tokens include both *o qual*-ARCs and *o qual*-RRCs.

¹⁴ For the occurrence of *qual* without article, see also Clarinda Maia (1986: 696) and Mattos e Silva (1989: 752, fn. 17). The examples cited in these studies seem to corroborate the idea that *o qual* without an article introducing a headed relative clause preferentially has a restrictive interpretation.

¹⁵ In the corpus edited by Martins (2001), *qual* (without the definite article) can also occur as a member of the correlative pair *tal ... qual* ‘such ... as’ (see (i)) and as an element introducing a nominal constituent (see (ii)).

- (33) que faça ou façã ende stromêto ou stromêtos
that make: PRES.SUBJ.3SG or make: PRES.SUBJ.3PL of.it deed or deeds
 quaes les o dito Steuã perez mãdar fazer
which to.them:CL the mentioned S. P. order:FUT.SUBJ make:INF
 ‘(And I order that) he make or they make the deed or deeds that the aforementioned Steuã Perez ordered them to make.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1278)

- (34) quays fforom presentes.
which were present
 ‘who were present: [list]’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1292)

- (35) e pera fazer ende carta [...] per qual Tabellion que a esta cousa
and to make:INF of.it letter by which notary that to this thing
 for demandado
be:FUT.SUBJ summoned
 ‘and to compose a letter (...) by whichever notary that be summoned to this thing.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1290)

Likewise, earlier stages of Portuguese allow *o qual*-ARCs to take human and non-human antecedents, as shown in (36)-(37).

- (i) e quaéés dereituras sêpre deu táées dares tu;
and which rents always gave:3SG such give:INF.2SG you
 ‘And you must pay the same rent that he paid.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1272)
- (ii) damos [...] a uos Afonso rodriguiz nosso irmão qual filho de nosso
give:IPL to you A. R. our brother which son of our
 padre quantos herdamentos nos auemos
father all...that lands we have
 ‘We give you Afonso Rodriguiz, our brother, as son of our father, all the lands that we have ...’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1295)

- (36) todollos herdamẽtos e Cassaes que nos Auemos [...] ãno Couto
all.the lands and hamlets that we have in.the ±property
 de negrelhos que este Alen doyro o qual ha nos fficou de
of N. that is ±beyond D. the which to us:CL stayed from
 parte de nosso padre
part of our father

‘all the lands and hamlets that we have in the property of Negrelhos, which is beyond Doyro, which was left to us by our father.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1304)

- (37) e leixo a dita mha tesstamẽteira por affom que auera
and leave:1SG to mentioned my executor by work that have:FUT.3SG
 dez libras aA qual dou e outorgo comprido poder
ten l. to.the which give:1SG and grant:1SG full power

‘And I leave ten *libras* [the currency] to my executor for the work that she will have; and I give and grant her full power (to distribute my money and property).’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1383)

Finally, *o qual*-ARCs in earlier stages of Portuguese pattern with their contemporary counterparts in the possibility of relativizing the subject (see (36)), the object of a preposition (see (37), which illustrates the relativization of the indirect object), and the direct object (see (38)).¹⁶

¹⁶ In the case of RRCs introduced by *o qual*, there is, however, one important contrast between CEP and earlier stages of Portuguese. As mentioned in fn.12, *o qual*-RRCs in CEP can only relativize the object of a preposition (see (i) below). This restriction does not hold, however, for earlier stages of Portuguese (cf. Lucchesi 1990). See, for instance, (i) and (ii) below, where the direct object is relativized.

- (i) O Millennium é o banco ao qual recorro mais vezes.
the M. is the bank to.the which resort.to:1SG more times
 ‘The Millennium is the bank I work with more often.’

- (ii) obligeo a uos e empenhoro hũa nossa Casa. a qual nos auemos
pawn:1SG to you and pledge:1SG a our house the which we have
 en Lixbõa.
in L.

‘I pawn and pledge to you a house of ours that we have in Lixbõa.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1286)

- (iii) que façã a eles entregar todas as herdades as quaes a esses
that make:3PL to them return:INF all.the lands the which to those
 Moesteiros e a essas Egreias tẽẽ e teuerõ filhados.
monasteries and to those churches have:3PL and had:3PL seized

(to be continued)

(38) per hũa procuraçõ feyta per mááo de Domĩgos stephães Tabellion
by one letter.of.attorney made by hand of D. S. notary
 das Alcaçouas. a qual eu Johã soarez Tabellion da Cidade de
of.the A. the which I J. S. notary of,the city of
 Lixbõa ui, líj
L. saw read

‘by one letter of attorney created by Domĩgos Stephães, notary of Alcaçouas, which I, Johã Soarez, notary of Lisboa, saw and read.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1286)

3.2. Two types of *o qual*-ARCs: diachronic evidence

Whereas the preceding section focuses on some properties with respect to which *o qual*-ARCs in CEP and in earlier stages of Portuguese behave alike, the present section offers an overview of the syntactic properties that differentiate the two constructions. Section 3.2.1 is devoted to three of these syntactic properties: the occurrence of an additional internal head (Section 3.2.1.1), the extraposition of the *o qual*-ARC (Section 3.2.1.2), and the restrictions on pied-piping (Section 3.2.1.3). The first two properties were partially discussed in Cardoso (2008, 2009) and Cardoso and De Vries (2010).

Section 3.2.2 considers additional syntactic properties, namely, clausal antecedents (Section 3.2.2.1), split antecedents (Section 3.2.2.2), coordination of the *wh*-pronoun with another DP (Section 3.2.2.3), illocutionary force (Section 3.2.2.4), and the presence of a coordinator (Section 3.2.2.5).

As the reader will notice, Section 3.2.1 will require more detail (and space) than Section 3.2.2. There are two reasons for this contrast. First, the properties discussed in Section 3.2.1 are robustly attested in earlier stages of Portuguese, whereas the phenomena reviewed in Section 3.2.2 are not as well represented in the corpora. Second, the properties discussed in 3.2.1 are subject to various restrictions and thereby require more complex descriptive devices.

‘(I order) that they make them return all the lands that they have seized from those monasteries.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; non-dated document, 13th century)

3.2.1. Some properties in contrast

3.2.1.1 Internal head

O qual-ARCs in CEP cannot exhibit an additional internal head. This impossibility is illustrated in (39) and (40) with a nominal and a non-nominal ARC, respectively:¹⁷

- (39) *Existem argumentos fortes a favor dessa análise,
there are arguments strong in favor of that analysis
 os quais argumentos apresentarei de seguida.
the which arguments present:FUT.ISG next

‘There are strong arguments in favor of that analysis, which arguments I will present next.’

- (40) *Os portugueses não gostam de música portuguesa,
the Portuguese people not like of music Portuguese
 o qual facto explica a escassa produção musical deste país.
the which fact explains the sparse production musical of this country

‘The Portuguese people do not like Portuguese music, which fact explains the sparse musical production of this country.’

Two alternative constructions can be used in these contexts: (i) the ARCs can be introduced by a relative pronoun (and no internal head is present) (see (41)a-(42)a); (ii) the appositive construction can surface with an additional external head (see Section 4.1.2); in this case, a noun phrase is modified by an RRC, and the complex (noun phrase + RRC) is in apposition to the antecedent (see (41)b-(42)b).¹⁸

¹⁷ The Portuguese judgments reported in this chapter are my own and are shared by the speakers I consulted. According to Brito (1991: 133), *o qual*-ARCs with an internal head are marginal in CEP. Smits (1988) also notes that this construction cannot have an additional internal head in CEP (see Table 2 below). I will come back to this topic in Section 6.2.

¹⁸ For more details on the constructions in (41)b and (42)b, see Peres and Mória (1995: 270-1), Brucart (1999: 423), and Brito and Duarte (2003: 674-5).

- (41) a. Existem argumentos fortes a favor dessa análise,
there.are arguments strong in favor of.that analysis
 os quais apresentarei de seguida.
the which present: FUT.ISG next
 ‘There are strong arguments in favor of that analysis, which I will present next.’¹⁹
- b. Existem argumentos fortes a favor dessa análise,
there.are arguments strong in favor of.that analysis
 argumentos que apresentarei de seguida.
arguments that present: FUT.ISG next
 ‘There are strong arguments in favor of that analysis, arguments that I will present next.’
- (42) a. Os portugueses não gostam de música portuguesa,
the Portuguese.people not like of music Portuguese
 o que explica a escassa produção musical deste país.
the which explains the sparse production musical of.this country
 ‘The Portuguese people do not like Portuguese music, which explains the sparse musical production of this country.’
- b. Os portugueses não gostam de música portuguesa,
the Portuguese.people not like of music Portuguese
 facto que explica a escassa produção musical deste país.
fact that explains the sparse production musical of.this country
 ‘The Portuguese people do not like Portuguese music, a fact that explains the sparse musical production of this country.’

However, languages do not behave uniformly regarding the occurrence of an additional internal head. Smits (1988) shows that there is cross-linguistic variation in this respect, as summarized in Table 2.

¹⁹ Note that the English translation of the sentence (41) is ambiguous because *which* can take as antecedent either *analysis* or *strong arguments*. In CEP, this ambiguity does not arise because the relative pronoun *os quais* is masculine plural, which indicates that the antecedent is necessarily *fortes argumentos* ‘strong arguments.’

Table 2. ARCs with an additional internal head (based on Smits 1988: 65)

Languages	Pronoun	Nominal-ARCs	Non-nominal-ARCs
Italian	il quale	+	-
Spanish	el cual	+	-
Catalan	el qual	arch	+
French	lequel	+	+
Dutch	welke	+	+
German	welcher	+	+
English	which	+	+
Swedish	vilken	+	
Norwegian	hvilken	+	
Danish	hvilken	+	
Portuguese			
Romanian		no form	
Icelandic			

Some examples of this construction are given in (43) to (48) below (taken from Smits 1988: 65, 306, 321, 369, 272, 288).²⁰

[Italian]

(43) %Cercavo una ragazza, con la quale ragazza uscire a cena.

‘I was looking for a girl, with which girl to go out and dine.’

[Spanish]

(44) %Los ejemplos de este fenómeno que he presentado, en los cuales ejemplos he pensado mucho, no dejan de confundirme.

‘The examples of this phenomenon that I presented, about which examples I have thought much, never cease to confuse me.’

[French]

(45) Toutes les idées que j'aurais à développer, lesquelles idées sont exposées en détail dans ce mémoire, ...

‘All the ideas that I would have to develop, which ideas are laid out in this report...’

[Dutch]

(46) Moby Dick, welk boek mij als kind al mateloos boeide, is nog steeds mijn favoriete roman.

‘Moby Dick, which book already as a child enthralled me, is still my favorite novel.’

²⁰ Smits (1988) uses the symbol ‘%’ to indicate that it is a highly formal and marked construction.

[Dutch]

(47) Er sagte "Guten Tag", welchen Gruß sie freundlich erwiderte.

‘He said “good day”, which greeting she friendly returned.’

[English]

(48) My dog, which faithful animal has guarded me for years, died last week.

Interestingly, variation is also found diachronically. Several authors have pointed out that *o qual*-ARCs in earlier stages of Portuguese can exhibit an additional internal head (see Dias 1933/1970: §93, Said Ali 1931/1971: §515-516, Huber 1933/1986: §347, Barreto 1911/1980: 141, Neto 1957/1970: 509, Maia 1986: 696-697, Lucchesi 1990: 181, A. Costa 2004: 419). A case in point is the following:

(49) entrego e outorgo. ao Mosteiro de san Saluador de Moreyra. hũu
give:1SG andconcede:1SG to.the monastery of S. S. de M. a

casal que e en Rial de Pereyra. o qual casal a dita dona
hamlet that is in R. de P. the which hamlet the mentioned D.

Mayor uêegas [...] mandou ao dito Mosteiro.
M. V. left to.the mentioned monastery

‘I give and concede a hamlet that is (located) in Rial de Pereyra to the monastery of San Saluador de Moreyra, which hamlet the aforementioned Dona Mayor Vêegas left to the monastery.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1282)

This possibility is also documented in Latin, as illustrated in (50). According to Ernout and Thomas (1972: 332) and Bassols de Climent (1967: 240), an additional internal head is found primarily in formal contexts, especially in legal documents. Its use can be explained as a strategy to avoid ambiguity when the relative and the antecedent are non-adjacent, as a way of conferring more precision to the utterance.

(50) erant omnino itinera duo, quibus itineribus domo exire
were in.all routes two by.which routes from.home leave:INF

possent
could:IMPERF.SUBJ.3PL

‘There were but two routes, by which routes they could leave home.’

(Caesar, *De Bello Gallico*, I, 6, cited in Finch 2006: 36)

In historical Portuguese, *o qual*-ARCs with an additional internal head are attested in texts from different periods and belonging to different textual typologies. Examples

(51)-(57) provide illustrations of the construction in different textual genders, namely, notarial documents (51), historiographic texts (52), dissertations (53), travel literature (54), theatre (55), letters (56), and religious texts (57). Note further that examples (49) and (51)-(57) range over different periods, from the 13th century to the 18th century.

- (51) E ffica ao dito Gomez perez e a ssa molher hũu prazo
and goes to.the mentioned G. P. and to his wife a ±contract
 que cõta que e de Orraca perez e de Affonso bēetíz pelo qual
that says that is of O. P. and of A. B. by.the which
prazo deuyã a dõna. Steuahya hũa soma de dinheiros.
±contract owed:3PL to D. S. an amount of money

‘And a contract – which was made with Orraca Perez and Affonso Bēetíz - is assigned to Gomes Perez and his wife, under which contract they owed an amount of money to Dona Steuahya.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1311)

- (52) ao quall foy emcomemdada outra torre que está junto com ha
to.the which was commissioned another tower that is close by the
 outra de Fez [...], a quall torre emtão hera chamada de Madraba
other of F. the which tower then was called of M.

‘who was commissioned another tower that is close by the other (tower) of Fez, which tower was then called Madraba.’

(Brocardo 1997; Gomes Eanes de Zurara, *Crónica do Conde D. Pedro de Meneses*, 15th century)

- (53) Esquizo são as primeiras linhas ou traços que se fazem com a pena,
sketch are the first lines or strokes that SE:CL make:3PL with a quill
 ou com o carvão, dados com grande mestria e depressa, os quaes
or with the coal made with great mastery and quicky the which
traços comprehendem a idea e invenção do que queremos fazer
strokes contain the idea and invention of what want:3PL do:INF

‘A sketch is the first lines or strokes that are made with a quill or with coal, drawn with great mastery and quickly, which strokes contain the idea and the invention of what we want to do.’

(TYC; Francisco de Holanda, *Da Pintura Antiga*; year 1548)

- (54) a origem do rio procedia de hum lago que se chamaua Pinator, que
the origin of.the river came of a lake that SE:CL called P. that
 demoraua a leste daquelle mar duzentas e sessenta legoas, no
took to east of.that sea two.thousand and sixty leagues in.the
 reyno de Quitirvão, o qual lago estaua cercado de grandes serranias
kingdom of Q. the which lake was ringed of huge mountains
 ‘the river had its source in a lake that is called Pinator, which was two thousand and
 sixty leagues east away from the sea in the kingdom of Quitirvão, which lake was
 ringed by huge mountains’

(TYC; Fernão Mendes Pinto, *Peregrinação*; 16th century)

- (55) imaginei ãa festa
imagined:1SG a party
 à nossa Júlia modesta
to.the our J. modest
 nacida per mão de Deos
born by hand of God
 a qual festa será esta.
the which party be:FUT this

‘I imagined a party / to our modest Júlia / born by the hand of God / which party will
 be as follows.’

(Camões 1999; Gil Vicente, *Inverno e Verão*; year 1529)

- (56) os padres totalmente desconfiam de os índios haverem de
the priests fully doubt of the Indians have:INF.3PL of
 descer sem violência a qual violência não é menos duvidosa
go.down:INF without violence the which violence not is less doubtful
 ‘the priests fully doubt that the Indians would go down without violence, which
 violence is no less doubtful’

(TYC; António Vieira, *Cartas do Padre António Vieira*; 17th century)

- (57) esta prodigiosa demonstraçãõ, foy a resposta que o Senhor lhe deu,
this prodigious demonstration was the answer that the Lord to.her:CL gave
 da qual veyo a entender hauia muyto que cortar [...].
of.the which came to understand there.was a.lot what cut
 A qual revelaçãõ se veyo a verificar
the which revelation SE:CL came to verify:INF

‘this prodigious demonstration was the answer that the Lord gave her, from which she
 realized that there was a lot to be cut (...). This revelation came true.’

(TYC; Maria do Céu, *Vida e Morte de Madre Helena da Cruz*; year 1721)

In what follows, I will present the essential properties of ARCs with an additional internal head in earlier stages of Portuguese. These properties are listed in A. through E. below and discussed in turn.

- A. *Categorial nature of the antecedent/internal head*
- B. *Semantic class of the nominal internal head*
- C. *Relation between the antecedent and the internal head*
- D. *Expansion of the internal head*
- E. *Contexts of occurrence*

A. *Categorial nature of the antecedent/internal head*

The antecedent of an ARC with an additional internal head can be nominal or non-nominal. ARCs with nominal antecedents are exemplified in (51) and (52) above. ARCs with non-nominal antecedents are illustrated in (58)-(59) below. Note that in (58) the antecedent is clausal, whereas in (59), it is an adverbial phrase (modified by an RRC).

- (58) os ditos cassaaes fforõ cõprados dos dinheiros do dito
the mentioned hamlets were bought of.the moneys of.the mentioned
 mosteiro polla quall Razom de derejto pertecem e perteyam
monastery by.the which reason of right belong:3PL and belonged:3PL
 ao dito mosteiro
to.the monastery

‘the aforementioned hamlets were bought with the money of the aforementioned monastery, for which reason they belong and belonged to the monastery by right’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1437)

- (59) A sombra não se ha de dar senão ali onde não alcança a lux
the shadow not SE:CL will fall:INF except there where not reach the light
 e claridade, o qual lugar fica logo inobre e é sombra.
and clarity the which place stays instantly unimpressive and is shadow

‘The shadow will not fall except there where the light and clarity do not flood, which place becomes instantly unimpressive; and it is shadow.’

(TYC; Francisco de Holanda, *Da Pintura Antiga*; year 1548)

Importantly, these examples show that, regardless of the category of the antecedent, the internal head is always nominal.

B. Semantic class of the nominal internal head

There seems to be no restriction on the semantic class of nouns that can appear as an internal head. For instance, the additional head can be a proper name (60), a count noun (61), or a non-count noun (62).²¹

- (60) o dito Johã viçente disse que a dita vjnha trouxera
the mentioned J. V. said that the mentioned vineyard had.brought
 ã outro tempo Luzia domingujz [...] A qual Luzia domingujz Era
in other time L. D. the which L. D. was
 ffinada deste mondo
deceased from.this world

‘The aforementioned Johã Viçente said that Luzia Domingujz once owned the aforementioned vineyard (by emphyteusis), which Luzia Domingujz was not anymore in this world.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1377)

- (61) cõfesamos que nos Recebemos de uos Martjn saluadorez Cjncoeëta
confess:1PL that we received of you M. S. fifty
 libras de dinheiros portugéeses as quaes #Lta libras a nos erõ
l. of currency Portuguese the which fifty l. to us were
 Julgadas per Sentêca
attributed by sentence

‘We confess that we received from you Martjn Saluadorez fifty *libras* [the currency] of the Portuguese currency, which fifty *libras* were attributed to us by sentence’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1365)

- (62) que dem e paguem de foro e pensom da
that give:PRES.SUBJ.3PL and pay:PRES.SUBJ.3PL of tenancy and rent of.the
 dicta qujntãa en cada hũu ãno hũu tonell de vinho, puro do
mentioned farm in each a year one vat of wine pure of.the
 que deus der nas vjnhas [...] o quall vinho sera vermelho
that god give:FUT.SUBJ in.the vineyards the which wine be:FUT red

lit. ‘And (I demand) that they give and pay as tenancy and rent for the aforementioned farm in each year one vat of wine, pure, of the one that God will give in vineyards, (...) which wine will be red’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1510)

²¹ Note that in (62) the antecedent *hũu tonell de vinho ...* ‘one vat of wine...’ contains a non-count noun (*vinho* ‘wine’) associated with a unit of measurement (*tonell* ‘vat’).

C. *Relation between the antecedent and the internal head*

There can be phonological and semantic identity between the antecedent and the internal head, as shown in (51)-(56) above. However, both elements can also differ. In this case, there are a number of possibilities. When the antecedent is nominal, the internal head can be a true synonym of the antecedent, as in (63). It can also express a defining property of the antecedent, as in (64), or a more specific classificatory property, as shown in (65).

- (63) *mostrarõ logo ã Jujzo húú testamẽto [...] na qual mãda*
showed:3PL immediately in judgment a testament in.the which will
fazia mẽçom Antre as outras coussas que A mãdara
made mention among the other things that it:CL had.ordered
fazer Sancha gil.
make:INF S. G.

‘They immediately showed a testament in trial (...), in which will it was mentioned, among other things, that Sancha Gil ordered (him) to make it.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1328)

- (64) *e começou a era de quatrocentos e oito: no quall ano, estando*
and started the era of four.hundred and eight in.the which year be:GER
el-rrei dom Henrique na villa de Touro, soube como
the.king D. H. in.the village of T. knew:3SG that

‘and the era of four hundred and eight started, in which year, being in the village of Touro, the king Dom Henrique became aware that ...’

(Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

- (65) *as quaees leteras forõ probicadas a dom lourenço [...] o qual arçebispo*
the which letters were addressed to D. L. the which archbishop
obedeçendo aas dictas leteras fez sobre ello seus processos
obey:GER to.the mentioned letters made on it his processes

‘which letters were addressed to Dom Lourenço (...), obeying the aforementioned letters, this archbishop made his processes (based) on it’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1402)

If the antecedent is clausal, the additional internal head is typically a general abstract noun such as *reason* in (58). However, it can also be a more specific verbal noun that is morphologically related to a verb introduced in the preceding context; see (66), where the noun *pitiçon* ‘request’ is morphologically related to the preceding verb *pedir* ‘to request.’

- (66) E como Eu dito priol lhe pedise e
and as I mentioned prior to.him:CL request:IMPERF.SUBJ and
 Mādase pedir A dita palha [...] A qual pitixon o
order: IMPERF.SUBJ request:INF the mentioned straw the which request the
 dito francisco martjnz cōtestou dela
mentioned F. M. contested of.it

‘And as I, the aforementioned prior, requested the aforementioned straw (...), which request the aforementioned Francisco Martjnz contested’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1370)

D. Expansion of the internal head

The internal head NP can be a conjoined phrase, as is shown in (67). In these contexts, there is typically first conjunct agreement for ϕ -features between the relative pronoun and the noun in the first conjunct (*casal* ‘hamlet’ in (67)).

- (67) ffazemos prazo [...] dūu Casal que auemos en Cūpustelá e dūu meio.
make:IPL ±contract of.a hamlet that have:IPL in C. and of.a half

Barco en verdugo O qual casal e meio Barco ora trage Maria
boat in V. the:SG which:SG hamlet and half boat now brings M.

‘We make a contract of a hamlet in Cūpustelá and half of a boat in Verdugo, which hamlet and half of a boat now belong to Maria.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1308)

The internal head can also be modified by different categories, such as a PP (see (68)) or even a relative clause (see (69)).

- (68) o quall casal com suas perteenças disse que trazia ã pregã
the which hamlet with its belongings said:3SG that brought in cry

‘which hamlet with its belongings he said was being announced for sale.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1472)

- (69) os quaaes dinheiros que uos eu assy hey de dar e pagar
the which moneys that you I so will give:INF and pay:INF

uos auya de pagar Joham
to.you:CL would pay:INF J.

‘which money, which I will give and pay you, would be given to you by Joham.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1436)

Furthermore, the internal head can be extended by a numeral, as in (61) above, or an adjectival possessive pronoun (70). Both are construed pre-nominally. The occurrence of universal quantifiers is also attested, typically in a post-nominal position; see (71):

(70) aos quaes meus procuradores dou cõprido poder
to.the which my attorneys give:1SG full power

lit. ‘to which my attorneys I give full power’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1341)

(71) das quaées coussas todas o dito Priol por ssj e pelo Conuêto de
of.the which things all the mentioned prior by him and by.the convent of
 sseu Mostejro pediu ende A m̃j dito tabaliõ este strumento
his monastery asked of.it to me mentioned notary this deed

‘of (all) which things the aforementioned prior in his name and in the name of the convent of his monastery asked me, the aforementioned notary, this deed.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1328)

E. Contexts of occurrence

As previously mentioned, the presence of an internal head can sometimes be explained as a way of avoiding ambiguity. Consider, for instance, the sentence in (72); in this example, the lack of the internal head could lead to ambiguity, as it would not be clear whether the antecedent was *sua força* ‘his force’ or *Autorydade* ‘authority.’

(72) o dito prioll per sua força e Autorydade lha tomara
the mentioned prior by his force and authority to.him:CL.it:CL had.taken
 fforçãdóo della A qual fforça djzia que
forcing.him:CL of.it the which force said:3SG that

‘the aforementioned prior had taken it from him by his force and authority, resorting to violence, which force he said that ...’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1379)

The presence of an additional internal head is also favored in contexts in which the antecedent and the ARC are not adjacent. In the corpus-based investigation presented in Cardoso (2008), I have shown that, in earlier periods of Portuguese until approximately the 17th century, ARCs with an additional internal head are more frequent than ordinary ARCs in contexts of extraposition, as in (73), and when the

relative clause is preceded by stacked or multiple embedded relative clauses, as in (74).

- (73) E o dicto Juiz per sentença defenetia asy o Julgou da
and the mentioned judge by sentence definitive this.way it:CL judged of.the
 quall sentença o dicto Reeo nō apellou
which sentence the mentioned defendant not appealed

‘And the aforementioned judge by a definite sentence judged this way (=passed this sentence), against which sentence the aforementioned defendant did not appeal.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1422)

- (74) Sabhã todos que [...] ã hũa ṽjnha que chamã o cõchouso
know:PRES.SUBJ.3PL all that in a vineyard that call:3PL the C.

que e A par da de Pero. caramos termho de Santarẽ A qual ṽjnha
that is next of.the of P. C. environs of S. the which vineyard

deziam que Era do Moesteyro dachellas
said:3PL that was of.the monastery of.C.

‘(I want you) all to know that in a vineyard that people call the Cõchouso, that is next to the (vineyard) of Pero Caramos, environs of Santarẽ, which vineyard people said belonged to the Monastery of Chellas’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1377)

However, as the examples in (61) above and (75) below illustrate, none of this is a necessary condition.

- (75) segundo [...] he cõheudo no estormento do dicto enprazamento do
as is contained in.the deed of.the mentioned emphyteusis of.the
 qual enprazamento ella dicta ljanor de ualladares he a segunda pessoa
which emphyteusis she called L. de V. is the second person

‘as (...) is contained in the deed of the aforementioned emphyteusis, of which emphyteusis she Ljanor de Ualladares is the second person’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1482)

To sum up, concerning the presence of an internal head in *o qual*-ARCs, the following conclusions can be drawn.

- CEP does not allow ARCs with an internal head.
- This construction was available in earlier stages of Portuguese (and in other languages, e.g., French, Dutch, and German).
- Essential properties of this construction are as follows:

- A non-nominal antecedent is allowed, as in regular appositive constructions.
- The internal head NP is often not identical to the external antecedent and can be of various semantic classes.
- The internal head is not just a nominal head but is a phrase that can be enriched by PPs or other modifiers; it can also be a conjoined noun phrase.
- The occurrence of an internal head confers more precision to the utterance, avoiding ambiguity and providing an important clue to recover the antecedent when heavy intervening material is present.

3.2.1.2 Extraposition

Based on examples like (76), Brito (2004) argues that extraposition of ARCs is not possible in CEP.²²

- (76) a. Vi o João, que é o meu amigo preferido.
saw:1SG the J. that is the my friend favorite
 ‘I saw João, who is my favorite friend.’
- b. O João, que é o meu amigo preferido, foi visto por mim.
the J. that is the my friend favorite was seen by me
 ‘João, who is my favorite friend, was seen by me.’
- c. *O João foi visto por mim, que é o meu amigo preferido.
the J. was seen by me that is the my friend favorite

(Brito 2004: 402)

The same claim has been made for other languages. For instance, Ziv and Cole (1974: 777-778), Emonds (1979: 234-235) and Alexiadou et al. (2000: 31) assume that ARCs cannot be extraposed in English, as exemplified in (77).²³

²² Following the same practice as in Chapter 3, the elements that intervene between the antecedent and the relative clause are referred to as *intervening material* and are underlined for expository purposes.

²³ According to Alexiadou et al. (2000), extraposition of ARCs appears to be marginally possible with presentative focus on the antecedent, as in (i)b.

(i) a. *John arrived, who happens to be an expert in aerodynamics.

(to be continued)

- (77) a. *A boy was kissing Mary, whom I had never seen before.
 b. *A boy was here, whom I had never seen before.
 c. *John was here, whom I had never seen before.
 d. *My father just came in, who runs his own business.

(Ziv and Cole 1974: 777-778)

Vergnaud (1974: 181) also asserts the non-extraposability of ARCs in French:

- (78) *Paul vient de passer qui portait un fedora.
 ‘Paul just passed wearing a fedora.’

However, this view has recently been challenged by various authors (see De Vries 2002, 2006b, Arnold 2007, Strunk 2007). On the basis of examples like (79) from Dutch, De Vries (2002, 2006b) refutes the traditional view, showing that it is plainly false that ARCs cannot be extraposed.

- (79) Gisteren heb ik mijn zuster bezocht, die blond haar heeft
yesterday have I my sister visited, who blond hair has
 (zoals je weet).
 (as you know)

‘Yesterday I have visited my sister, who has blond hair (as you know).’

(De Vries 2006b: 254)

Additional counterexamples can be found in English and German, as shown, respectively, in (80), taken from Arnold (2007: 306), and (81), taken from Strunk (2007):

- (80) I was also given a Jubilee mug at school, which I still have.

(BN6/0630)

-
- b. ?? **John** arrived, who happens to be an expert in aerodynamics.

(Alexiadou et al. 2000: 31)

- (81) Allerdings habe er mit Prodi gesprachen, zu dem er "ein
however has he with Prodi talked to who he a
 enges und intensives Verhältnis" pflege.
close and intensive relationship cultivates
 'However, he has spoken with Prodi, with who he has a close and intensive
 relationship.'
 (TübingerBaumbankdes Deutschen/ Schriftsprache(TüBa-D/Z), sentence 10665)

Likewise, and contrary to traditional belief, ARCs can be extraposed in CEP; see (82), taken from a CEP newspaper corpus:²⁴

- (82) O leiloeiro, para não levantar suspeitas, utilizava ainda um outro
the auctioneer to not arouse:INF suspicions used still an other
 indivíduo nos negócios, o qual muitas vezes aparecia a arrematar
man in.the negotiations the which many times showed.up A buy:INF
 os bens em seu lugar.
the goods in his place
 'Not to arouse suspicion, the auctioneer used yet another man in the negotiations, who
 showed up many times buying goods at auction in his place.'
 (CETEMPúblico 1.7 v. 4.0)

In some discourse contexts, ARC extraposition may even be obligatory, as illustrated in (83) below. In this case, the event referred to in the ARC is subsequent to that referred to in the main clause. Hence, ARC extraposition must take place to respect the sequence of events: the crash into a lamppost occurs after the passenger is thrown into the air. This type of ARC has been referred to in the literature as a continuative appositive clause (see, e.g., Jespersen 1949 and Loock 2007) and supplementary ARC (see Huddleston, Pullum, and Peterson 2002).

²⁴ Given the object of study of the present chapter, I will henceforth focus the discussion on extraposition of *o qual*-ARCs. It seems to me that the observations made here may apply, with the relevant adaptations, to ARCs introduced by other relativizers. However, further investigation is necessary in this area.

- (83) a. O carro despistou-se, projectando um passageiro pelo ar, o qual
the car skidded throwing a passenger by.the air the which
 foi embater contra um poste.
was crash:INF against a lamppost
 ‘The car skidded, throwing a passenger into the air, who crashed into a lamppost.’
 (adapted from Peres and Mória 1995: 367)
- b. *O carro despistou-se, projectando um passageiro, o qual foi
the car skidded throwing a passenger the which was
 embater contra um poste, pelo ar.
crash:INF against a lamppost by.the air

However, just as observed for RRCs (see Chapter 3), there are severe restrictions on the extraposition of *o qual*-ARCs in CEP. These restrictions are listed in A. through D. below and discussed in turn.

A. *The definiteness effect*

B. *Restriction on extraposition from embedded positions*

C. *Restriction on extraposition from pre-verbal positions*

D. *The intervening material*

A. *The definiteness effect*

In CEP, the antecedent of an extraposed *o qual*-ARC can be a weak noun phrase but not a strong noun phrase (in the sense of Milsark 1974). The contrasts given in (84)-(86) illustrate this point: an extraposed *o qual*-ARC can be made acceptable if the antecedent is changed from a strong noun phrase to a weak noun phrase.²⁵

²⁵ At least for some speakers, it seems possible to have an extraposed *o qual*-ARC with a strong noun phrase (in pre-verbal position) as antecedent. This is illustrated in (i) below, taken from a newspaper corpus of CEP. Note, however, that, according to my own judgments and the informants I consulted, this sentence is ungrammatical.

- (i) Na região da Trofa, dos quatro fogos registados, o mais difícil
in.the region of.the T. of.the four fires registered the most difficult
 de combater ocorreu em S.Mamede do Coronado, o qual implicou
of fight:INF occurred in S.M. d. C. the which required
 ainda a ajuda dos bombeiros da Maia, Matosinhos e Santo Tirso.
also the help of.the firemen from.the M. M. and S. T.
 ‘In the region of Trofa, four fires took place. The most difficult fire to fight occurred in S. Mamede do Coronado, which also required the help of the firemen from Maia, Matosinhos, and Santo Tirso.’

- (84) Em França, um grupo de skinheads atirou *o/um jovem marroquino
in F. a group of skinheads threw the/a young.man Moroccan
ao rio Sena, o qual acabaria por morrer afogado.
to.the river S. the which end.up:COND by die:INF drowned
 ‘In France, a group of skinheads threw the/a young Moroccan man into the river
 Seine, who would end up drowning.’
- (85) Deverá ser construída brevemente *a /uma ponte no Barreiro,
shall:FUT be:INF built soon the/a bridge in.the B.
 a qual terá mais de 5 quilómetros e cerca de 5 faixas de rodagem.
the which have:FUT more of 5 kilometer and about of 5 lanes
 ‘The/a new bridge, which will be more than 5 kilometers and have about 5 lanes, will
 be built soon in Barreiro.’
- (86) A presidente demitiu-o e nomeou *o outro/outro conselheiro para
the president fired.him:CL and appointed the other another counselor to
o seu lugar, o qual tomou posse logo no dia seguinte.
the his place the which took office right in.the day following
 ‘The president fired him and appointed the other/another counselor in his place, who
 took office right on the following day.’

B. Restriction on extraposition from embedded positions

Extraposition of *o qual*-ARCs is not allowed when the antecedent is the object of a preposition. Examples (87)-(89) illustrate this impossibility.²⁶

- (87) *Foi preso o mestre de uma embarcação ontem,
was arrested the:MASC skipper:MASC of a:FEM boat:FEM yesterday
 na qual foram encontrados 10 quilos de cocaína.
in.the:FEM which:FEM were found 10 kilos of cocaine
 ‘The skipper of a boat was arrested yesterday; 10 kilos of cocaine were found in the
 boat.’

²⁶ The judgments concerning these sentences vary with the particular example. Moreover, some speakers consider the example (87) even more marginal than examples (88)-(89) (hence, the */?? contrast). This may well be due to the fact that the object of a preposition is more embedded in (87) than in (88)-(89); whereas (88)-(89) show extraposition from within a PP, (87) shows extraposition from a PP within a DP.

- (88) *Discuti com um amigo meu ontem, o qual teima em
argued:1SG with a friend mine yesterday the which insists on
 dizer que não vai votar nas próximas eleições.
say:INF that not goes vote:INF in.the next elections
 ‘Yesterday I argued with a friend of mine; he insists on saying that he is not going to
 vote in the next elections.’

- (89) ??Vou candidatar-me a uma câmara no próximo ano,
go:1SG run:INF to a town.council in.the next year
 na qual permanecerei até final do mandato.
in.the which stay:FUT.1SG until end of.the mandate
 ‘I am going to run for a position on a town council; I will stay there until the end of
 the mandate.’

C. Restriction on extraposition from pre-verbal positions

C1. Restriction on extraposition from pre-verbal subjects

Extrapolated *o qual*-ARCs can take post-verbal subjects as antecedents, as shown in (90)a-(91)a. However, if the subject is construed pre-verbally, the sentence is out, as shown in (90)b-(91)b.

- (90) a. Terá lugar uma reunião no dia 21 de Setembro,
have:FUT place a meeting on.the day 21 of S.
 na qual se discutirá a viabilidade do projecto.
in.the which SE:CL discuss:FUT.3SG the viability of.the project
 ‘A meeting will take place on September 21; the viability of the project will be
 discussed there.’

- b. *Uma reunião terá lugar no dia 21 de Setembro,
a meeting have:FUT place on.the day 21 of S.
 na qual se discutirá a viabilidade do projecto.
in.the which SE:CL discuss:FUT.3SG the viability of.the project

- (91) a. Será adoptado um novo modelo de avaliação de professores
be:FUT adopted a new model of evaluation of teachers
no próximo ano lectivo, do qual todos os professores discordam.
in.the next year school of.the which all the teachers disagree
 ‘A new evaluation model of teachers will be adopted in the next school year; all
 teachers disagree with it.’

- b. *Um novo modelo de avaliação de professores será adoptado no próximo ano lectivo, do qual todos os professores discordam.
a new model of evaluation of teachers be:FUT adopted in.the next year school of.the which all the teachers disagree

C2. Restriction on extraposition from other pre-verbal constituents

Extrapolated *o qual*-ARCs cannot take as antecedent a topicalized constituent. This impossibility is illustrated by the contrasts displayed in (92) and (93):²⁷

- (92) a. *Filmes cómicos, não aprecio, com os quais todos se riem às gargalhas.
movies comic not appreciate:1SG with the which all SE:CL laugh to.the laughter

‘I do not appreciate comedy movies, with which everyone roars with laughter.’

- b. Filmes cómicos, com os quais todos se riem às gargalhadas, não aprecio.
movies comic with the which all SE:CL laugh to.the laughter not appreciate:1SG

- (93) a. *Com pessoas destas, não vale a pena trabalhar, com as quais nunca se pode contar.
with people of.these not is.worth work:INF with the which never se:CL can count:INF

‘It is not worth working with these kinds of people, whom we can never count on.’

- b. Com pessoas destas, com as quais nunca se pode contar, não vale a pena trabalhar.
with people of.these with the which never SE:CL can count:INF not is.worth work:INF

However, this restriction does not hold for other constituents at the left periphery. As shown in (94) and (95), extraposition is allowed when the antecedent is a *wh*-constituent or a preposed focus:²⁸

²⁷ Examples (92)b and (93)b are a little marked (given the heaviness of the topicalized constituent), but acceptable in general.

²⁸ The description of RRC extraposition set out in Chapter 3 also takes into account sentences involving preposed emphatic/evaluative phrases. This context is not considered here because *o* (to be continued)

(94) Que desporto é que tu praticas, sem o qual não sobreviverias?
what sport is that you practice without the which not survive:COND.2SG
 lit. ‘What sport do you practice, without which you would not survive?’

(95) Outras pessoas se manifestaram contra a barragem,
another people SE:CL demonstrated against the dam
 com as quais eu concordei inteiramente.
with the which I agreed fully
 ‘Another group of people, whom I fully agreed with, demonstrated against the dam.’

D. *The intervening material*

In CEP, if the antecedent occurs in a post-verbal position, only prepositional phrases (see (96)) and adverbs (see (97)) can break the adjacency between the antecedent and *o qual*-ARC.²⁹ These constituents can be either adjuncts (as in (96)-(97)) or complements of the verb (as in (98)-(99)).

qual-ARCs cannot take emphatic/evaluative phrases as antecedent, as shown by the non-extraposed variant in the *b.* example below.

- (i) a. *Muito whisky o João bebeu, com o qual ficou completamente embriagado!
a.lot.of whisky the J. drank with the which got completely drunk
 ‘João drank a lot of whisky; he got completely drunk on it.’
 b. *Muito whisky, com o qual ficou completamente embriagado, o João bebeu!
a.lot.of whisky with the which got completely drunk the J. drank

²⁹ A time-denoting expression as the *after*-phrase given in (i) may also be licensed as intervening material. However note that, according to Mória (2000: 199), a null locating operator precedes the sequence (\emptyset_{EM} dois dias depois / \emptyset_{IN} two days after); hence, it can be assumed that (i) also involves a PP as intervening material.

- (i) Este documento mereceu um despacho de Carlos Borrego, dois dias depois, no qual o ministro concorda com o exposto
this document deserved a dispatch from C. B. two days after in.the which the minister agrees with the exposed
 lit. ‘This document deserved a dispatch from Carlos Borrego, two days later, in which the minister agrees with the provisions exposed’

- (96) De seguida os assaltantes fizeram um buraco na parede,
afterwards the burglars made a hole in.the wall
 através do qual conseguiram aceder à caixa forte.
through of.the which managed gain.access:INF to.the strongbox
 ‘Afterwards, the burglars made a hole in the wall through which they managed to gain access to the strongbox.’
- (97) Vi um filme ontem, do qual nunca mais me esquecerei.
saw:1SG a movie yesterday of.the which never more forget:FUT.1SG
 ‘Yesterday I saw a movie, which I will never forget.’
- (98) Foi oferecido um jantar aos congressistas, no qual estava
was offered a dinner to.the congresspersons in.the which was
 presente o Presidente da República.
present the President of.the Republic
 ‘A dinner, at which the President of the Republic was present, was offered to the congresspersons.’
- (99) Ele deu instruções precisas aos trabalhadores, sem as quais o
he gave instructions precise to.the workers without the which the
 plano teria fracassado.
plan have:COND failed
 ‘He gave the workers precise instructions, without which the plan would have failed.’

When extraposition takes place from a preposed focus or *wh*-constituent, other material may intervene between the antecedent and the *o qual*-ARC. See, e.g., example (100) (repeated from (94) above), where the complex *é que* ‘is that’ (which appears in focalized *wh*-questions; see Duarte 2000), the subject and the verb occur as intervening material.

- (100) Que desporto é que tu praticas, sem o qual não sobreviverias?
what sport is that you practice without the which not survive:COND.2SG
 lit. ‘What sport do you practice, without which you would not survive?’

Having seen the restrictions that hold for CEP, let us now turn to the properties of extraposition in earlier stages of Portuguese. For ease of comparison, the same set of properties used for CEP is inspected for earlier stages of Portuguese.

A. *The definiteness effect*

In earlier stages of Portuguese, the extraposition of *o qual*-ARCs was not sensitive to the definiteness effect. Sentences (101)-(104) exemplify extraposed *o qual*-ARCs taking strong noun phrases as antecedents. In (101) and (104), the antecedent is introduced by a definite article, in (102) by a definite article (followed by a possessive), and in (103) by a demonstrative.

- (101) *depos morte da dicta dona Gyralda fficou o dicto*
after death of.the mentioned D. G. stayed the mentioned
herdamento ao dicto Moesteyro de suso nomeado. o qual
land to.the mentioned monastery of above mentioned the which
herdamento est assy como os manios Çinquenta astiís
land is such as the ±untilled.grounds fifty a.
 ‘After Dona Gyralda’s death, the aforementioned monastery got the aforementioned land, which land has, like the untilled grounds, fifty *astiís* [medieval agrarian measure].’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1294)
- (102) *mãdamos dar esta Sentemça Seelada do nosso Seelo*
demand:1PL give:INF this sentence stamped of.the our stamp
ao dicto Conuêto no qual screuemos nosso nome data
to.the mentioned convent in.the which wrote:1PL our name date
 ‘We demand that this sentence with our stamp, in which we wrote our name and date, be given to the aforementioned convent.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1364)
- (103) *Eu Affomso goterrez [...] que este stromento pera ho dicto lujs*
I A. G. that this deed to the mentioned L.
EAnes scripuý no quall meu Sjgnal fjz que tall e%
E. wrote on.the which my sign made:1SG that such is
 lit. ‘I, Affomso Goterrez, who wrote this deed to the aforementioned Lujs EAnes, in which I made my sign, which is as follows % [sign].’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1412)
- (104) *cõ outras confrontaçõescõ que de dereito os ditos bões*
with other limits with that of right the mentioned properties
deuẽ departir; os quaees elas [...]enprazam nouamente aa dicta antonja
should border:INF the which they give again to.the mentioned A.
 ‘(...) with other limits on which the aforementioned properties should border by right. They give again these properties (in emphyteusis) to the aforementioned Antonja.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1520)

B. Restriction on extraposition from embedded positions

In earlier periods of the history of Portuguese, an extraposed *o qual*-ARC can take as an antecedent the object of a preposition. See (105)-(108) below; in (105)-(106), the PP containing the antecedent is the indirect object, and in (107)-(108), it is an oblique constituent.

- (105) Emtestao Eixido que perteeçe aas dictas cassas com A rrua
ends the land that belongs to.the mentioned houses with the street
de ffellgeiras as quaaes cassas dezia o dicto vaasco gonçallvez
of F. the which houses said the mentioned V. G.
 Autor que trazia de ssua Maa Joham gonçallvez
plaintiff that brought of his hand J. G.

‘the land that belongs to the aforementioned houses leads to Street of Ffellgeiras, and the aforementioned Vaasco Gonçallvez, the plaintiff, said that Joham Gonçallvez owned those houses (in emphyteusis).’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1397)

- (106) Joham Lourenço mandou rrecado a sua molher que sse fosse
J. L. sent message tohis wife that SE:CL goes.IMPERF.SUJ
pera elle: da quall ja tiinha hũu filho, que chamavom Alvaro
to him of.the which already had:3SG a son that call:3PL A.

‘Joham Lourenço sent his wife a message saying that she should go back home. He already had a son from her called Alvaro.’

(Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

- (107) que este emprazamento valha e se cumpra
that this emphyteusis be.valid:PRES.SUBJ. and SE:CL carry.out:PRES.SUBJ.
 como se nesta carta contem haa quaL dou minha auctoridade
as SE:CL in.this letter contains to.the which give:1SG my authority

‘I want this emphyteusis to be valid and to be carried out as it is written in this letter, to which I give my authority’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1538)

- (108) os Reys comarcaõs della o mandaraõ visitar por seus
the king neighboring of.it him:CL ordered visit:INF by their
 Embaixadores, & darlhe os parabês da sua
ambassadors and give.to him:CL the congratulations of.the his
capitania, [...], entre os quais veyo hum del Rey dos Batas
appointment among the which came an of.the king of.the B.

‘the neighboring kings sent their ambassadors to visit him and congratulate him on his appointment (with offers to renew the peace and friendship treaties they had

maintained with the king of Portugal). Among those who came was an (ambassador) of the king of the Battak.’

(TYC; Fernão Mendes Pinto, *Peregrinação*; 16th century)

Examples (105)-(108) display extraposition from embedded positions because the antecedent is contained within a PP. However, even more surprising is that the PP may be further embedded in another constituent. See, e.g., (109) below, where extraposition takes place from a PP within a DP.

- (109) os quaes posam penhorar [...] em quaesquer bẽes dos ditos
the which can seize in any belongings of the mentioned
 enprazadores honde quer que achados fforem os quaes nam terã
lessees wherever found were the which not have:FUT
 poder de tolher o dito penhor
power of block:INF the mentioned seizure

‘(so that) they can seize any properties of the aforementioned lessees, wherever they are, and the lessees have no power to block the aforementioned seizure.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1499)

C. Restriction on extraposition from pre-verbal positions

C1. Restriction on extraposition from pre-verbal subjects

In earlier periods of its history, Portuguese allowed for extraposed *o qual*-ARC with a pre-verbal subject as antecedent. Sentences (110)-(113) below attest to the relevant pattern.

- (110) diserom que os dictos logares danboroes e mõte valem todo
said:3PL that the mentioned lands of.A. and hill are.worth all
onze maravedis da boa moeda cõ ho dicto monte os quaes
eleven m. of.the good coin with the mentioned hill the which
 logares danboroes partem cõ erdade darouqua
lands of.A. border with land of.A.

‘They said that the aforementioned lands of Anboroes and the hill are worth eleven *maravedis* [the currency] of the good coin, including the hill. These lands of Anboroes border on land of Arouqua.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1414)

- (111) E mortas as dictas pessoas o dicto enprazamento cõ toda
and died the mentioned people the mentioned emphyteusis with all
 sua benffeitoria ficar liure e desenbargado ao dicto
its improvements become:INF free and ±without.embargo to.the mentioned
Moesteiro cuJo he Ao qual enprazamẽto Eu [...] dou m̃nha autoridade
Moesteiro whose is to.the which emphyteusis I givemy authority
 ‘And the aforementioned people having died, the aforementioned emphyteusis with
 all its improvements remain free for the aforementioned monastery to which it
 belongs. I (...) give my authority to this emphyteusis.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1441)

- (112) e toda a outra cidade era devassa, na quall moravam
and all the other city was opened in.the which lived
 muitas gentes avondadas de grandes rriquezas e bẽes
many people full of big wealth and belongings
 ‘And the rest of the city, in which many rich people lived, could be easily attacked.’

(Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

- (113) onde então o Rey dos Batas se estaua fazẽdo prestes para yr
where then the king of.the B. SE:CL was making ready to go:INF
sobre o Achẽ, o qual tanto que soube do presente & carta que
over the A. the which as.soon.as knew of.the gift and letter that
 lhe eu leuaua do Capitão de Malaca, me mandou receber
to.him.CL I took from.the captain of M. me:CL ordered receive.INF
 pelo Xabandar
by.the X.

‘(from the city of Panaajũ,) where the king of the Battak was busy with preparations
 to attack the Achinese. As soon as he heard about the gift and letter that I was
 taking to him, he sent out Xabandar to welcome me.’

(TYC; Fernão Mendes Pinto, *Peregrinação*; 16th century)

C2. Restriction on extraposition from other pre-verbal constituents

In earlier stages of Portuguese, *o qual*-ARCs may take a preposed focus as antecedent. Example (114) illustrates the point at hand.

- (114) Arato[...] fez tirar em publico as outras pinturas dos tiranos,
A. made remove:INF in public the other paintings of:the tyrants
 mas a de Aristrato determinava de quebrar, a qual pintura
but the of A. determined:3SG of break:INF the which painting
 era nobre á maravilha;
was noble to.the wonder

‘Arato ordered the other paintings of the tyrants to be removed in public, but he was determined to break the painting of Aristrato, which was very impressive.’

(TYC; Francisco de Holanda, *Da Pintura Antiga*; year 1548)

However, in the corpora inspected thus far, no clear occurrence of *o qual*-ARCs with a topic as antecedent was found.³⁰

C3. *New contexts of extraposition from pre-verbal positions*

Earlier stages of Portuguese had a richer clausal structure than CEP, making available more syntactic positions with specific interpretative effects. This is the case with the multiple Specifier positions that were available in the IP domain, which were responsible for the IP scrambling (or medial scrambling) attested to in earlier stages of Portuguese (cf. Chapter 3, Section 4.3). Not surprisingly, extraposition may emerge in this context; see (115) (repeated from (107)).³¹

³⁰ There are some complex structures in which the alleged antecedent of the ARC is introduced into the universe of discourse and then referred to anaphorically by different elements, such as *wh*-constituent *o qual* (*N*). This gives rise to complex sequences, such as the one displayed in (i) below. For ease of reading, the antecedent is marked in italic and the anaphoric links in bold.

- (i) *a preza de Ribell* tem este casall daredor **dela**[...] e **da**
the dam of R. has this hamlet around of:it and of:the
dita **preza** tem no verã daugua **della** hũ dia cada
mentioned dam has:3SG in.the summer of:water of:it one day each
 somana e asy en todo ãno **a quall** he de muyto pouca augua
week and such in all year the which is of very little water

‘The dam of Ribell has this hamlet around it (...); there is water in the dam one day each week during all the year; the dam has very little water.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1545)

Note, however, that although one of the intermediate chain links is a topic in a left dislocation construction (*da dita preza* ‘of the mentioned dam’), the anaphoric link that is nearer the *wh*-constituent *o qual* is *ella* ‘she,’ which is not in a topic position.

³¹ Note that the scrambling of the *nesta carta* ‘in this letter’ in (115) is confirmed by the relative position of this constituent with respect to the verb and the clitic. According to Martins (2002), clitics in clauses with interpolation set the border between left-dislocated/focused constituents and scrambled constituents. Hence, in (115), because *nesta carta* ‘in this letter’ is interpolated (i.e., occurs between the proclitic and the verb), it is necessarily a scrambled constituent. If it occurred to the left of the clitic, it would be a left-dislocated/focused constituent.

- (115) que este empraçamento valha e se cumpra
that this emphyteusis be.valid:PRES.SUBJ. and SE:CL carry.out:PRES.SUBJ.
 como se nesta carta contem haa qual dou minha auctoridade
as SE:CL in.this letter contains to which give:1SG my authority
 ‘I want this emphyteusis to be valid and to be carried out as it is written in this
 letter, to which I give my authority.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1538)

D. *The intervening material*

One of the most remarkable differences between contemporary extraposition and extraposition in earlier stages of Portuguese concerns the number and heaviness of constituents that may intervene between the antecedent and the *o qual*-ARC.

In addition to the PPs and adverbs found in CEP, there is a wide range of constituents that may break the adjacency between a post-verbal antecedent and the verb. These intervening constituents belong to different morphosyntactic categories (e.g., adverbs, prepositional phrases, noun phrases) and play diverse syntactic functions (e.g., object, modifier). To mention a few examples, consider (101), (102), and (105) above, where a PP appears as intervening material, corresponding to an indirect object (in (101), (102)) and to an oblique complement (in (105)). Another possibility is that a noun phrase or nominal conjuncts appear as intervening material, as shown in (116)-(117) below.

- (116) mãdej Ao dicto Martím balastro que apresentasse
ordered:1SG to.the mentioned M. B. that present:IMPERF.SUBJ:3SG
 perante o dicto Tabeliõ e perante domjgos martĩjz
before the mentioned notary and before D. M.
as Testemonhas conteudas no dicto estormẽto Aos quaees eu
the witnesses contained in.the mentioned deed to.the which I
 mãdej que as pergũtassem
ordered that them:CL interrogate:IMPERF.SUBJ:3PL
 ‘I ordered the aforementioned Martím Balastro to present the witnesses referred to
 in the deed before the aforementioned notary and before Domjgos Martĩjz, whom I
 ordered to interrogate them.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1353)

- (117) *testemunhas que Eram presentes llopo martjz orjuez e alluaro*
witnesses that were present L. M. jeweler and A.
gomcalluez barbeJro e bento veloso ao quall llopo martiz a dita
G barber and B. V. to.the which L. M. the mentioned
catarjna periz rrogou que asynasse por sy e por ella
C. P. asked that sign:IMPERF.SUBJ:3SG for him and for her
 ‘witnesses that were present: Llopo Martjz, a jeweler, Alluaro Gomcalluez, a barber, and Bento Velloso; the aforementioned Catarjna Periz asked Llopo Martiz to sign for him and for her.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1522)

The intervening material can also be an even higher constituent, namely, a clause. The three main types of dependent clauses are attested in the intervening position: nominal clauses (see (106) above), adjectival clauses (see (109) above, which involves a free relative clause), and adverbial clauses (see (118), which involves a reason clause). Note that the presence of an internal head in some of these contexts is crucial for the identification of the antecedent; e.g., in (118) the lack of the internal head *pardieiro* ‘ruinous building’ could lead to ambiguity, because it could be unclear whether the antecedent was *moesteiro* ‘monastery’ or *pardieiro* ‘ruinous building’ (see also Section 3.2.1.1).

- (118) *era concertada cõ Joham goncaluez [...] de lhe auer*
wasconcerted:3SG with J. G. of to.him.CL have:INF
denprazar o dicto pardieiro porque asy ho aujã por
of.give:INF the mentioned ruinous.building because such it:CL had:3PL by
seruiço de deus e proueito da dicta dona mjcia e do dicto
service of god and benefit of.the mentioned D. M. and of.the mentioned
seu moesteiro; o quall pardieiro lhe logo enprazarõ
his monastery the which ruinous.building to.him:CL immediately gave:3PL
 ‘She had a deal with Joham Goncaluez to give him (in emphyteusis) the aforementioned ruinous building because they regarded it as service of God and as a benefit to the aforementioned Dona Mjcia and to her monastery; they immediately gave him the ruinous building (in emphyteusis)’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1494)

However, in earlier stages of Portuguese, not only do we find a variety of constituents in the intervening position, but two or more constituents can simultaneously co-occur in this position. Given the multiple combinations of constituents that may appear in the intervening position, it is not possible to exhaustively enumerate them here. To mention but a few possible combinations, see examples (119)-(121) below.

[verb + PP]

- (119) *aqueste prazo fizi. é én testemoyo destas cousas meu
this ±document made:1SG and in testimony of.these things my
sinal pusi en elle o qual sinal tal este.
sign put:1SG on it the which sign such is*

‘I made this document and, as testimony of these things, I put my sign on it; the sign is as follows’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1279)

[adverb + clitic + verb]

- (120) *E o dicto Juiz per sentença defenetiua asy o Julgou
and.the mentioned judge by sentence definitive this.way it:CL judged
da quall sentença o dicto Reeo nō apellou
of.the which sentence the mentioned defendant not appealed*

‘And the aforementioned judge by a definite sentence judged this way (=passed this sentence), against which sentence the aforementioned defendant did not appeal.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1422)

[conjoined non-dependent clause + conjoined non-dependent clause + ...]

- (121) *E logo os ssobreditos lançarõ sortes das ditas partições E
and immediately the mentioned drew.lots of.the mentioned divisions and
acõteçeo A ljonor gonçaluez os oljuaaes e herdades que som ã
happened to L. G. the olive.groves and lands that are in
santarẽ [...]E acõteçeo Ao dito Affonso martijz e a ssa molher
S. and happened to.the mentioned A. M. and to his wife
A herdade de mõte maior [...]as quaes partições os ssobreditos outorgarõ
the land of M. M. the which divisions the mentioned granted*

‘And the aforementioned (people) drew lots of the aforementioned divisions. And the olive groves and the lands located in Santarẽ went to Ljonor Gonçaluez; the land in Mõte Maior went to Affonso Martijz and to his wife. The aforementioned people granted all these divisions.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1370)

As can be noted, there is great variability in the number and heaviness of constituents in the intervening position. Surprisingly, this possibility can lead to a far more radical situation: the units that appear as intervening material may belong not to the sentence level but rather to the discourse level. More specifically, a textual fragment may intervene between the antecedent and the ARC. See, for instance (122), where a document is transcribed before the *o qual*-ARC:

(122) luis dallmeida dom prioll do mosteiro de uillarinho e francisco
L. D. D. prior of.the monastery of U. and F.

fernandez [...] me enviaram dizer per sua pitição o seguinte [...]
F. to.me:CL sent say:INF by their petition the following

a quall pitição vista per mjm mãdei vasar carta de vedoria
the which petition seen by me ordered:1SG make:INF letter of assessment

‘Luis de Allmeida, prior of the monastery of Uillarinho, and Francisco Fernandez sent me a petition saying the following [copy of the petition]. Having seen this petition, I ordered a letter of assessment to be made.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1545)

In (122), the extraposed ARC clearly relates to an antecedent across discourse. Another possibility is that the antecedent and the *o qual*-ARC appear in different utterances. See, for instance, (123) and (124), where the antecedent appears in a first-person direct speech (punctuated with an introductory dash), whereas the ARC appears in the third-person narration.

(123) - Senhor, chegou ally o allmocadê, e parece-me que diz que lhe he necessario de vos fallar llogo amte que amanheça.

‘- Sir, the Moorish captain arrived there and it seems to me he is saying he needs to speak to you immediately before it dawns.’

O qual o comde mamdou que viesse.
the which the count ordered that come:IMPERF.SUBJ:3SG

‘The count ordered the Moorish captain to come.’

(Brocardo 1997; Gomes Eanes de Zurara, *Crónica do Conde D. Pedro de Meneses*, 15th century)

(124) -Ora – disse o comde – ã abasta que vos esto comteis a mÿ soo, mas quer que o digaes assy presente todos estes fidallgos que aquy sã.

‘- Well – said the count – I want you to tell this story not only to me, but also to all the noblemen here.’

Os quaes forã mui comtemtes do que lhe as escuitas disserão
the which were very happy of.the that to.them:CL the eavesdroppers said

‘The noblemen became very happy with what the eavesdroppers said.’

(Brocardo 1997; Gomes Eanes de Zurara, *Crónica do Conde D. Pedro de Meneses*; 15th century)

To sum up, the main conclusions concerning the extraposition of ARCs in Portuguese are as follows.

- *O qual*-ARCs can be extraposed both in CEP and in earlier periods of the history of Portuguese.
- Extraposition of *o qual*-ARCs obeys a number of restrictions in CEP:
 - The antecedent must be a weak noun phrase.
 - The antecedent cannot be the object of a preposition.
 - The antecedent cannot be a pre-verbal subject.
 - Extraposition cannot take place from topics, but can take place from *wh*-constituents and preposed foci.
 - There are restrictions on the constituents that may occur as intervening material. The antecedent cannot be separated from the *o qual*-ARC across the discourse.
- Extraposition of *o qual*-ARCs in earlier stages of Portuguese does not obey the restrictions found in CEP, because:
 - The antecedent can be a strong noun phrase.
 - The antecedent can be the object of a preposition.
 - The antecedent can be a pre-verbal subject.
 - There seems to be no restrictions on the type and number of elements that may appear as intervening material. The antecedent can be separated from the *o qual*-ARC within the sentence or across the discourse.

Finally, by comparing the properties of RRC extraposition in CEP, which are presented in Chapter 3, with the properties of *o qual*-ARC extraposition in CEP, it emerges that both constructions display the same cluster of properties and restrictions. This may suggest that they involve the same syntactic structure and derive extraposition in a similar manner. This hypothesis is considered in more detail in Section 5.2.

As for earlier stages of Portuguese, things are not so straightforward. There are important similarities between RRC extraposition and ARC extraposition, but there are also some divergences, for instance, with respect to the intervening material. In particular, extraposed *o qual*-ARCs can be separated from the antecedent across the discourse, contrary to extraposed RRCs.

3.2.1.3 Pied-piping

The term *pied-piping* refers to a phenomenon whereby a particular movement operation, designated to displace an element X, actually displaces a larger phrase in which X is embedded. Piped-piping occurs in various contexts, e.g., questions, *wh*-exclamatives and relative clauses. The focus of this section is the occurrence of pied-piping in relative clauses, as in (125). When applied to relativization, pied-piping involves the movement to the C-domain not only of the relative noun phrase but also of its surrounding structure (a PP, in (125)).

(125) the man *to whom* I gave the book

In CEP, there are category-specific restrictions with respect to the constituent that can be pied-piped in *o qual*-ARCs. As illustrated in (126) and (127), pied-piping is allowed if the constituent to be raised is a PP or an AdvP.

(126) Recomendo este livro, [PP no qual] podes encontrar toda a
recommend:1SG this book in.the which can:2SG find:INF all the
 informação que procuras.
information that look.for:2SG

‘I recommend this book, in which you can find all the information you are looking for.’

(127) os proprietários da garagem são os subscritores do pedido de
the owners of.the garage are the subscribers of.the request of
 licenciamento que deu entrada na autarquia, [AdvP relativamente ao
licensing that gave entrance in.the council relatively to.the
 qual] a ACIB foi convidada a pronunciar-se.
which the ACIB was invited to pronounce.SE:CL

‘The owners of the garage are the subscribers of the licensing request that was submitted to the council, on which the ACIB was invited to pronounce.’

(CETEMPúblico 1.7 v. 4.0)

However, pied-piping is not allowed if the constituent to be raised is a DP³² or an AP (see (128) and (129), respectively).

³² Brito (1991), Peres and Mória (1995) claim that examples that parallel (128) are grammatical in CEP. However, according to my own judgments and the informants that I consulted, the ungrammaticality of (128) above and (i)-(ii) below is sharp. I will come back to the variation found in the synchronic dimension in Section 6.2.

(to be continued)

- (128) *O Pedro, [_{DP} a mulher do qual] conhecestes ontem, perguntou por ti.
the P. the wife of.the which met:2SG yesterday asked by you
 ‘Pedro, the wife of whom you met yesterday, asked for you.’
- (129) *Vou convidar o João, [_{AP} admirador do qual] eu sempre fui.
go:1SG invite:INF the J. admirer of.the which I always was
 lit. ‘I will invite João, admirer of whom I have always been.’

An apparent exception to the generalization that DPs cannot get pied-piped concerns the contexts involving partitive constructions.³³ In these cases, when the relative pronoun is the complement of the preposition, the whole partitive construction can get pied-piped along with the relative pronoun. This possibility is illustrated in (130):³⁴

- (130) Este acto terá levado o industrial a disparar três tiros,
this act have:FUT led the industrialist to fire:INF three shots
 [_{QP} dois dos quais] terão atingido o filho no abdómen.
two of.the which have:FUT hit the son in.the stomach
 ‘This act might have led the industrialist to fire three shots, two of which might have hit his son in the stomach.’
 (CRPC [jornal_ anotado_RL, AT-0334])

- (i) O João, [_{DP} a amiga do qual] tu conheces, telefonou agora mesmo.
the J. the friend of.the which you know called now right
 ‘João, the friend of whom you know, called right now.’
 (Brito 1991: 132)
- (ii) Foram apresentados vários filmes portugueses muito interessantes, [_{DP} os realizadores dos quais] o Estado deveria apoiar.
were presented various movies Portuguese very interesting the directors of.the which the state should support
 ‘Various interesting Portuguese movies were presented; the state should support the directors of those movies.’
 (Peres and Mória 1995: 278)

³³ A partitive construction typically has the following structure: expression of quantity + of + noun phrase. The complement of the preposition designates a set out of which certain individuals are selected. An example is given in (i):

(i) Two of the girls showed up.

³⁴ In the label associated with the pied-piped constituent, I assume that partitive constructions involve a Quantificational Phrase (QP). See Section 5 for more details.

In (130), the pied-piped constituent is a quantifying phrase headed by a numeral. Another possibility is that it involves a non-numeral quantifier (such as *algumas* ‘some’ in (131)).³⁵

- (131) Nas últimas provas de natação, foram seleccionadas vinte
in.the last competitions of swimming were selected twenty
 crianças, [QP algumas das quais] o Paulo tinha treinado.
children some of.the which the P. had coached
 ‘In the last swimming competitions, there were selected twenty children, some of whom Paulo had coached.’

(Peres and Mória 1995: 278)

An additional restriction on pied-piping concerns the contexts in which the constituent to be moved is a non-finite clause.³⁶ As shown in (132)-(134), infinitival, gerundive, and participial clauses cannot get pied-piped in CEP:³⁷

³⁵ Interestingly, the pied-piping of the partitive construction is also possible in appositions, where no verb occurs. This is illustrated in (i) below:

- (i) Com a sua prisão já são cinco as pessoas detidas no âmbito do
with the his prison already are five the people arrested in.the context of.the
 processo Lasa e Zabala, [QP quatro das quais] comandos e militares
process L. and Z. four of.the which commandoes and military.men
 da guarda.
of.the guard

‘With his prison, there are already five people arrested in the process Lasa and Zabala, four of which (are) commandoes and men of the military guard.’

(CRPC [jornal_ anotado_RL, Ref: J19128/])

This construction may provide evidence for an analysis of appositions as involving a (implicit) clausal structure with a null copula, as proposed by Cardoso and De Vries (2010). See Section 4.1.2 for more details.

³⁶ There is no consensus among authors with respect to the analysis of constructions such as (i) below, taken from Horvath (2007), and originally reported by Nanni and Stillings (1978):

- (i) The elegant parties, [to be invited to one of which] was a privilege ...

Some authors assume that they involve pied-piping (more precisely ‘heavy’ or ‘massive’ pied-piping) (see Heck 2008, Cable 2007); others claim that they do not involve true instances of pied-piping, but rather topicalization (see Emonds 1976, 1979), Webelhuth 1992). Trusswell (forthcoming), when analyzing sentences such as (ii) below, attested in the 16th–19th century English, claims that they do not involve pied-piping but rather base-generation of the clause in a left-adjoined position.

- (ii) This seemed to be done in distrust of the privy council, as if they might stifle his evidence;
 [[which to prevent], he put it in safe hands].
 (BURNETCHA-E3-H,1.2,163.329) [Gilbert Burnet, *History of my own time*, 1683–1713]

Here I assume that these constructions are true instances of pied-piping. I will come back to this topic in Section 5.

[infinitival clauses]³⁸

- (132) *Entregaram-me ontem os documentos, [CP para analisar
delivered:3PL.me:CL yesterday the documents to analyze:INF
 os quais], preciso de pelo menos um mês.
the which need:1sg at least a month

‘They delivered the documents to me yesterday; to analyze them, I need at least a month.’

[gerundive clauses]³⁹

- (133) *Convocámos os responsáveis, [CP reflectindo com os quais]
called:1PL the people.in.charge reflect:GER with the which
 chegámos a uma conclusão.
came:1PL to a conclusion

‘We called the people in charge for a meeting; reflecting with them, we came to a conclusion.’

³⁷ In labels associated with the pied-piped clauses, I assume, following Lobo (2003), that gerundive and participial clauses involve a CP projection. The same analysis is adopted for infinitival clauses, under the assumption that the connective introducing the infinitival clause (as *para* in (132)) occupies the C-position. These are the criteria for Portuguese examples reported here; for the examples taken from other authors, I will adopt the original bracketing and labels (if present).

³⁸ Peres and Mória (1995) claim that examples that parallel (132) are grammatical in CEP (see (i) below). However, for me and for the people I consulted, the ungrammaticality of (132) and (i) is sharp. I will return to the variation found in the synchronic dimension towards the end of the chapter (Section 6.2).

- (i) Foram descobertas novas provas, [CP para analisar as quais]
were found new proofs to analyze:INF the which
 o tribunal precisa de muito tempo.
the court needs of much time

‘New proofs were found; to analyse them, the court will need much time.’

(Peres and Mória 1995: 279)

³⁹ Peres and Mória (1995) claim that examples that parallel (133)a are grammatical in CEP (see (i) below). However, according to my own judgments and the people that I consulted, the ungrammaticality of (133) and (i) is sharp. I will come back to this in Section 6.2

- (i) Foram descobertas novas provas, [CP considerando as quais] o tribunal mudou
were found new proofs consider:GER the which the court changed
 de opinião.
of opinion

‘New proofs were found; considering these proofs, the court changed its opinion.’

(Peres and Mória 1995: 279)

[participial clauses]⁴⁰

- (134) *A direção vai apresentar os resultados, [CP conhecidos os quais]
the management go.FUT present:INF the results known the which
 algumas soluções estratégicas serão discutidas.
some solutions strategic be:FUT discussed

‘The management will present the results; the results known, some of the strategic solutions will be discussed.’

To sum up, the restrictions reviewed up to this point are presented in Table 3:

Table 3. Restrictions on pied-piping (Portuguese *o qual*-ARCs)

	DPs	APs	CPs	AdvPs	PPs	Partitive constructions
CEP	-	-	-	+	+	+

Interestingly, these restrictions do not universally hold but are subject to cross-linguistic variation. In fact, whereas some languages are very strict about the category of the pied-piped constituents, other languages seem to be much less constrained, allowing generalized pied-piping in ARCs. This is reported, e.g., by Cinque (2008), for Italian *il quale*-ARCs; see (135):

⁴⁰ A word is in order regarding the apparent pied-piping of participial clauses. There is a special context in which the construction seems to be possible in CEP, as shown in (i) and (ii) below. However, it is only possible with the verb *terminar* ‘to expire’, as in (i), or with a synonym of it, as *findar* in (ii). A change of the verb blocks its viability, as indicated in (134). This fact can be explained if we assume that the apparent pied-piping of participial clauses is not a productive syntactic structure in CEP (as opposed to the situation in earlier stages of Portuguese). Hence, the sequence *terminado/findo o qual* ‘expired which’ behaves as a fixed expression, involving specific lexical items and not admitting the occurrence of other verbs.

- (i) Será definido um período, [CP terminado o qual] ninguém poderá
be:FUT established a period ended the which nobody can:FUT
 reclamar.
complain:INF
 ‘A period will be defined; this period ended, nobody can complain.’
 (Peres and Mória 1995: 279)

- (ii) Os analistas estimam que estas negociações [...] se prolonguem por um prazo
the analysts estimate that these negotiations SE:CL extend by a period
 entre 12 e 18 meses, [CP findo o qual] deverá haver um acordo.
between 12 and 18 months ended the which should:FUT have:INF a deal
 ‘The analysts estimate that these negotiations will be extended for a period of 12-18 months; this period ended, there must be a deal.’
 (CETEMPúblico 1.7 v. 4.0)

- (135) a. Inviterò anche Giorgio, [PP del quale] avete certamente sentito parlare.
 ‘I will invite also G., of whom/that you have certainly heard.’
 b. Inviterò anche Giorgio, [DP il fratello del quale] è uno dei nostri più cari amici.
 ‘I will invite also G., the brother of whom is one of our dearest friends.’
 c. Inviterò anche Giorgio, [AP affezionato al quale] per altro non sono.
 ‘I will also invite G., fond of whom at any rate I am not.’
 d. Inviterò anche Giorgio, [CP liberarmi del quale] non mi è proprio possibile.
 ‘I will invite also G., to get rid of whom is really not possible for me.’
 e. Inviterò anche Giorgio, [AdvP diversamente dal quale] io non serbo rancore.
 ‘I will invite also G., differently from whom I bear no grudge.’
 (Cinque 2008: 101)

Similar possibilities of pied-piping are reported for English. Heck (2008: 168) shows that English ARCs allow for the pied-piping of PPs, APs, DPs, and clausal constituents (see (136)). Fabb (1990: 64) also reports the pied-piping of DPs ((137)a) and partitive constructions ((137)b):

- (136) a. Egbert, [PP to whom] you were talking only yesterday, . . .
 b. ? this earth quake, [AP affected by which] the area was , . . .
 c. the royal family, [DP pictures of whom] are permanently on sale, . . .
 d. Egbert, [α to hire whom] would be a real scoop, . . .
- (137) a. The man, [the mother of whom] I met yesterday, is a French speaker.
 b. The men, [some of whom] I like, arrived yesterday.

Interestingly, the restrictions on pied-piping are also subject to variation in the diachronic dimension. When we compare the properties of pied-piping in CEP with the ones in earlier periods of Portuguese, the differences are remarkable. The general scenario is that earlier periods of Portuguese pattern with contemporary English and Italian in allowing generalized pied-piping.

To be more concrete, *o qual*-ARCs in earlier stages of Portuguese allow pied-piping of PPs, partitive constructions, and AdvPs, just like their contemporary counterparts. This is illustrated in (138)-(140).

- (138) Reçebemos de Giral dominguiz [...] Cem libras de dinheiros
receive:1PL of G. D. one.hundred l. of currency

portugaeses [PP polos quaes] lhj nós vendemos [...]
Portuguese by.the which to.them:CL we sold

‘We received from Giral Dominguez one hundred *libras* [the currency] of the Portuguese currency, for which we sold them [two houses that we have in the aforementioned village] ...’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1353)

- (139) nos matou logo seis homens, [QP hum dos quais] foy
us:CLkilled outright six men one of.the which was

Diogo Vaz Coutinho filho do Capitão mòr
D. V. C. son of.the admiral

‘killed six of our men outright, one of whom was Diogo Vaz Coutinho, the admiral’s son.’

(TYC; Fernão Mendes Pinto, *Peregrinação*; 16th century)

- (140) taes são os importantes objectos, [AdvPrelativamente aos quaes]
such are the important topics relatively to.the which

devem os factos ser escolhidos, e detalhados.
should the facts be:INF selected and detailed

‘The facts that should be selected are detailed, taking into account the important topics just mentioned.’

(CdP; António Leite Ribeiro, *Theoria*; 1818)

However, historical Portuguese, contrary to CEP, allows pied-piping of DPs and clausal constituents.⁴¹ Examples (141)-(145) illustrate pied-piping of DPs; notice that in these examples the gap corresponds either to the subject (as in (141)-(143), (145)-(146)) or to the direct object position (as in (144)).

[pied-piping of DPs]

- (141) receb| hua procuraço do Abade san Joane da pendorada
received:1SG one letter.of.attorney of.the abbot S. J of.the P.

e do Conuêto [DP o teor da qual] atal e de ueruo. a ueruo
and.of.the convent the tenor of.the which such is of word to word

‘I received one letter of attorney from the abbot of San Joane of Pendorada and from the convent; the tenor of the letter is the following, word for word.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1278)

⁴¹ In the corpora of historical Portuguese inspected, pied-piping of APs is not attested. For this reason, in this section I mainly focus on the pied-piping of DPs and clausal constituents.

- (142) como mais larguamente consta dapeguação que aqui
as more extensively is.reported of.the.±possession.letter that here
 mandei treladar de verbo a verbo [DP o trelado da qual]
ordered:1SG copy:INF of word to word the copy of.the which
 he o seguinte
is the following
 ‘As it is more extensively reported in the possession letter that I ordered to be copied here, word for word; the copy of the letter is as follows.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1545)
- (143) A composição dos edefícios consta de symetria,
the composition of.the buildings consists of symmetry
 [DP a razão da qual] os deligentes arquitetos hão de entender.
the reason of.the whichthe diligent architects must understand:INF
 ‘The composition of the buildings encompasses symmetry; the diligent architects must understand its logic.’
 (TYC; Francisco de Holanda, *Da Pintura Antiga*; year 1548)
- (144) e fe os particulares devem fer amparados na fua menor
and if the common.people should be:INF protected in.the their minor
 idade, quanto mais o deve fer hum Rey; [DP a boa criação
age let alone it:CL should be:INF a king the good education
 do qual] fe dirige ao bem de muitos, ao serviço de Deos,
of.the which se:CL directs to.the good of many to.the service of God
 e à protecção da Religião Catholica;
and to.the protection of.the Religion Catholic
 ‘And if the common people should be protected when they are underage, let alone the king, whose good education benefits not only the well-being of many, but also the service of God and the protection of the Catholic Religion.’
 (TYC; Manuel dos Santos, *História Sebástica*; year 1735)
- (145) Agora falarei nos requisitos para a intelligência da
now talk:FUT.1SG in.the requirements for the understanding of.the
 dita língua, [DP a falta dos quais] não se deve
mentioned language the lack of.the:PL which:PL not SE:CL should
 contar entre os menores abusos;
number:INF among the minor abuses
 ‘I will now talk about the requirements for the understanding of the aforementioned language; the lack of these requirements must not be numbered among the minor abuses.’
 (TYC; Luís António Verney, *Verdadeiro Método de Estudar*; year 1746)

Over the course of its history, Portuguese also allowed for pied-piping of non-finite CPs. By way of illustration, see examples below, which involve participial clauses ((146)-(152)), gerundive clauses ((153)-(157)), and infinitival clauses ((158)-(160)).

[pied-piping of participial clauses]

- (146) E sobre o negado ffoy ffilhada Enqueriçõ [CP A qual vista per m̃j]
and about the ±denial was made examination the which seen by me
 Julgey que o dito prioll prouaua quanto Auõdaua
judged: ISG that the mentioned prior proved all.that was.sufficient

‘And an examination was made about the denial; the examination seen, I judged that the aforementioned prior has proved conclusively (that he was right).’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1379)

- (147) Dona Thareyía martiis dona da Chelas mostrou hũa carta de nosso senhor El Rey e sselada do seu seelo pendête da qual o tẽhor atal he. [...]

‘Dona Thareyía Martiis, Dona of Chelas, showed a letter from the King, stamped with his hanging stamp; the tenor of the letter is as follows: [transcription of the letter]’

[CP A qual carta mostrada e leuda] a dita Thareyía martiis
the which letter showed and read the mentioned T. M.

comprou tres courelas de vinhas en Barathoio per outoridade da
bought three lands of vineyards in B. by authority of.the

dita carta
mentioned letter

‘This letter shown and read, the aforementioned Thareyía Martiis bought three vineyards in Barathoio by the authority of the aforementioned letter.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1317)

- (148) Eu Nicollao de ffreitas tabaliam del Rey na dicta villa de guimarãães que esta procuraçom per mādado e outorgamêto da dicta Maria fernandez screpuj e aquy meu synal fiz que tal. he.

‘I, Nicollao de Ffreitas, notary of the king in the aforementioned village of Guimarães, who wrote this letter of attorney by mandate of the aforementioned Maria Fernandez, here I put my sign, which is as follows:’

[CP A quall presentada] os dictos procuradores do dicto
the which presented the mentioned attorneys of.the mentioned

Moesteiro disserom que antre elles Era preito
monastery said that between them was legal.dispute

‘This letter of attorney shown, the attorneys of the aforementioned monastery said that they were involved in a legal dispute.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1411)

- (149) E com os ingleses viinha o alferes do duque d'Allancastro
andwith the English came the flag.bearer of.the duke o.A.
 [...], que tragia sua bandeira; [_{CP} a quall tendida na batalha],
that brought his flag the which stretched in.the battle
 braadavom os ingleses todos
yelled the English all
 ‘And the flag bearer of the duke of Allancastro, carrying his flag, came among the English knights. The flag stretched out, the English knights started yelling out: (“Castella and Leon are for the king Joham of Castella ...”).’
 (Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)
- (150) O Capitão mór entendendo quão importante cousa esta era, lhe aceitou a promessa, & lhe concedeo de nouo as pazes,
 ‘The captain, fully aware of the gravity of the situation, accepted her promise and renewed the peace,’
 [_{CP} as quais juradas aly logo & confirmadas de ambas as
the which sworn there immediately and confirmed of both the
 partes com as cerimonias costumadas entre aquelles Gentios], a
parties with the ceremonies used among those heathen.people the
 Raynha buscou todos os meynos possiveis para cumprir a sua palaura
queen tried all the means possible to keep:INF the her word
 ‘The peace sworn to there and then and confirmed by both parties in accordance with the local ceremonies, the queen tried in every way possible to keep her word.’
 (TYC; Fernão Mendes Pinto, *Peregrinação*; 16th century)
- (151) Depois de faber ler, e escrever, ouviu ElRey Grammatica,
after of know:INF read:INF and write INF heard the.king grammar
 [_{CP} na qual instruido] paffou ao estudo de Autores Latinos
in.the which instructed moved to.the study of authors Latin
 ‘After learning to read and write, the king learned grammar; instructed in grammar, he started studying the Latin authors.’
 (TYC; Manuel dos Santos, *História Sebástica*; year 1735)

- (152) e no anno de 699. foy mandado ouvir Artes no Real
and in.the year of 699 was:3SG sent listen:INF Arts in.the Royal
 Mosteiro de Santa Maria de Ceiça, e Theologia no nosso
Monastery of S. M. d. C. and Theology in.the our
 Collegio de S. Bernardo de Coimbra; [CP acabados os quaes Curfos],
College of S. B. d. C. ended the which courses
 fe graduou de Doutor Theologo
SE.CL graduated:3SG of D. T.

‘In the year of 1699, he was sent to attend Arts in the Royal Monastery of Santa Maria de Ceiça and Theology in our College of S. Bernardo de Coimbra; when these courses ended, he graduated as *Doutor Theologo* [± Doctor in Theology].’

(TYC; Manuel dos Santos, *História Sebástica*; year 1735)

[pied-piping of gerundive clauses]

- (153) enprazou a afonso periz de lestosa e a sua molher marja anes
gave:3SG to A. P. from L. and to his wife M. A.
 e a hũu filho ou filha dantre ambos [CP o qual hi nom
and to a son or daughter of both the which there not
 avendo] a hũa pessoa qual ho postumeiro que deles mais
have:GER to a person which the last that of.them more
 viuer nomear
live:FUT.SUBJ appoint:INF

‘He gave (it in emphyteusis) to Afonso Periz from Lestosa and to his wife Marja Anes and to a son or daughter of them; if they do not have any child, he gives it to a person whom the last of them to die will appoint.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1489)

- (154) avendo primeiro salvo-conducto de dona Johana, rainha entom d’aquella
have:GER first safe-conduct of D. J. queen then of.that
 provencia; [CP na quall estando per pouco tempo], Pero Bernaldez,
province in.the which be:GER by short time P. B.
 cossairo d’Aragom, chegou hi com gallees armadas
corsair from.A, arrived there with galleys armed

‘(They went to the city of Neapolli,) having a safe-conduct given by Dona Johana, then queen of that province. Being there for a short time, Pero Baernaldez, a corsair from Aragon, arrived there with armed galleys.’

(Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

- (155) estando hi em cabido scilicet o Reueremdo senhor lujs dalmeida prioll do dito mosteiro e manuell JorJe conjgo do dito mosteiro
 ‘The Reverend Sir Lujs dalmeida, prior of the aforementioned monastery, and Manuell JorJe, canon of the aforementioned monastery, being there gathered for the chapter’

[_{CP} o quall prioll e conjgo estando no dito cabido Juntos
the which prior and canon be:GER in.the mentioned chapter together

per som de campam tamgida como tem de seus costumes] o
by sound of bell rung as has of its costumes the

dito prioll dise que
mentioned prior said that

‘The prior and the canon being gathered in the aforementioned chapter at the sound of the bell ring, as usual, the aforementioned prior said that (...)’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1534)

- (156) a suessaõ delRey D. Joaõ III. filho primogenito delRey D. Manoel, acabou em ElRey D. Sebastiaõ seu neto; e tornando aos filhos do mesmo Rey D. Manoel, naõ achou varaõ vivo, mais que o Cardeal D. Henrique,

‘the succession of the king D. Joaõ III, firstborn son of the king D. Manoel, ended at the king D. Sebastiaõ, his grandson; among king D. Manoel’s children, there was no living son besides Cardinal D. Henrique

[_{CP} o qual morrendo sem suessaõ, e sem irmaõ, ou irmãa,
the which die:GER without succession and without brother or sister

a quem deixasse o Reyno], necessariamente havia de hir
to whom leave:IMPERF.SUBJ the kingdom necessarily would go:INF

a hum de muitos sobrinhos seus
to one of many nephews his

‘If the Cardinal D. Henrique dies without succession and without a sister or a brother to whom to leave the kingdom, the succession will necessarily pass to one of his many nephews.’

(TYC; Manuel da Costa, *Arte de Furtar*; 17th century)

- (157) e me disse como se lhe pedissem juramento,
and to.me:CL said:3SG that if to.her:CL ask:IMPERF.SUBJ.3PL oath

o daria na verdade deste cazo; [_{CP} o qual relatando ao
it.CL make:COND in.the truth of.this case the which tell:GER to.the

mesmo Padre], lhe respondeo, que...
same Priest to.her:CL replied:3SG that

‘And she told me that, if she was asked to make an oath, she would make it in the name of the truth of this case. Telling this case to the same Priest, he told her that (...)’

(TYC; Maria do Céu, *Vida e Morte de Madre Helena da Cruz*; year 1721)

[pied-piping of infinitival clauses]

(158) no Latim há três Gerúndios, um em Di, outro em Do, outro em Dum,
in.the Latin has three gerunds one in d. other in d. other in d.

[_{CP} para explicar os quais] se serve a língua Portuguesa da
to express:INF the which SE:CL uses the language Portuguese of.the

voz do Infinitivo com alguma preposição
voice of.the infinitive with some preposition

‘In Latin, there are three gerunds ending in –di, –do and –dum. The Portuguese language uses the infinitive with a preposition to express these gerunds.’

(TYC; Jeronimo Contador de Argote, *Regras da Língua Portuguesa*; 1725)

(159) se descobriu em mim culpas, [_{CP} para remir as quais] me
if found:3SG in me faults to cleanse:INF the which to.me:CL

marcou esta penitência, bem vê com que resignação eu a
gave:3SG this penance well see:3SG with what resignation I it:CL

aceito
accept

‘If you found faults in me, and gave me this penance to cleanse me from them, you can well see with how much resignation I accept it.’

(CdP; Júlio Dinis, *Os Fidalgos da Casa Mourisca*; 1871)

(160) Burlado até na esperança de colher às mãos o audaz primo do
deceived even in.the hope of find:INF the bold cousin of.the

senhor de Cresconhe, Egas, que ele supunha em Guimarães, e [_{CP} para
S. d. C. E. that he presumed in G. and to

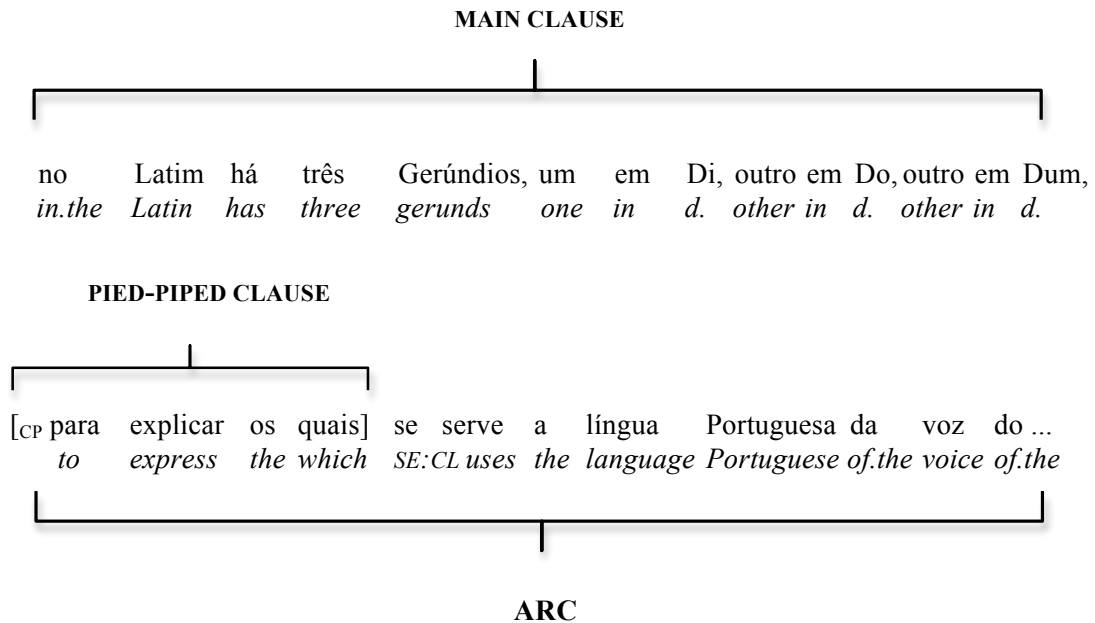
achar o qual] tinham sido vãs as mais severas pesquisas
find:INF the which had been vain the more severe researches

‘Deceived even in the hope of finding the bold cousin of the Senhor de Cresconhe, who he presumed to be in Guimarães, and whom he made in vain several attempts to find (...)’

(CdP; Alexandre Herculano, *O Bobo*; 1843)

A closer inspection of the above examples reveals that they involve a rather complex syntactic environment containing at least three different clauses: the clause that contains the antecedent (*the main clause*), the embedded clause (*the ARC*), and the *pied-piped clause* contained within the ARC. See the schematic representation in (161):

(161)



There is a lot going on in these examples, but there are three aspects that I would like to highlight: (i) the chronology; (ii) the position of the relative pronoun; and (iii) the clause types involved.

A. *The chronology*

In the corpora inspected in this research, the pied-piping of non-finite clausal constituents is attested in earlier periods of Portuguese. However, it is not evenly distributed across non-finite clauses but is found almost exclusively in participial and gerundive clauses. For instance, in the texts edited by Martins (2001), pied-piping of participial and gerundive clauses is attested, but pied-piping of infinitival clauses is not. I found it in other corpora, but only in latter periods (see examples given in (158)-(160)). Further evidence from larger corpora is needed to assess whether this is real or corresponds to an accidental gap.

B. *The position of the relative pronoun*

Within the pied-piped clause, the relative pronoun can occur in its base position or can undergo internal movement to the CP domain. In (152) and in (158)-(160) above,

the relative pronoun stays in its base position within the pied-piped clause, whereas in (146)-(157), it undergoes internal movement to the CP domain. The latter case corresponds to the so-called *internal wh-movement* or *secondary wh-movement* (see Bianchi 1999, Smits 1988, Cable 2007, Heck 2008, Trusswell forthcoming).⁴²

In the earliest texts inspected, the internal *wh*-movement is predominant. In the corpus edited by Martins (2001), all participial and gerundive clauses involve internal *wh*-movement. The occurrence of the *wh*-pronoun in its base position is attested in latter texts, as shown by the examples in (152) and in (158)-(160) above, involving a participial clause and infinitival clauses, respectively.

In the data inspected thus far, internal *wh*-movement also displays the following properties.

- It may involve pied-piping of a PP (see (151) and (154)).
- The relative pronoun may occur with an additional internal head, as in (147) and (152); this additional internal head can be a conjoined phrase, as in (72) (cf. Section 3.2.1.1).
- It may involve across-the-board extraction of the relative pronoun out of coordinate participial clauses (see (147) and (150)).

⁴² Trusswell (forthcoming) reports the existence of internal *wh*-movement in earlier stages of English, as shown in (i) and (ii) below. Bianchi (1999: 143), in turn, reports the possibility of internal *wh*-movement in earlier stages of Italian, as illustrated in (iii) below.

(i) a sarmon, somthing better then that in the morninge: [_{CP} which ended, with all Ceremones], I returned to my lodginge.
(Lady Margaret Hoby, *Diary*, 1599–1601)

(ii) Mr Hoby, my Mother, and my selfe, went to visitt some freindes [_{CP} who, beinge not at home], we returned
(Lady Margaret Hoby, *Diary*, 1599–1601)

(iii) Non si meravigli dunque alcuno se lunga è la digressione della mia scusa, ma, sì come necessaria, la sua lunghezza paziente sostenga. [_{CP} La quale proseguendo], dico che...
lit. ‘hence nobody be astonished if the digression of my justification is long, but, as (it is) necessary, its length with patience tolerate. Which (digression) continuing, I say that...’
(Dante, *Convivio*, I, X)

C. *The clause types involved*

In the examples provided above, the clausal pied-piping involves non-finite clauses. However, pied-piping finite adverbial clauses is also attested in earlier stages of Portuguese. By way of illustration, see (162)-(164) below.⁴³

- (162) E emtom a molher disse ao segundo marido que matasse
and then the wife said to.the second husband that kill:IMPERF.SUBJ.3SG
 o primeiro marido e que ella teria a elle por seu marido.
the first husband and that she have:COND A him as her husband
 [CP O quall como nom quisesse fazer tamanha traiçom], a
the which since not want:IMPERF.SUBJ make:INF such betrayal the
 dita molher matou ao dito primeiro marido em no celeiro.
mentioned wife killed to.the mentioned first husband in in.the barn

‘And then the wife told to the second husband to kill the first husband and that, in that case, she would take him to be her husband. Because he did not want to make such a betrayal, the aforementioned wife killed the first husband in the barn.’

(CdP; *Crónica da Ordem dos Frades Menores* (1209-1285); 15th century-manuscript)

⁴³ Trusswell (forthcoming) reports similar constructions for 16th–19th century English. Some examples are given below:

- (i) receive then this Draught [[with which when thou art refresh’d], thou mayst more strongly proceed to other Matters which yet remain].
 (BOETHPR-E3-H,201. 466) [Richard Preston (tr.), *Of the Consolation of Philosophy*, 1695]
- (ii) I make a square, that is G.H.K.L, [[In which square if I draw crosse lines frome one side to the other, according to the diuisions of the line G.H], then will it appear plaine, that the theoreme doth affirme].
 (RECORD-E1-H,2.F1R.312) [Robert Record, *The Path-way to Knowl- edg, Containing the First Principles of Geometrie*, 1551]
- (iii) but not so easie work found Ethelfrid against another part of Britans that stood in arms, [[whom though at last he overthrew], yet with slaugh- ter nigh as great to his own souldiers].
 (MILTON-E3-H,X,149.76) [John Milton, *The history of Britain, that part especially now call’d England*, 1670]

- (163) Admite além disso a nossa língua com grande elegância,
admits besides that the our language with great elegance
 e particular graça as metáforas, [CP as quais como se podem aplicar
and particular grace the metaphors the which because SE:CL can:3PL apply
 a tantas cousas], fica uma mesma sentença servindo a muitos sentidos
to so many things stays a same sentence serve:GER to many meanings
 ‘With great elegance and particular grace, our language also admits metaphors.
 Because metaphors can apply to many things, the same sentence can have many
 meanings.’

(CdP; Manuel Severim de Faria, *Discursos Vários Políticos*, 1631)

- (164) nem tenham diante dos-olhos estas circunstâncias: [CP as quais
nor have:PRES.SUBJ.3PL before of.the.eyes these circumstances the which
 se eu nam tivesse executado], totalmente me-faltaria
if I not have:PRES.SUBJ.3PL executed totally to.me:CL.lack:COND
 aquela benevolencia, que certamente me-mostram, os que examinam as
that benevolence that certainly to.me:CL-show the that examine the
 minhas asoens
my actions

‘(...) nor have before the eyes these circumstances. If I had not executed them, I
 would lack that benevolence that the ones who examine my actions show.’

(TYC; Luís António Verney, *Verdadeiro Método de Estudar*; year 1746)

The number of attested cases is small; up to now, I have only found 14 tokens in the *Corpus do Português* (CdP) and 1 in the *Tycho Brahe Corpus* (TYC). All the examples found involve internal *wh*-movement, and the relative pronoun always precedes the connective introducing the adverbial clause.

Observe furthermore that the adjacency between the relative pronoun and the connective introducing the adverbial finite clause and the fact that the relative pronoun does not play any function within the main clause clearly show that the relative pronoun is not extracted from the adverbial clause, but rather internally moved to the left periphery.⁴⁴

⁴⁴ Peres and Mória (1995: 287) report a construction from a 16th century Portuguese text that, in my opinion, is similar to the ones discussed here (see (i)). However, they claim that it involves extraction of the relative pronoun from the subordinate clause. I depart from their analysis (and interpretation) because, as clearly shown by the translation in (i), the relative pronoun does not play any function in the main clause.

(i) Esta é a ditosa pátria minha amada,
this is the delightful homeland my beloved
 À qual se o céu me dá que eu sem perigo
to.the which if the heaven to.me:CL gives that I without danger
 (to be continued)

To summarize, the contrasts between CEP and earlier stages of Portuguese with respect to pied-piping are displayed in Table 4.

Table 4 . Restrictions on pied-piping (CEP vs. earlier stages of Portuguese)

	DPs	APs	CPs	AdvPs	PPs	Partitive constructions
CEP	-	-	-	+	+	+
Earlier stages of Port.	+	?	+	+	+	+

3.2.2. Additional properties in contrast

3.2.2.1 Clausal antecedent

In CEP, *o qual*-ARCs cannot take clausal antecedents (see (165)).⁴⁵

- (165) *O João chegou a horas, **o qual** muito me surpreendeu.
the J. arrived on time the which very.much me:CL surprised
 ‘João arrived on time, which surprised me very much.’

The only relativizers that can introduce clausal antecedents are *o que* lit. ‘the that’ and *que* lit. ‘that’. This is illustrated in (166)-(167) below:⁴⁶

Torne com esta empresa já acabada
return with this war already ended
 Acabe-se esta luz ali comigo.
end:PRES.SUBJ.3SG.SE:CL this light there with.me
 ‘This is my own beloved delightful land / to which if heaven accord me safe / return, with this war ended, / there may the light of life leave me.’
 (L. de Camões, *Os Lusíadas*, Book III, 21, 16th century, cited in Peres and Mória 1995: 287)

⁴⁵ Interestingly, the use of *o qual* with a clausal antecedent is found in the *Syntactic Annotated Corpus of Portuguese Dialects* (CORDIAL-SIN), (see (i)). However, such a sentence is out of the standard variety.

- (i) e era tudo pregado com cravetes, **o qual** desta forma é mais
and was everything nailed with ±metal.slivers the which of.this way is more
 fácil, com menos despesa
easy with less expense
 ‘And everything was nailed with metal slivers, which was the easiest and less expensive way to do it.’

(CORDIAL-SIN - Graciosa - GRC12)

- (166) O João chegou a horas, **o que** muito me surpreendeu.
the J. arrived on time the that very.much me:CL surprised
 ‘João arrived on time, which surprised me very much.’
- (167) O João faltou à reunião, **que** era o que eu devia ter feito.
the J. missed to.the meeting that was the that I should have done
 ‘João missed the meeting, which was what I should have done.’

Earlier stages of Portuguese behave differently in this respect. As examples (168)-(172) show, *o qual*-ARCs can take clausal antecedents; in this case, the ARC is introduced by an invariable *o qual*.

- (168) e se obrygou de pagar os dytos duzentos Reaes e dous fframguãos e a dyta galinha de fforo despoys do ffaleçimento da dyta molher do dito alluaro fernandez em cada hũu Ano pelo dito dia de natall
 ‘and he committed himself to pay every year, on Christmas day, the aforementioned two thousand *reaes* [the currency], two cockerels, and one hen as rent, after the death of Alluaro Fernandez’s wife.’
 pera **o qual** loguo obrygou seus bões
for the which immediately pawned:3SG his belongings
 ‘for which he pawned his belongings’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)
- (169) e tantas lagrimas e gritos e taaes pallavras diziam, que nom havia homem que as ouvisse que nom ouvesse d’ellas compaixom e doo;
 ‘And (the women) cried so many tears, let out so many screams, and said such words that all the man that could hear them felt compassion and pity for them.’
o quall tanto esforço fez cobrar aos que dentro eram que
the which such strength made gather:INF to.the that inside were that
 rrijamente aderçarom pera aquell logar em que os mouros stavom
sturdily went:3PL to that place in that the Moors were
 ‘These facts made the men that were inside the city gather so much strength that they sturdily went to the place where the Moors were (and fought with them with great courage).’
 (Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

⁴⁶ Note that, in (167), the ARC *que era ... feito* ‘which was ... done’ accidentally contains a free relative clause. For further examples of ARCs with clausal antecedents in CEP, see Brito and Duarte (2003: 674-675).

- (170) se assentou com este mercador por esta maneyra, que o padre lhe desse duzentos taeis, que saõ trezentos cruzados da nossa moeda, & que auia de yr daly da nao ate a cidade sempre cos olhos tapados porque se caso fosse que por elle ser estrangeyro, a justiça entendesse nelle, como estaua certo que auia de ser, & pondoo a tormento lhe dissessem que confessasse quem o aly trouxeraõ elle o não soubesse dizer, nem conhecesse quem o aly trouxera, porque se temia que se fosse descuberto lhe mãdassem por isso cortar a cabeça,

‘They agreed with this merchant as follows: the father was to give him two hundred *taeis* [the currency] – which is worth three hundred *cruzados* [the currency] in our money – to take him from where the *nao* was anchored all the way to the city with his eyes blindfolded, so that in case – because he was a foreigner – the police got hold of him, as was bound to happen, and tried to make him confess under torture who had brought him there, he would not be able to tell them nor recognize the one who had brought him there, for fear that if he were discovered they would have his head chopped off’

o qual o padre aceitou com todos estes partidos
the which the father accepted with all these conditions

‘The father accepted this agreement with all these conditions.’

(CdP; Fernão Mendes Pinto, *Peregrinação*; 16th century)

- (171) E depois de feito Deos e home deitou outro pregão sobre o mesmo caso dizendo aos discípulos: nam convém a vós outros saber o que está por vir, porque isso pertence à onnipotência do padre.

‘And after making God and the man, He announced to his Disciples: it is not in your interest to know what will happen in the future because that belongs to the Father’s omnipotence.’

Polo qual mui maravilhado estou dos letrados mostrarem-se
by.the which very surprised am of.the lettered.men be:INF.SE:CL

tam bravos contra tam hórridos pregões
so furious against such horrible notices

‘For this reason, I am very surprised with the lettered man being so furious with such horrible notices.’

(Camões 1999; Gil Vicente, *Tormenta*; year 1531)

- (172) acrescentando ele suplicante [...] que por obedecer levaria os papéis e apontamentos que tinha feito no estado em que estivessem como lhe era mandado.

‘He, supplicant, added that he would bring them the papers and the notes he had made, exactly how they were, just as was ordered of him.’

Em cumprimento **do qual** foi ele suplicante ao Santo Oficio em
in observance of.the which went he supplicant to.the S. O. on

14 do dito mês
14th of.the mentioned month

‘In the observance of this he, supplicant, went to the Santo Oficio [tribunal of the Inquisition] on July 14th.’

(Muhana 1995; *Os Autos do Processo de Vieira na Inquisição*; year 1665)

Further examples making the same point are given in (173)-(175). These examples contrast with (168)-(172) above in that an additional internal head follows the relative pronoun. Recall from Section 3.2.1.1 that, in the contexts of ARCs with a clausal antecedent, the additional internal head is typically a general abstract noun such as *cousa* ‘thing’ (as in (173)) or *razom* ‘reason’ (as in (174)-(175)).

- (173) E dou por firme e por estauil pera todo sempre totalas cousas que
and give:1SG as firm and as steady forever all.the things that
 forem feytas e procuradas per este meu procurador [...]. No
be:FUT.SUBJ made and represented by this my attorney in.the
 testemoyo **da qual cousa** roquey Domĩgos esteueiz tabelliom das
testimony of.the which thing asked D. E notary of.the
 alcaceuas que mi fezesse ende esta procuraçom.
A. that to.me make:IMPERF.SUBJ:3SG of.it this letter.of.attorney
 ‘I confirm whatever my attorney will do; and as a testimony of it, I asked Domĩgos Esteueiz, notary of Alcaceuas, to make this letter of attorney.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1291)
- (174) nom declarar que os ditos cassaaes fforõ cõprados dos dinheiros
not declare that the mentioned hamlets were bought of.the money
 do dito mosteiro polla **quall Razom** de derejto perteçem
of.the mentioned monastery by.the which reason by right belong:3PL
 e perteçyam ao dito mosteiro
and belonged to.the mentioned monastery
 ‘to not declare that the aforementioned hamlets were bought with the money of the aforementioned monastery, for which reason they belong and belonged by right to the aforementioned monastery.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1437)
- (175) Bem sabe el-rrei dom Henrique, meu irmão e amigo, como el-rrei
well knows the.king D. H. my brother and friend that the.king
 de Graada tem tomados navios e averes e gentes cativas de minha
of G. has taken navies and goods and people captive of my
 terra, por **a quall rrazom** eu ei com ell guerra
land by the which reason I have with him war
 ‘The king Dom Henrique, my brother and friend, knows very well that the king of Graada has my navies, goods, and people in his possession, for which reason I am at war with him.’
 (Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

There are some contemporary languages that pattern with earlier stages of Portuguese in this respect. Cinque (2008) reports that Italian *il quale*-ARCs may take clausal

antecedents; see (176), where the relative pronoun is followed by an internal head. The same point can be made for English. As shown in (177), the relativizer *which* can take a clausal antecedent, optionally followed by an internal head.

- (176) Carlo lavora troppo poco. **La qual cosa** verrà certamente notata
 ‘Carlo works too little. Which thing will certainly be noticed.’
 (Cinque 1988, cited in Cinque 2008: 106)

- (177) a. Little Joey snatched the letter away, **which** infuriated his sister.
 b. They are said to have taught a baboon to write, **which claim** has immediately been ridiculed by most scholars.
 (Smits 1988: 287)

3.2.2.2 Split antecedents

In CEP, *o qual*-ARCs cannot have split antecedents. This impossibility is illustrated in (178):⁴⁷

⁴⁷ Brito (1991) asserts that examples like (178) are grammatical in CEP (see (i) and (ii), taken from Brito 1991: 133). However, for me, the ungrammaticality of (178), (i), and (ii) is sharp.

- (i) Como a Maria_i não se estava a dar muito bem com o António_j,
as the M. not SE:CL was A:PREP get:INF along with the A.
 os quais_{i+j} de facto não têm muito em comum, ele resolveu aceitar o emprego
the which in fact not have much in common he decided accept:INF the job
 em Lisboa.
in L.
 ‘As Maria was not getting along with António, who in fact do not have much in common,
 he decided to accept the job in Lisboa.’
- (ii) Como a Maria_i veio ao Porto com o Henrique_j, com os quais_{i+j} eu
as the M. came to.the P. with the H. with the which I
 já não estava há muito tempo, fui jantar com eles.
already not was has much time went:1SG dine:INF with them
 ‘As Maria came with Henrique to Porto, whom I was not with for long time, I had dinner
 with them.’

- (178) *Se o Carlos_i já não gosta da Maria_j, os quais_{i+j} nunca se
if the C. already not likes of.the M. the which never SE:CL
 deram nada bem, então acho que não vale a pena continuarem
got along then think:1SG that not is.worth stay:INF.3PL
 juntos.
together
 ‘If Carlos no longer loves Maria, who never got along with each other, then I think they should not stay together.’

The same, however, is not true of earlier stages of Portuguese. As shown in (179)-(182) below, *o qual*-ARCs with split antecedents are documented in the history of Portuguese. In the corpora inspected in this research, two options are available: (i) the ARC may be introduced by the plural form of the relative pronoun, as shown in (179)-(180); (ii) the relative pronoun may be followed by an additional internal head, which may be a conjoined noun phrase, as in (181)-(182) below.⁴⁸

- (179) Julgo per sentença que este êprazamento valha e se
judge:1SG by sentence that this emphyteusis be.valid:PRES.SUBJ. and SE:CL
 cūpra como se nesta carta_i cõtẽ, e no vltimo
fulfill:PRES.SUBJ.3SG as SE:CL in.this letter contains and in.the last
 consentimento do dicto prior e convento_j faz menção Aos
approval of.the mentioned prior and convent makes mention to.the:PL
 quaes_{i+j} dou e hey por dada mynha autorydade
which:PL give:1SG and have:1SG by given my authority
 ‘I order this contract to be valid and fulfilled, as it is contained in this letter and as it is mentioned in the last approval of the aforementioned prior and convent; I give my authority to this letter and to the approval (of the mentioned prior and convent).’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1540)
- (180) E por séer mays firme esta carta_i seelamos dos nossos séelos e
and by be:INF more firm this letter stamp:1PL of.the our stamps and
 outra tal_j. das quaes_{i+j} deue téer o dicto ffernã yohanes hũa
other such of.the which should have:INF the mentioned F. Y. one
 e a dicta dona outra.
and the mentioned D. other
 ‘And, to be irrevocable, we stamp this letter and a duplicate of it, one of which goes to Ffernã Yohanes and the other to the aforementioned Dona.’
 (Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1273)

⁴⁸ Recall from Section 3.2.1.1 that in the latter contexts there is typically first conjunct agreement for ϕ -features between the relative pronoun and the noun in the first conjunct.

- (181) per a dicta soprioresa ffuj logo apresentada hũa carta
by the mentioned vice-prioress was immediately shown a letter
 dEl Rey_i [...] na quall ffazya mençõ antre as outras cousas
from.the king in.the which made mention among the other things
 que Em Ella Era conthyudo hũa clausulla_j [...] a quall carta_i E
that in it was contained a clause the which letter and
 clausulla_j Em Ella conthyuda asy amostrada [...]
 clause in it contained this.way shown

‘A letter from the king was immediately shown by the aforementioned vice-prioress, in which it was mentioned, among other things, that a clause was contained in it. After showing the letter and the clause contained in it (the aforementioned vice-prioress said that...).’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1429)

- (182) E pagam de cada casal ou courella dezasete alqueires de pam_i [...]Item
and pay:3PLof each hamlet or land seventeen bushels of bread also
 pagam mais em dinheiro_j [...] mjl & trezentos & trimta
pay:3PLmore in money one.thousand and three.hundred and thirty
 Reaaes.O qual pam_i & dinheiro_i sam obrigados repartirem antre ssy
r. the which bread and money are.3PL forced share:INF.3PL between them

‘And they pay for each hamlet and land seventeen bushels of bread. They also pay in money one thousand three hundred and thirty *reais* [the currency]. They are forced to share the bread and money between them.’

(CdP; *Forais Manuelinos*; year 1496-1520)

Interestingly, the same pattern is reported for other languages. Cinque (2008) points out that Italian *il quale*-ARCs can take split antecedents, as in (183); Arnold (2007) reports the same behavior for English ARCs, as illustrated in (184).

- (183) Se Carlo_i non amava più Anna_j, i quali_{i,j} d'altra parte non si erano mai voluti veramente bene, una ragione c'era.

‘If Carlo was no longer in love with Anna, who at any rate never really loved each other, there was a motive.’

(Cinque 2008: 104)

- (184) Kim likes muffins_i, but Sandy prefers scones_j, which_{i+j} they eat with jam.

(Arnold 2007: 274)

3.2.2.3 Coordination of the *wh*-pronoun with another DP

In contemporary *o qual*-ARCs, coordinating the *wh*-pronoun with another DP results in ungrammaticality.

- (185) *O presidente elogiou o João, [o qual e a sua mulher]
the president praised the J. the which and the his wife
 têm desenvolvido um óptimo trabalho naquela instituição.
have developed a great work in.that institution
 ‘The president praised João; he and his wife have been developing a great work in that institution.’

However, such coordination is possible in earlier stages of Portuguese; see (186)-(187) below.

- (186) filho de hum seu filho chamado per nome dom Henrique, o qual era lidimo e, segundo conta a cronica, era o primeiro filho que o dito rei de Ungria ouve.

‘son of one of his sons called Dom Henrique, who was legitimate and, according to the chronicle, was the first son that the king of Ungria had.’

[O qual dom Henrique e hum seu tio, irmão de sua madre],[...] se
the whichD. H. and a his uncle brother of his moher SE:CL

vierão a Castela aa corte, donde o dito rei dom Affonso estava
came:3PL toC. to.the court where the mentioned kingD. A. was

‘The aforementioned Dom Henrique and an uncle of his, his mother’s brother, came to Castela, to the court, where the king Dom Affonso was.’

(CdP; *Cronica de Portugal*; year 1419)

- (187) [As quais razões e outras muitas que o padre-mestre Francisco lhe
the which reasons and other many that the Father.Master F. to.him.CL

dava], o rei gentio de Bungo ouviu e entendendo de maneira que
gave the king heathen of B. heard and understood of way that

deu em pródigo com os pobres.
became prodigal with the poor

‘The heathen king of Bungo heard these and many other reasons that the Father Master Francisco gave him and became prodigal, helping the poor people.’

(CdP; João de Lucena, *Historia da vida do Padre S. Francisco Xavier*; year 1600)

Similarly, *o qual* and a DP can occur as the object of prepositions within conjoined PPs. See examples displayed in (188) and (189) below.

- (188) pedindo-lhe usasse livremente dos poderes que
ask:GER.to.him:CL use: IMPERF.SUBJ.3SG freely of.the powers that
 trazia de Sua Santidade, [com os quais e com sua doutrina e
had from His Holiness with the which and with his doctrine and
 exemplo] estava mui certo havia de fazer grandes serviços a Deus
example was very sure had of make great services to God
 lit. ‘asking him to freely use the power that he was given from His Holiness; with
 which and with his doctrine and example, he would certainly make great things to
 serve God.’
 (CdP; João de Lucena, *Historia da vida do Padre S. Francisco Xavier*; year 1600)

- (189) em que aponta as conveniências de se fazer a impressão
in that points.out the advantages of SE:CL make:INF the printing
 antes em Madrid que em Lisboa, [com as quais e com o
rather in M. than in L. with the which and with the
 partido que oferece] eu me conformei
payment that proposes I myself:CL resigned
 lit. ‘in which he points out the advantages of making the printing in Madrid rather
 than in Lisboa, to which and to the payment he proposes I resigned myself.’
 (CdP; António Vieira, *Cartas*; 1626-1692)

Note that although there is a tendency for the occurrence of an additional internal head in these contexts (see (186)-(188)), the head internal need not necessarily be spelled out. This is illustrated in (189) above, where the *wh*-pronoun *as quais* lit. ‘the:FEM.PL which:FEM.PL’ occurs per se within the first PP.

Once again, Italian *il quale*-ARCs and English ARCs pattern with earlier stages of Portuguese in allowing coordination of the relativizer with a DP. This possibility is reported in Cinque (2008: 108, 115):

- (190) a. ?Gianni e Mario, [le rispettive consorti e i quali] non si erano mai potuti soffrire...
 ‘Gianni and Mario, the respective wives and whom had never been able to stand
 each other, ...’
 b. Gianni e Mario, [fra le rispettive consorti e i quali] non c’era mai stato un grande
 affiatamento,...
 ‘Gianni and Mario, between their respective wives and whom there never was a
 real understanding,...’
 (191) He recalled the name of the solicitor, [between whom and himself] there had been
 occasional correspondence.

(Jespersen 1949: 191)

3.2.2.4 Illocutionary force

The distinction between contemporary and historical *o qual*-ARCs in terms of illocutionary force is formally reflected in the fact that the two construction types behave differently with respect to the syntactic markers of illocutionary force, such as the different basic clause types (declarative, interrogative, and imperative).

Let us first look at CEP. Contemporary *o qual*-ARCs can be declarative, even if the matrix is interrogative or imperative. This is illustrated in (192) and (193) below. In these examples, the ARC is declarative and the matrix is interrogative and imperative, respectively:

- (192) Será que o João, com o qual pudemos sempre contar,
be:FUT.3SG that the J. with the which could:3PL always count:INF
 estará disponível desta vez?
be: FUT.3SG available of.this time
 ‘Will João, who we have always counted on, be available this time?’

- (193) Telefona aos teus pais, os quais estarão certamente
phone: :IMP.2SG to.the your parents the which be:FUT.3PL certainly
 disponíveis para te ajudar!
available to you:CL help:INF
 ‘Phone your parents, who will certainly be available to help you!’

The reverse does not hold, however. *O qual*-ARCs in CEP do not allow any clause types beyond the declarative, as shown by the unacceptability of the interrogative in (194) and the imperative in (195).

- (194) *O único que te apoiou foi o João, ao qual já
the only that you:CL supported was the J. to.the which already
 agradeceste devidamente por tudo o que te fez?
thanked:2SG properly by all the which to.you:CL made:3SG
 ‘The only person who supported you was John; have you already thanked him properly for everything he made for you?’

- (195) *Acabou de chegar o João, ao qual vai já oferecer
has.just of arrive:INF the J. to.the which go:IMP.2SG now offer:INF
 uma bebida!
a drink
 ‘João has just arrived; offer him a drink now!’

Earlier stages of Portuguese allow the equivalent of (195), however. See (196)-(198) below, where the matrix is declarative and the relative clause has imperative force.⁴⁹

- (196) e posto que hūuas pallavras sejam contra as outras, e todas em soma contradigam aa verdade, nós porém creemos que suas erradas rrazoões nom foi per malicia dos autores mas per inorancia da verdade,

‘and although some information is contradictory and clearly far from the truth, we nevertheless believe that the mistakes result not from the author’s malice but rather from ignorance of the truth.’

a quall sabee que foi d’esta guisa.
the which know:IMP:2PL that was of.this way

‘Know that the truth was as follows.’

(Macchi 1975; Fernão Lopes, *Crónica de D. Fernando*; 15th century)

- (197) ho prior do moesteiro de uilarinho do dicto arcebispado me enviou dizer que sentindo por proueito do dicto mosteiro queria enprazar como de feito enprazou a quebrada de penellas que o dicto mosteiro tem sita na freguesia de sam frausto a fernam correa escudeiro morador em a villa de guimarães e a sua molher mjcía fferrnandez [...]

‘the prior of the monastery of Uilarinho of the aforementioned archbishopric ordered me to say that, for the aforementioned monastery’s benefit, he wanted to give in emphyteusis – as in fact he did – the land of Penellas, which the monastery has in the parish of Sam Frausto, to Fernam Correa, squire, inhabitant of Guimarães, and to his wife, Mjcia Fferrnandez’

Os quaeesa Jam e pessuam a dicta
the which have:PRES.SUBJ and possess:PRES.SUBJ the mentioned

quebrada Com todas suas casas vinhas sotos
land with all its houses vineyards thickets

‘Fernam Correa and Mjcia Fferrnandez have and possess the aforementioned land with all its houses, vineyards, thickets, (...)’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1534)

⁴⁹ Note that in Portuguese, imperative sentences use the imperative mood for the second person. For other grammatical persons and for every negative imperative sentence, the subjunctive is used.

- (198) Com o teor do qual mandei passar esta carta testemunhável ao dito Bento Henriques, à qual mando que seja dada tanta fé e autoridade, em juízo e fora dele, e onde quer que fôr apresentada, quanta por direito se lhe deve dar.

‘I ordered to send this letter, with the content of the aforementioned document, to Bento Henriques; I order that all the faith and authority – recognized by law within our jurisdiction or outside of it, and wherever it will be shown – to be given to this letter.’

O qual uns e outros assim cumpram e al
the which some and others as.such obey:PRES.SUBJ:3PL and another
 não façais
not make:PRES.SUBJ:2PL

‘All the intervenient persons obey this and do not make it differently.’

(Pereira 1987; *Doc. para a História da Inquisição em Portugal*; 1578)

Furthermore, both the matrix and the ARC may have non-declarative force. See (199) below, where the matrix and the ARC have imperative force.

- (199) E ponha ã corporall posissom della o dicto prioll de vilarinho. ou seu certo procurador scilicet per pedra terra telha altar ljuros calezes chaves vestimētas E per outros quaeesquer hornamentos e bēes que em ella forem achados,

‘And give the aforementioned prior of Vilarinho or his attorney possession of the church, with its land, tile, altar, books, chalices, keys, vestments, and any other adornments and belongings that might be found there,’

dos quaees lhe seJa facto Enuentairo segundo Costume
of.the which to.it.CL be:PRES.SUBJ:3SG made inventory as usual

‘Make an inventory of all these things, as used.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1450)

Other languages pattern with earlier stages of Portuguese in allowing the ARC to have a non-declarative illocutionary force. This is reported by Cinque (2008: 102, 103) for Italian *il quale*-ARCs (see (200)-(201)) and English ARCs (see (202)-(203)). Examples below display ARCs with interrogative force (see (200), (202)) and imperative force (see (201) and (203)).

- (200) L’unico che potrebbe è tuo padre, il quale potrà, credi, perdonarci per quello che abbiamo fatto?

‘The only one who could is your father, by whom will we ever be forgiven, you think, for what we have done?’

- (201) Ci sono poi i Rossi, per i quali, ti prego, cerca di trovare una sistemazione!

‘There are then the Rossi’s, for whom please try to find an accommodation!’

- (202) a. There is then our father, by whom will we ever be forgiven for what we have done?
 b. It may clear up, in which case would you mind hanging the washing out?
 c. She may have her parents with her, in which case where am I going to sleep?
 (Huddleston, Pullum, and Peterson 2002: 1061, for *b.* and *c.* examples)
- (203) a. Please accept my check for \$3.69, which find enclosed!
 (Martin 1972: 5)
 b. He said he'd show a few slides towards the end of his talk, at which point please remember to dim the lights!
 (Huddleston, Pullum, and Peterson 2002: 1061)
 c. My friend, who God forbid you should ever meet,...
- (John Lyons, cited in Werth 1974, fn.4)

3.2.2.5 Coordinator

In CEP, *o qual*-ARCs cannot be preceded by a coordinator. The contrast given in (204) illustrates this point: sentence *a.* becomes ungrammatical if the *o qual*-ARC is preceded by the coordinator *e* 'and.'

- (204) a. Foi detectado um erro grave na prova de química, para o
was detected an error serious in.the exam of chemistry to the
 qual ainda não foi apresentada nenhuma explicação.
which yet not was provided none explanation
 'A serious error was detected in the chemistry exam, for which no explanation has been provided yet.'
- b. *Foi detectado um erro grave na prova de Química, e para
was detected an error serious in.the exam of Chemistry and to
 o qual ainda não foi apresentada nenhuma explicação.
the which yet not was given none explanation

However, this was possible in earlier stages of Portuguese. As shown in (205)-(207) below, *o qual*-ARCs could be preceded by the coordinator *e/et* 'and'.

- (205) *customarõ dauer e ouuerom no dicto Monsteiro bõa raçom e
used:3PL of.have and had:3PL in.the mentioned monastery good ration and
mãtjímêto de pam aluo boroa. carne e vïho e o qual mãtjímêto
provisions of bread white corn.bread meat and wine and the which provisions
os Priores [...] auiã e som theudos de dar ao dicto conuêto
the priors had and are compelled of give:INFto.the mentioned convent*
‘They used to have and had in the aforementioned monastery good ration and provisions of white bread, corn bread, meat, and wine, which provisions the priors had and were compelled to give to the aforementioned convent.’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1364)

- (206) *me outorgo por biẽ pagada deste dicto herdamêto &
me:CL declare:1SG by well paid of.this mentioned land and
cousas que aqui en esta carta som en ella escriptos & cõteudos. Et
things that here in this letter are in it written and contained and
o qual herdamêto & cousas sobredictas hã jazença no logar que
the which land and things aforesaid have ±location in.the place that
chama de Curraes
call:3PL of C.*

‘I declare that I was paid for the aforementioned land and things referred in this letter. The aforementioned land and things are located in a place that people call Curraes.’

(CIPM; *História do Galego-Português* [HGP 105]; year 1289)

- (207) *E nos [...] outorgamos sse formos contra este prazo en todo
and we declare if go:FUT.SUBJ.3PL against this ±document in all
ou en parte que peytemos aos sobreditos [...] cen mr
or in part that pay:PRES.SUBJ.1PL to.the aforementioned one.hundred m. old
uelhos de pãa. E a qual pãa pagada ou nõ, este prazo e
old of penalty and the which penalty paid or not this ±document and
as cousas que neel sson cõtehudafiquen en ssa firmydõe.
the things that in.it are contained stay:PRES.SUBJ.3PL in its firmness*

‘And we declare that, if we go against this document, we must pay to the aforementioned (persons) one hundred *maravedis* [the currency] as penalty. And independently of the payment of this penalty, this document and the things contained in it should be valid.’

(CIPM; *História do Galego-Português* [HGP 143]; year 1313)

Interestingly, the same pattern may occasionally be found in English, as shown in (208)-(211) below. In (208)-(210) the ARCs are introduced by the coordinator *and*, whereas in (211) it is introduced by the coordinator *but*.

- (208) and the new capitol is here, of course, too, built five years before she was born, and which she has always associated with learning Latin.

(COCA; Willie Morris, *Miss Eudora*, 1995)

- (209) " I'm inept -- how do you like that word? -- at everything but my work and getting to and from it, " was how he liked to phrase it whenever she asked him to do a chore, and which she said was his alibi for doing nothing around the house

(COCA; Stephen Dixon, *Interstate 2*, paragraph 1, 1993)

- (210) Well, Pickering gave me an earful, not directed at me, and which I much enjoyed.

(COCA; Emile Capouya, *The rising of the moon*, 1995)

- (211) Eventually I found one willing to sell me a camel at what would have been an exorbitant price under ordinary circumstances , but which I was all too willing to pay.

(COCA; Ted Chaing, *The Merchant and the Alchemist's Gate*, 2007)

In these examples, a coordinator may show up (i) when the nominal antecedent is already modified by a constituent – as in (211), where the antecedent *price* is modified by *exorbitant*;⁵⁰ and (ii) when no such a modifier is present – as in (209), where the antecedent is non-nominal.

3.3. Summary

In Sections 3.2.1 and 3.2.2, I have shown that *o qual*-ARCs in CEP behave differently from *o qual*-ARCs in earlier stages of Portuguese with respect to a number of

⁵⁰ Beatrice Santorini (p.c.) reports to me that earlier stages of English behave in a similar fashion. See, for instance, (i)-(iii) below, from Earlier Modern English (1500–1700), where the ARC is preceded by a coordinator (and the antecedent is modified by one or more adjectives):

- (i) and hopes the Pope will not any longer delay gratifying him in so reasonable a request, and which his Majesty desires so earnestly from his Holinesse.

(PPCEME, SPENCER-1680-E3-H,3.4,315.33)

- (ii) but the greater power and working of wine may be spied more plainly in colde and withered bodies, and wherein is lesse naturall heat, as in olde men, and in such as are amended of their sicknesse.

(PPCEME, TURNER-E1-P1,E2R.198)

- (iii) That had been too wild and extravagant a supposition, and which it is likely in those days had never entered into any mans mind.

(PPCEME , TILLOTS-C-E3-P2,457.52)

syntactic properties. The contrasting properties analyzed thus far are summarized in Table 5. Here a minus ‘-’ indicates that *o qual*-ARCs may display the relevant properties and a plus ‘+’ indicates that they may not.

Table 5. Properties distinguishing *o qual*-ARCs in earlier stages of Portuguese and in CEP

	<i>o qual</i> -ARCs CEP	<i>o qual</i> -ARCs in earlier stages of Portuguese
Internal head	-	+
Generalized extraposition	-	+
Generalized pied-piping	-	+
Clausal antecedents	-	+
Split antecedents	-	+
Coordination of the <i>wh</i> -pronoun with another DP	-	+
Non-declarative illocutionary force	-	+
Coordinator preceding the <i>wh</i> -pronoun	-	+

Additionally, I have demonstrated that Italian *il quale*-ARCs and English ARCs pattern with *o qual*-ARCs in earlier stages of Portuguese with respect to the same syntactic properties. This is summarized in Table 6:

Table 6. Properties of Italian *il quale*-ARCs and English ARCs

	Italian <i>il quale</i> -ARCs	English ARCs
Internal head	+	+
Generalized extraposition	+	+
Generalized pied-piping	+	+
Clausal antecedents	+	+
Split antecedents	+	+
Coordination of the <i>wh</i> -pronoun with another DP	+	+
Non-declarative illocutionary force	+	+
Coordinator preceding the <i>wh</i> -pronoun	⁵¹	+

Cross-linguistic and language internal contrasts found in the synchronic and diachronic dimension are discussed in more detail towards the end of the chapter (see Section 6).

⁵¹ At this point, I do not have information on whether Italian allows for a coordinator (preceding the *wh*-pronoun *il quale*).

4. An analysis of the two types of ARCs

Having reviewed the contrasting properties of *o qual*-ARCs in CEP and in earlier stages of Portuguese, let us now consider the impact that such contrasts have in the theoretical analysis of ARCs. At this point, it is possible to follow two possible directions.

Hypothesis I. The first is to assume that there is only one syntactic structure that derives *o qual*-ARCs both in earlier stages of Portuguese and in CEP.

Hypothesis II. The second approach consists of saying that there are two different syntactic structures that generate *o qual*-ARCs: one for *o qual*-ARCs in earlier stages of Portuguese and another for *o qual*-ARCs in CEP.

I submit that Hypothesis I is untenable because it would leave the eight contrasting properties discussed in Section 3.2 either unexplained or explained by ad hoc and language-specific stipulations. It is no coincidence, I think, that these properties arise from the comparative approach carried out through the preceding sections. Such contrasts simply follow from the dual approach to *o qual*-ARCs advocated here, according to which *o qual*-ARCs in earlier stages of Portuguese are derived from specifying coordination (see Section 4.1), whereas *o qual*-ARCs in CEP are derived from raising (see Section 4.2). Both analyses may be qualified as constituency analyses, in the terms discussed in Section 2.1.2 above.

4.1. *o qual*-ARCs in earlier stages of Portuguese

The main claim of this section is that *o qual*-ARCs in earlier stages of Portuguese involve the specifying coordination analysis proposed by De Vries (2006b). The basic tenets of this proposal are outlined in Section 4.1.1. Then, in Section 4.1.2, I will show that not only ARCs but a wide range of appositive constructions can be derived from the coordinate-style account, as proposed by Cardoso and De Vries (2010).

4.1.1. The specifying coordination analysis

A. A coordination account of apposition

De Vries (2006b) argues that appositional constructions involve a coordinating relationship between the anchor and the apposition. More recently, Heringa (2007, 2009) has made the same claim.⁵²

One of the arguments for treating appositional constructions in terms of coordination comes from the fact that a coordinator may occasionally show up in appositions. This is illustrated in (212), taken from Heringa (2007: 69):

- (212) a. The United States of America, or America for short...
 b. You could cut the atmosphere with a knife, and a blunt knife at that.
 c. John is interested in science, but especially linguistics.
 (*a.* and *b.* examples are from Quirk et al. 1985:1311-12)

The connection between the two DPs may also be made explicit by phrases such as *that is (to say)*, *namely*, or *for example*. What these elements have in common is that they are specifying phrases, i.e., elements that introduce a DP that adds information to the anchor.

For more arguments on treating appositional constructions in terms of coordination, see De Vries (2006b) and Heringa (2007, 2009). For now, it is sufficient to point out that this analysis implies that there are (at least) four semantic types of coordination. See (213):

- (213) a. the Netherlands and Belgium (additive)
 b. the Netherlands or Belgium (disjunctive)
 c. not the Netherlands, but Belgium (adversative)
 d. the Netherlands, or Holland (specifying)
 (Heringa 2009)

The main difference between the traditional types of coordination and the type involved in appositional constructions is semantic. Whereas the conjuncts denote two

⁵² Other authors have highlighted the parallel between coordination and appositive constructions. For example, Quirk et al. (1985:1301/2) state that: “Apposition resembles coordination in that not only do coordinate constructions also involve the linking of units of the same rank, but the central coordinators *and* and *or* may themselves occasionally be used as explicit markers of apposition.”

different entities in examples *a.* to *c.*, they refer to one and the same entity in example *d.*. In the latter case, the second conjunct specifies, i.e., gives more information about, the anchor. It is precisely this relationship that is dubbed *specifying coordination*.

In syntactic terms, however, the different types of coordination involve the same structure. Following Kayne (1994) and Johannessen (1998), De Vries (2006b) represents coordination as $[_{CoP} XP [_{Co'} YP]]$. Appositions involve a coordination phrase (CoP), with a coordinator as the head and with the two conjuncts as the Specifier and complement of this head, as demonstrated in (214):⁵³

- (214) $[_{CoP} [_{DP} \text{ anchor}] [_{Co} [_{DP} \text{ apposition}]]]$
 e.g. *John,* *a nice guy*

The coordinative head is often phonologically null, but, as already mentioned, it can also be made overt by a specifying phrase.

B. *A coordination account of ARCs*

According to De Vries' (2006b), the ARC also involves specifying coordination. More concretely, the ARC is treated as a complex apposition that is coordinated with the antecedent:

- (215) $[_{CoP} [_{DP} \text{ antecedent}] [_{Co} [_{DP} D [_{CP} \text{ ARC}]]]]$

The abstract coordinator involved is semantically specialized; it constitutes a relationship of specification between the two DP conjuncts. Within the second conjunct, the relative clause is the complement of D; this corresponds to a raising-style configuration of a full relative construction (a DP containing a relative clause). The surprising aspect of this analysis is, therefore, that an ARC is, in fact, an RRC in apposition to the overt antecedent. Given that the second conjunct normally

⁵³ De Vries (2006b) symbolically represents the coordinative head by an ampersand plus a colon, i.e., '&:'. Here, I will simply make use of the more general denotation Co for coordinative head. (cf. Chapter 3, Section 2.1.3, fn. 2).

does not contain an overt antecedent itself, the RRC comprises a semi-free relative construction. Thus, (216)a is analyzed roughly in the same manner as (216)b:

- (216) a. Jack, who lives in Paris
 b. Jack: person who lives in Paris

The structural representation of (216)a is displayed in (217):

- (217) [_{CoP} DP Co [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} t_k]]]]
 e.g. *Jack* , \emptyset \emptyset *who* \emptyset *lives in Paris*

The determiner heading the second conjunct together with the raised abstract head NP can be considered a pronoun that behaves in a similar way to an E-type pronoun requiring co-reference with some objects (Evans 1980).⁵⁴ This means that it is able to pick up an appropriate antecedent without requiring any particular syntactic configuration, similarly to how definite anaphoric or demonstrative pronouns refer to a phrase across discourse. This freedom is restrained, however, by the semantics of the specifying coordination, which require that the second conjunct give additional information to the phrase in the first conjunct. This prevents the null pronoun from taking as its antecedent a phrase outside the first conjunct.

C. Some properties of ARCs derived

The coordinate-style account offers a natural explanation for the interpretative properties of ARCs. For the sake of illustration, let us consider three of these properties in some detail: (i) the scope of the determiner, (ii) the lack of reconstruction effects, and (iii) the opacity for binding. For a detailed presentation of how the specifying coordination analysis derives other properties of ARCs, see De Vries (2006b).

⁵⁴ See Del Gobbo (2008) for a recent discussion of the E-type character of the referential link between (regular) appositive relative clauses and the antecedent.

I. *Scope of the determiner*

ARCs, in contrast to RRCs, are not within the scope of the determiner/quantifier that belongs to the antecedent; see (218):

- (218) a. the students that passed the exam [RRC]
 b. the students, who passed the exam [ARC]

In (218)a, the determiner *the* takes scope over the noun and the relative clause in the RRC; from the interpretative point of view, it implies that there is a group of students that did not pass the exam. In contrast, in (218)b, the determiner *the* takes scope over only the noun; consequently, it refers to all of the students regardless of whether they passed the exam.

Now compare the representations in (219)a and (219)b, which involve the raising analysis and the specifying coordination analysis, respectively. For the sake of clarity, the visible antecedent is underlined in both structures.

- (219) a. [_{DP} D [_{CP} NP ... relative IP]] [RRC]
 b. [_{CoP} [_{DP} D NP] [Co [_{DP} ... relative IP]]]] [ARC]

Clearly, the relevant D in (219)a c-commands the head NP and the relative clause, but the antecedent D (and also N) in the appositive configuration in (219)b does not c-command the relative clause because of the coordination structure; both are embedded inside the first conjunct. Following the standard assumption that scope is dependent on c-command, the scopal difference between RRCs and ARCs is derived.

II. *Reconstruction effects*

RRCs and ARCs behave differently with respect to reconstruction effects. For instance, some idiomatic expressions allow the relativization of the idiomatic object in RRCs but not in ARCs:

- (220) a. The horrible face that Harry made at Peter scared him. [RRC]
 b. *The horrible face, which Harry made at Peter, scared him. [ARC]
 (Emonds 1979: 233)

In the raising analysis, the head is base-generated inside the relative clause; hence, it can be reconstructed in that position. Following the assumption that the constituents of the idiomatic expression must be adjacent in the LF representation, the grammaticality of (220)a follows. In the specifying coordination analysis, although there is raising of the (abstract, pronominal) head NP within the second conjunct, the visible antecedent is base-generated in the first conjunct. There is no movement chain between the antecedent and the position of the gap inside the relative CP; thus, the constituents of the idiomatic expression cannot reconstruct in a relative-clause internal position.

III. *Opacity for binding*

Pronoun-binding by a quantifier is possible if the pronoun surfaces in an RRC but not in an ARC; see (221):

- (221) a. I gave every assistant_i who loved his_i uniform a new one. [RRC]
 b. *I gave every assistant_i, who loved his_i uniform, a new one. [ARC]
 (Emonds 1979: 236)

In the raising analysis, the grammaticality of (221) is derived from the fact that the antecedent c-commands the pronoun inside the relative clause. In the specifying coordination analysis, such a relationship cannot be established because second conjuncts are *invisible* for the higher context in terms of c-command (see De Vries 2005 and Chapter 3, Section 6.3, fn. 51).⁵⁵

⁵⁵ In the contexts involving a pronoun that might potentially be bound by material higher up in the matrix (as in (i)b), the same reasoning applies, i.e., the pronoun cannot be bound because second conjuncts are shielded from c-command relationships.

- (i) a. Everyone_i spoke about the museum that he_i had visited. [RRC]
 b. *Everyone_i spoke about the Millennium Dome, which he_i had visited. [ARC]
 (De Vries 2006b: 256)

4.1.2. The expansion of the specifying coordination analysis

One of the most promising aspects of the specifying coordination analysis of ARCs is that it may account not only for the syntax of ARCs but also for a wide range of appositive constructions. As shown in Cardoso and De Vries (2010), differences lie primarily in the choice of which elements are spelled out and in their respective positions.

In what follows, I will summarize the main findings of Cardoso and De Vries' (2010) paper. The presentation focuses mainly on two issues: (i) the predictions of the specifying coordination analysis, and (ii) the correlation between these predictions and the existence of a wide range of appositional constructions.

A. Predictions of the specifying coordination analysis

It is uncontroversially accepted that RRCs exhibit variation in the choice of elements that can be spelled out in the CP domain. See (222):

- (222) a. the girl which I saw
 b. the girl that I saw
 c. the girl I saw

As shown in (222), the sources of variation include the following:

- The presence/absence of an overt relative pronoun D_{rel} .
- The presence/absence of an overt complementizer C .

Additional sources of variation include:

- The presence/absence of an overt head noun.
- The presence/absence of an overt external determiner.
- The position of the head NP.

This yields the difference between fully headed, semi-free, free, and internally headed free relative clauses; see the English examples in (223), which are taken from Cardoso and De Vries (2010).

- | | | |
|----------|---|-----------------------------------|
| (223) a. | the pirate who Jack admires | [headed relative] |
| b. | he/those/someone/the one who Jack admires | [semi-free relative] |
| c. | who Jack admires; what Jack did | [free relative] |
| d. | whichever man Jack admires | [internally headed free relative] |

If the specifying coordination analysis of ARCs involves a complete RRC in the second conjunct, the same type of variation is expected to occur in appositive constructions. More precisely, variation is expected to be found in:

- The presence/absence of an overt relative pronoun D_{rel} .
- The presence/absence of an overt complementizer C .
- The presence/absence of an overt (additional) external D .
- The presence/absence of an overt (additional) head NP.
- The position of the additional head NP, if present.

As will become clear below, these predictions are confirmed by the existence of various appositive construction types.

B. Overview of the construction types

Cardoso and De Vries (2010) show that the predicted patterns are attested in the synchronic and diachronic dimensions. Below, (224) outlines some of the relevant possibilities; for ease of exposition, they are illustrated with English words, and only overt elements are indicated.

- | | | |
|----------|------------------------|--------------------------------------|
| (224) a. | DP, C ... | Jack, that is my best friend |
| b. | DP, D_{rel} ... | Jack, who is my best friend |
| c. | DP, D C ... | Jack, the that is my best friend |
| d. | DP, D D_{rel} ... | Jack, he who is my best friend |
| e. | DP, NP D_{rel} ... | Jack, man who is my best friend |
| f. | DP, NP C ... | Jack, man that is my best friend |
| g. | DP, D NP C ... | Jack, the man that is my best friend |
| h. | DP, D NP D_{rel} ... | Jack, the man who is my best friend |
| i. | DP, D_{rel} NP ... | Jack, which man is my best friend |
| j. | DP, ... | Jack, my best friend |

The patterns listed above can be grouped together into five appositive construction types.

I. *The regular ARC*

The regular ARC may correspond to the pattern illustrated in (224)a or (224)b. In this construction, the D° and NP remain silent, but D_{rel} or C° can be spelled out.

The choice between D_{rel} and C° seems to be subject to minor parametric choices. For instance, in Italian, the ARC can be introduced by a complementizer (see (225)a, taken from Cinque 2008: 100), whereas in English, this option is not available; in this language, ARCs must be introduced by a relative pronoun D_{rel} (see (225)b).

- (225) a. Inviterò anche Giorgio, che abita qui vicino.
lit. ‘I will invite also Giorgio, that lives nearby.’
- b. Jack, who is my best friend ...

The two different options are illustrated in (226):

- (226) [C_{OP} DP Co [DP D [CP NP_i [D_{rel} t_i]_k C [IP t_k]]]]
 e.g. *Giorgio*, \emptyset \emptyset \emptyset *che* *abita qui vicino*
Jack, \emptyset \emptyset *who* \emptyset *is my best friend*

II. *The semi-free appositive construction*⁵⁶

The semi-free appositive construction may correspond to the pattern illustrated in (224)c or (224)d. Here there is an additional D element (possibly combined with a light noun) that can be spelled out as an article or pronoun.

⁵⁶ Semi-free relatives (also called light-headed or false free relatives) are a variant of regular RRCs. The main difference concerns the nature of the head NP. Regular restrictives have an overt, full nominal head. In contrast, in semi-free relative constructions, the external determiner is spelled out in the form of a pronoun or article, and the nominal head remains abstract or can be considered to be part of the pronoun or pronominal complex. The result is a semantically (and often morphologically) light antecedent. This is different from true free relatives, where there is no external element whatsoever (see Cardoso and De Vries 2010).

There is cross-linguistic variation with respect to the light elements that can introduce semi-free relative constructions (see Lehmann 1984, Smits 1988, Rebuschi 2001). For instance, CEP allows for a definite article, as in (227)a, but not for a personal pronoun. In English, however, the light element can be a personal pronoun, as in (227)b, but not a definite article.

- (227) a. A Ana e a Maria, *as* que ganharam uma bolsa de estudo,
the A. and the M. the that won a grant of study
 acabaram de entrar na sala.
have just entered in the room
 ‘Ana and Maria, the (ones) that won the grant have just entered the room.’
 (Alexandre 2000: 30)
- b. Jack, he who is my best friend

Notice also that Portuguese uses a complementizer⁵⁷ and English a relative pronoun. The structural representation of the sentences in (227) is given in (228):

- (228) [_{CoP} DP Co [_{DP} D [_{CP} NP_i [_{DPrel} D_{rel} t_i]_k C [_{IP} t_k]]]]
 e.g. A A. e a M., *as* \emptyset \emptyset *que ganharam uma bolsa...*
 Jack, he \emptyset who \emptyset is my best friend

III. *The appositive construction with an additional external head*

The appositive construction with an additional external head may correspond to the patterns illustrated in (224)e/f/g/h. Here, there is an additional full NP that is left peripheral within the embedded clause. This NP is dubbed *additional external head*.

The additional external head may be preceded by an external D^o and/or followed by an internal D_{rel} and/or C^o. Two of the possible combinations are illustrated in

⁵⁷ There is no consensus in the literature regarding the status of the Portuguese *que* that introduces relative clauses. Traditional grammar analyzes the *que* as a relative pronoun comparable to *quem* ‘who.’ However, it has been claimed that there are good reasons for identifying this *que* with the complementizer that introduces other subordinate clauses (see Brito 1991, Brito 1995, Brito and Duarte 2003). This analysis has, however, been recently challenged by Kato and Nunes (2009), who claim that when introducing relative clauses, *que* is always a relative pronoun and that the *que/quem* alternation can be derived in the morphological component.

I will not go into this discussion here. Following Brito (1991), Brito (1995), and Brito and Duarte (2003), I will simply assume that *que* can be analyzed as a complementizer when introducing subject and object relative clauses.

(229)a, from CEP, and (229)b, from English. In these examples the additional external head corresponds to the NP *viagem* ‘trip’ and *man*, respectively:

- (229) a. *Vínhamos de viagem, viagem que acabava na Avenida da Liberdade.*
returned:1PL of trip trip that finished in.the A. d. L.
 ‘We were returning from a trip, a trip that finished in the Avenida da Liberdade.’
 b. Jack, the man who is my best friend

Again, notice that Portuguese uses a complementizer, and English uses a relative pronoun. These two options are illustrated in (230):

- (230) [_{CoP} DP Co [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} t_k]]]]
 e.g. *viagem, ø viagem ø que acabava na Avenida ...*
Jack, the man who ø is my best friend

IV. *The appositive construction with an additional internal head*

This appositive construction type corresponds to the pattern illustrated in (224)i. Here, there is an additional full NP c-commanded by a dependent relative pronoun D_{rel}. This constituent is dubbed an *additional internal head*. See, for instance (231), taken from Smits (1988: 287), where the additional internal head corresponds to the NP *faithful animal*.

- (231) My dog, which faithful animal has guarded me for years, died last week.

The structural representation of (231) is displayed in (232):

- (232) [_{CoP} DP Co [_{DP} D [_{CP} [D_{rel} NP]_k C [_{IP} t_k]]]]
 e.g. *My dog, ø which f. animal ø has guarded me for years*

V. *The regular appositional construction*

The regular appositional construction corresponds to the pattern in (224)j. To demonstrate the elements that are present/absent in the structure, a short *excursus* is necessary to present Cardoso and De Vries’ (2010) analysis of regular appositional

constructions. Here, I will focus the discussion on attributive appositions. For more details on the analysis of identifying appositions, see the referenced paper.

An apposition might be a simple DP that is linked to the antecedent (i.e., the anchor) by means of specifying coordination. However, there are indications that there is a (implicit) clausal structure in appositional constructions as well.

Cardoso and De Vries (2010), in line with O'Connor (2008) and Heringa (2007, 2009), show that this hypothesis is corroborated by several facts. For instance, all types of adverbs, including sentential and even speech act adverbs, can be used in appositions; see (233), from English, and (234), from CEP.

- (233) a. Norman Jones, *then* a student, wrote several bestsellers.
(Quirk et al. 1985: 1314)
- b. Keith, *once* a drug addict, now leads a rehabilitation centre.
(Heringa 2009: 6)
- c. Racial profiling, *unfortunately* a frequent occurrence in American society, must be stopped.
(O'Connor 2008: 97)
- d. This book, *frankly* not my favourite, won a prize.
(Heringa 2009: 7)
- (234) George W. Bush, *então* o “homem mais poderoso da terra” ...
G. W. B. *then* the man more powerful of the earth
'George W. Bush, then the most powerful man on the earth...'
(www.fundacao-mario-soares.pt/ms/textos/002/105.pdf)

Furthermore, the tense, modality, and illocutionary force of the secondary proposition may differ from that of the primary one:

- (235) Should Jane, once the best doctor in town, marry John?
a. Should Jane marry John?
b. Jane was once the best doctor in town.

In addition, a subordinator may show up in appositions; see (236) in English and (237) in CEP:

- (236) a. John, *though* no longer a coward, was still a weakling.
(Wulf Sachs, *Black Hamlet*; 1937)

b. The victim, *whether* a nice person or not, has to be helped.

(Heringa 2009: 10)

(237) O Belenenses, *embora* vencedor da jornada anterior, não está
the B. although winner of.the round preceding not is

no melhor da sua forma individual e colectiva.
in.the best of.the its form individual and collective

‘Belenenses, although winner of the preceding round, is not in its best individual and collective form.’

(CETEMPúblico v1.7)

Additionally, appositional constructions may apparently involve *wh*-movement. See (238), from CEP, repeated from fn. 35. Here, the DP *quatro das quais*, a partitive construction, is apparently pied-piped along with the relative pronoun to the CP domain.

(238) Com a sua prisão já são cinco as pessoas detidas no
with the his prison already are five the people arrested in.the

âmbito do processo Lasa e Zabala,_{DP} quatro das quais]
context of.the process L. and Z. four of.the which

comandos e militares da guarda.
commandoes and military.men of.the guard

‘With his prison, there are already five people arrested in the process Lasa and Zabala, four of which (are) commandoes and men of the military guard.’

(CRPC [jornal_annotado_RL, Ref: J19128/])

These facts point to the conclusion that regular appositional constructions contain a more extensive functional structure than has hitherto been assumed. The fact that appositional constructions have their own tense, possibly modified by adverbs, suggests that at least TP is projected in the structure. Moreover, the eventual presence of a subordinator, the independent illocutionary force, and the movement of a *wh*-constituent indicate that CP is also projected.

Turning now to the structural representation of regular appositional constructions, the main idea is that an example such as *John, a nice guy* can now be

compared to the appositive relative construction *John, who is a nice guy*.⁵⁸ The difference is that, in the regular appositional construction, not only the CP domain but also the verbal part of the predicate, which corresponds to an abstract copula, is silent. This structure is illustrated in (239):

- (239) [_{CoP} DP [_{Co} [_{DP2} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} t_k BE]]]]]
 e.g. *Jack, ∅ ∅ ∅ ∅ ∅ my best friend*

The existence of such a null copula (or zero copula) in this construction is not particularly surprising because it has been observed in many languages that copulas can be omitted (for a cross-linguistic overview, see Stassen 2008). In CEP, for instance, the omission of the copula is allowed for at least some constructions. Matos (2003: 875) reports that the copula can be omitted from some dependent clauses, as in (240) and (241).

- (240) O cargo pode-lhe ser atribuído desde que [-] compatível com
the position can.to.him:CL be:INF given as.long.as compatible with
 as funções que actualmente exerce.
the duties that currently carries.out

‘The position may be given to him, as long as it is compatible with the duties that he currently performs.’

- (241) Embora [-] cansada, a Maria dispunha-se a acabar o trabalho
although tired the M. was.willing to finish:INF the work
 antes de se ir deitar.
before of SE:CL go:INF to.the.bed

‘Although Maria was tired, she was available to finish the work before going to bed.’

The omission of the copula also occurs in non-standard varieties of Portuguese. See (242)-(244), which involve, respectively, a passive, a cleft and a modal auxiliary:

⁵⁸ Other authors have suggested a relationship between appositions and appositive relative clauses; for earlier ideas, see Smith (1964), Delorme and Dougherty (1972), Halitsky (1974), Klein (1977). Quirk et al. (1985: 1314), for instance, suggest that a regular appositional construction, such as that in *The two men, one a Norwegian and the other a Dane*, may involve a reduced relative clause: *The two men, one (of whom was) a Norwegian and the other (of whom was) a Dane*.

(242) INQ Às vezes até é assim de tijolo, não é?

‘Interviewer: Sometimes they are made out of brick, aren’t they?’

INF Pois. Muitas [-] feitas de tijolo; e outras são feitas só no Verão
±yes many made of brick and others are made only in.the summer

‘Informant: Yes. Many are made out of brick; others are made only in the summer.’

(CORDIAL-SIN - Serpa – SRP13)

(243) E depois essa água que ficava dessa cera escaldada [-] que
and then that water that remained of.that wax heated that

fazia-se os rebolos
made.SE:CL the ±balls

‘And then it was from that water that remained from the heated wax that the balls were made.’

(CORDIAL-SIN - Vale Chaim de Baixo– LUZ31)

(244) Pode [-] que eu esteja enganado!

can:3SG that I be: PRES.SUBJ. wrong

‘It can be that I am wrong!’

(CORDIAL-SIN - Covo– COV17)

The same is true of earlier stages of Portuguese. As shown in (245)-(246), a copula can be omitted, for instance, in clefts (see (245)) and dependent clauses (see (246)):

(245) o q lhe pouco disser [-] q nunca em minha vida não vi
the that to.you:CL can say:INF that never in my life not saw:1SG

nem ovi o q aqui te visto e ovisto
nor heard the that here have:1SG seen and heard

‘What I can tell you is that I have never seen nor heard in my life what I have been seeing and hearing here.’

(CARDS0001, year 1822)

(246) e agora fis hum requerimento q [-] o ministro hir
and now made:1SG a request that the minister go:INF

la e chamar a Juiza
there and call:INF the judge

‘And now I made a request, which is for the minister to go there and call the judge’

(CARDS0006, year 1822)

The next question addressed by Cardoso and De Vries (2010) is how the anchor is represented in the copular sentence. Their proposal is that the subject of the embedded clause is the additional external D in (239) (possibly with an incorporated N). Recall

that these elements are also silent in some of the appositional constructions already discussed, e.g., in regular ARCs.

To conclude, Cardoso and De Vries (2010) show that (attributive) appositions involve an implicit relative copular clause. Given the similarities between regular appositions and the complex appositional constructions already analyzed, they claim that the same structure can be realized in a number of ways; see (247). The differences lie primarily in the choice of which elements are spelled out and in the respective positions of these elements.

(247)	[_{CoP} [_{DP1} anchor] [_{Co} [_{DP2} D [_{CP} NP _i [D _{rel} t _i] _k C [_{IP} t _k BE predicate]]]]]]
	<i>Jack</i> <i>a nice guy</i>
	<i>Jack</i> <i>who</i> <i>is</i> <i>a nice guy</i>
	<i>Jack</i> <i>he</i> <i>who</i> <i>is</i> <i>a nice guy</i>
	<i>Jack</i> <i>the one</i> <i>who</i> <i>is</i> <i>a nice guy</i>
	<i>Jack</i> <i>some one</i> <i>who</i> <i>is</i> <i>a nice guy</i>

4.1.3. Summary

In the preceding sections, I have introduced the analysis to be adopted for *o qual*-ARCs in earlier stages of Portuguese, i.e., the specifying coordination analysis (De Vries 2006b). The basic tenets of this analysis are as follows:

- ARCs involve structural coordination of the ARC to the antecedent.
- The relationship between the two conjuncts encodes specifying coordination.
- The second conjunct involves a full RRC, i.e., a DP including a CP, in which raising takes place.

Then, I argued (in line with Cardoso and De Vries 2010) that a wide range of appositive constructions can be derived from the specifying coordination account: (i) regular ARCs, (ii) semi-free appositive constructions; (iii) appositive constructions with an additional external head; (iv) appositive constructions with an additional internal head; and (v) regular appositional constructions. Differences arise because of the particular choice of which elements remain implicit and which ones are actually spelled out in the second conjunct.

Now let us consider the analysis to be adopted for *o qual*-ARCs in CEP.

4.2. *O qual*-ARC in CEP

The main claim of this section is that *o qual*-ARCs in CEP involve the raising analysis, which was originally proposed by Brame (1968) and later developed by Schachter (1973), Carlson (1977), and especially Vergnaud (1974). In Section 4.2.1, I will outline the basic tenets of this proposal, focusing on Kayne's (1994) and Bianchi's (1999) implementations. Because the raising analysis has already been introduced in some detail in Chapter 2, I will present it here only briefly, paying special attention to the distinction between RRCs and ARCs (generated by the raising analysis).

4.2.1. The raising analysis

A. The raising analysis of RRCs

As already mentioned, the main idea underlying the raising analysis is that the head NP (the antecedent) of an RRC originates at the relativization site inside the subordinate clause and then raises to the left edge. The relative clause itself is generated as the complement of the so-called external determiner, with which the head NP may associate after raising. A relative pronoun or operator is then analyzed as a *relative determiner* originally belonging to the internal head NP. See the representation in (248). Normally, there are two movement steps: movement of the operator phrase DP_{rel} to the CP domain, and subsequent movement of the head NP to the left of D_{rel} .

$$(248) \quad \begin{array}{c} [_{DP} D \quad [_{CP} [_{DP_{rel}} NP [D_{rel} \quad t_{NP}]] C [_{IP} \dots \dots t_{DP}]]] \\ \text{e.g. } \textit{this} \quad \quad \quad \textit{book} \textit{ which} \quad \quad \quad \textit{I read} \end{array}$$

The operator phrase DP_{rel} moves to the CP domain to check the *wh*-feature on C. For the subsequent movement of the head NP to the left of D_{rel} , I will adopt Bianchi's (1999) proposal, according to which the external D° bears a strong N-feature that needs to be checked by a [+N] category. Because the CP category itself (the complement of D°) has no such feature, the head NP inside CP must be moved to a position governed by (or in the minimal domain of) the external D° .

For the landing site of this movement, I will assume (in line with De Vries 2002) that the head NP is moved to [Spec, DP_{rel}] in sentences such as (248). However, when DP_{rel} is embedded in another constituent and this constituent is dragged along with D_{rel} to the CP domain (i.e., when pied-piping is involved), I assume that the head NP moves to the highest Specifier position within the pied-piped constituent (see Kayne 1994, De Vries 2006a). See, for instance (249), which involves pied-piping of a PP to the CP domain:

$$(249) \quad \begin{array}{cccccc} [\text{DP } D [\text{CP} [\text{PP } \text{NP}_i [\text{P}' \text{ P} [\text{DP}_{\text{rel}} \text{ D}_{\text{rel}} \text{ t}_i]]]]_k \text{ C} [\text{IP} \dots \text{ t}_k]]] & & & & & \\ \textit{the} & \textit{bed} & \textit{in} & \textit{which} & & \textit{he sleeps} \end{array}$$

B. *The raising analysis of ARCs*

Kayne (1994) extends this analysis to ARCs and proposes that ARCs differ from RRCs only at the level of logical form: ARCs involve covert remnant movement (at LF) of the relative IP to the Specifier position of the external determiner D, where it is no longer in the scope of either D or the head NP. The representation is given in (250):

$$(250) \quad \begin{array}{ll} \text{a. } [\text{DP } D [\text{CP} [\text{DP}_{\text{rel}} \text{NP} [\text{D}_{\text{rel}} \quad \text{t}_{\text{NP}}]]] \text{ C} [\text{IP} \dots \dots \text{ t}_{\text{DP}}]]] & \text{(pre-LF)} \\ \text{b. } [\text{DP } \text{IP} [\text{DP } D [\text{CP} [\text{DP}_{\text{rel}} \text{NP} [\text{D}_{\text{rel}} \text{ t}_{\text{NP}}]]] \text{ C} \text{ t}_{\text{IP}}]]] & \text{(LF)} \end{array}$$

According to Kayne (1994), the movement of IP to the Specifier position of D is overt in pre-nominal or head-final relatives (cf. Chapter 2, Section 2.3).

According to this approach, it follows that all differences found between RRCs and ARCs generated by the raising analysis are determined by the different derivation in LF.

C. *Some properties of ARCs derived*

Let us now see how this approach derives some interpretative properties of ARCs, such as the scope of the determiner and the lack of reconstruction effects.

I. *Scope of the determiner*

RRCs differ from ARCs in that only the former are in the scope of the external D (see (218) above). This contrast can be easily derived under the raising analysis. Clearly,

the external D in the configuration in (250)a c-commands the relative clause. The same does not hold, however, for the appositive configuration in (250)b; in this case, after LF movement, the IP of the ARC is no longer c-commanded by the external D.

II. *Reconstruction effects*

RRCs differ from ARCs in that they allow relativization (and concomitant reconstruction) of the head. This fact is illustrated in sentences (218) above, which involve relativization of the object of an idiomatic expression. The lack of reconstruction effects in ARCs is initially unexpected under an analysis that combines head-raising with covert IP movement; the head, being generated inside the relative clause, should in principle be able to reconstruct in a relative clause-internal position. However, as Alexiadou et al. (2000: 32) note, head-raising only opens the possibility for the reconstruction from the head; it does not force it. The lack of reconstruction effects can be consistent with head-raising if independent principles ensure that the head cannot reconstruct in ARCs.

An analysis along these lines is put forth by Bianchi (1999), who claims that the relativization of the idiomatic object in ARCs involves a structure like (251)b below:

- (251) a. * The headway, which we made
 b. LF: [_{DP} [_{IP} we made t_i] [_{DP} the [_{CP} [_{DP} [_{NP} headway] [_{DP} which t_{NP}]]; [_{CP} C° t_{IP}]]]]
 (Bianchi 1999: 148)

Bianchi claims that if the head were reconstructed within IP, the c-command domain of the external determiner would be empty in LF because it would not contain any variable to be bound by it. This would be an instance of vacuous quantification, and it would be ruled out by the Full Interpretation Principle.

III. *Opacity for binding*

The opacity for binding, which is illustrated in (252)b (repeated from (221)), can be explained by assuming that, after LF movement, the IP of the ARC (where the pronoun is placed) is no longer c-commanded by the quantifier.

- (252) a. I gave every assistant_i who loved his_i uniform a new one. [RRC]
 b. *I gave every assistant_i, who loved his_i uniform, a new one. [ARC]
 (Emonds 1979: 236)

For the contexts involving a pronoun that might potentially be bound by material higher up in the matrix (as in (253)b, repeated from fn. 55), Kayne (1994: 163-164, fn 69) and Bianchi (1999: 152-153) suggest that IP is moved further out of DPrel, “to a topic-like position of matrix clause, where it is not c-commanded by any matrix binder” (Bianchi 1999: 152).

- (253) a. Everyone_i spoke about the museum that he_i had visited. [RRC]
 b. *Everyone_i spoke about the Millennium Dome, which he_i had visited. [ARC]
 (De Vries 2006b: 256)

5. Deriving the contrasting properties of the two types of *o qual*-ARCs

The main aim of this section is to show how the contrasting properties reviewed in Section 3.2 can be accounted for by a dual approach to ARCs. Concretely, it is proposed that the contrasting properties follow from the fact that *o qual*-ARCs in earlier stages of Portuguese are derived from the specifying coordination analysis, whereas *o qual*-ARCs in CEP are derived from the raising analysis. It is not the purpose of this section to discuss if other analyses may or may not account for the data presented here.

5.1. Internal head

O qual-ARCs in CEP differ from *o qual*-ARCs in earlier stages of Portuguese in that they disallow an additional internal head (see Section 3.2.1.1). Such a contrast can be explained in a straightforward manner if we assume that *o qual*-ARCs in CEP are derived from the raising analysis and that *o qual*-ARCs in earlier stages of Portuguese are derived from the specifying coordination analysis.

According to the raising analysis of relative clauses, the head NP originates as the complement of the relative determiner D_{rel} , as represented in (254). Thus, there is simply no room for an additional internal head because the only NP position available

is already filled with the visible head. This explains why *o qual*-ARCs in CEP cannot take an additional internal head.

(254) [DP D [CP NP_i [DP_{rel} *o qual* t_i]_k C [IP ... t_k ...]]]

In contrast, under the specifying coordination account, there are two NP positions in the appositive construction: the external antecedent in the first conjunct and the NP position within the second conjunct, as shown in (255). Because the antecedent is base-generated in the first conjunct, the NP in the second conjunct may be spelled out as an additional internal head. This configuration explains why *o qual*-ARCs in earlier stages of Portuguese can take an additional internal head.

(255) [CoP [DP antecedent] Co [DP D [CP [DP_{rel} *o qual* [NP internal head]]]_k C [IP ... t_k ...]]]]

Furthermore, it should be noted that the structure in (255) also explains the following properties of the internal head in earlier stages of Portuguese.

A. *Categorial nature of the antecedent/internal head*

The eventual non-categorial identity between the antecedent (which may be non-nominal) and the internal head (which must be nominal) can be explained by the structure in (255): regardless of the category of the constituent at which the second conjunct is attached, the internal head is always nominal because it is the complement of D_{rel}.

B. *Semantic class of the nominal head/C. Relationship between the antecedent and the internal head*

The facts that there is no restriction on the semantic class of nouns that appear as internal head and no necessary identity between the antecedent and the internal head are compatible with the structure in (255). Because there is no movement chain between the antecedent and the internal head, nothing forces phonological or semantic identity between both elements, and no restriction is imposed on the semantic class of the noun that shows up as internal head.

D. *Expansion of the internal head*

The internal head NP can be expanded by different elements, e.g., numerals, PPs, and relative clauses. This is straightforwardly derived from (255). As a determiner, D_{rel} can take as its complement a NP (as has been assumed thus far) and also a higher constituent, such as a QP. Moreover, as in regular noun phrases, the internal head can be associated with different post-nominal modifiers.

E. *Contexts of occurrence*

The occurrence of the internal head as a strategy to avoid ambiguity (for instance, in contexts of non-adjacency between the antecedent and the ARC) follows from the structure in (255); if the first conjunct contains more than one potential antecedent, an additional internal head may be spelled out to resolve the ambiguity.

5.2. Extraposition

As already mentioned in Section 3.2.1.2, *o qual*-ARC extraposition is possible in all periods of the history of Portuguese. However, the restrictions on extraposition are much less constrained in earlier stages of Portuguese than in CEP. I submit that this contrast relies on two related aspects. First, *o qual*-ARCs in CEP and in earlier stages of Portuguese involve different structures and derivations. Second, the different structures that derive *o qual*-ARCs correlate with different strategies to generate extraposition.

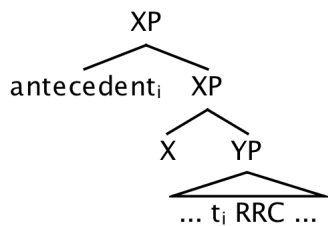
Before proceeding with the analysis of *o qual*-ARC extraposition, let us briefly review the analysis of RRC extraposition put forward in Chapter 3. Table 7 presents the guidelines of the proposal:

Table 7. (Non)-extraposed RRCs (Chapter 3)

	CEP	Earlier stages of Portuguese
non-extraposed RRCs	raising analysis (Kayne 1994)	raising analysis (Kayne 1994)
extraposed RRCs	stranding analysis (Kayne 1994)	specifying coordination analysis (plus ellipsis) (De Vries 2002)

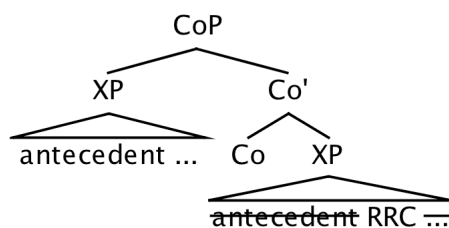
In CEP, RRCs are derived from the raising analysis (see Kayne 1994, among others). The same structure is involved in RRC extraposition. In this case, the antecedent is base-generated inside the RRC and undergoes leftward movement, stranding the RRC *in situ*, as schematically represented in (256):

(256)



In earlier stages of Portuguese, both extraposed and non-extraposed RRCs are generated by a raising structure. However, in extraposed RRCs, the raising structure occurs in the second conjunct of a specifying coordination structure, as proposed by De Vries (2002) (see (257)). The second conjunct has the same categorial status as the first conjunct; it repeats the material contained in the first conjunct, adding the extraposed RRC in its canonical position. Then, the repeated material is phonologically deleted.

(257)



Recall from Chapter 3 that this non-uniform approach to RRC extraposition is necessary to account for the contrasting properties found in the diachronic (and cross-linguistic) dimension.

The same line of reasoning is adopted in this chapter to account for *o qual*-ARC extraposition. As shown in Section 3.2.1.2, *o qual*-ARC extraposition in CEP exhibits the same restrictions as RRC extraposition in CEP. Thus, I submit that:

- Non-extrapolated *o qual*-ARCs in CEP are derived from a raising structure (just as non-extrapolated RRCs in CEP).
- Extrapolated *o qual*-ARCs are derived from a stranding structure (just as extrapolated RRCs in CEP).

For earlier stages of Portuguese, I propose that:

- Non-extrapolated *o qual*-ARCs are derived from specifying coordination (at the DP level).
- Extrapolated *o qual*-ARCs are derived from specifying coordination (at different levels of projection).

Note that in earlier stages of Portuguese, *o qual*-ARCs already involve the specifying coordination analysis, as shown in (258).

(258) [_{CoP} DP Co [_{DP} D [_{CP} [*o qual* (internal head)]_k C [_{IP} ... t_k ...]]]]

Thus, extraposition of *o qual*-ARCs is derived by assuming that the second conjunct can be attached at different levels of projection; see (259), where XP stands for any category (e.g., IP, CP, ...):

(259) [_{CoP} XP Co [_{DP} D [_{CP} [*o qual* (internal head)]_k C [_{IP} ... t_k ...]]]]

Here, I depart from De Vries' (2002) analysis of extraposition in assuming that only one syntactic structure of coordination is involved in ARC extraposition. Under De Vries' account, ARC extraposition would, in principle, require two distinct syntactic configurations of specifying coordination: a specifying coordination structure to derive ARCs and an additional specifying coordination (plus ellipsis) structure to derive extraposition, as in (260) (cf. De Vries 2002: 279).

(260) [_{CoP} [_{XP1} ... antecedent YP] [Co [_{XP2} [_{CoP} [_{DP1} antecedent]
Co [_{DP2} D [_{CP} *o qual* (internal head)]_k C [_{IP} ... t_k ...] YP]]]]]

I opt for the simpler structure represented in (259), which involves only one structure of coordination, i.e., the one that is independently necessary to derive *o qual*-ARCs in

earlier stages of Portuguese.⁵⁹ It should be noted that this position is compatible with the hypothesis pursued in Chapter 3, i.e., that extraposition is not a uniform construction-type.

In summary, an overview of the analysis proposed in this chapter for *o qual*-ARCs is given in Table 8.

Table 8. (Non)-extraposed *o qual*-ARCs (Chapter 4)

	Earlier stages of Portuguese	CEP
non-extraposed <i>o qual</i> -ARCs	specifying coordination analysis (De Vries 2002)	raising analysis (Kayne 1994)
extraposed <i>o qual</i> -ARCs	specifying coordination analysis	stranding analysis (Kayne 1994)

With this background in mind, the properties of *o qual*-ARCs in CEP described in Section 3.2.1.2 can be derived. As the technical details of the analysis have already been introduced in Chapter 3 (for RRC extraposition), only a brief recapitulation of the argument is necessary here.

A. *The definiteness effect*

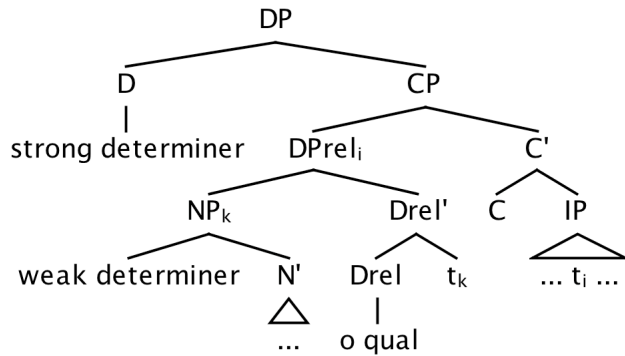
As mentioned in Section 3.2.1.2, extraposed *o qual*-ARCs can take as antecedent weak noun phrases but not strong noun phrases. This property can be explained if we assume, following Bowers (1988), that strong and weak noun phrases differ in structure: strong quantifiers are of category D, whereas weak quantifiers are adjectives and attach within NP.

Extending Bower's proposal to the raising analysis of relative clauses, I assume that strong quantifiers are located in the external determiner, whereas weak determiners are within NP. This straightforwardly explains why extraposed *o qual*-ARCs can only take weak noun phrases as antecedents. As depicted in (261), weak noun phrases can be moved leftward as a constituent, whereas strong noun

⁵⁹ It can be argued that the structure in (259) has the disadvantage of involving coordination of unequal categories (a problem that does not arise in (260) because the conjuncts are of the same category). It should be noted, however, that the coordination of unequal categories is independently necessary to account for the instances of *o qual*-ARCs with a clausal antecedent (see Section 5.4 below).

phrases cannot, because the strong determiner and the noun phrase (NP_k) do not form a constituent.

(261)



B. Restriction on extraposition from embedded positions

In CEP, extraposition of *o qual*-ARCs is not allowed if the antecedent is the object of a preposition. Again, this restriction is straightforwardly derived under the standard assumption that movement only applies to constituents. For instance, in sentences such as (88) above (here repeated as (262)), the preposition and the noun phrase in [Spec, CP] do not form a constituent (excluding the ARC); as a result, they cannot undergo leftward movement and strand the ARC *in situ*.

- (262) *Discuti com um amigo meu ontem, o qual teima em
argued:1SG with a friend mine yesterday the which insists on
 dizer que não vai votar nas próximas eleições.
say:INF that not goes vote:INF in.the next elections
 ‘Yesterday I argued with a friend of mine, who insists on saying that he is not going to vote in the next elections.’

C. Restriction on extraposition from pre-verbal positions

In CEP, extraposed *o qual*-ARCs can take post-verbal subjects but not pre-verbal subjects as antecedents. Additionally, extraposed *o qual*-ARCs can take *wh*-constituents and preposed foci as antecedent but not topicalized constituents. Just as proposed for RRC extraposition (see Chapter 3, Section 5), I submit that the

explanation for these contrasts rests upon the semantic interpretation of the antecedent. More precisely, I claim that extraposition of *o qual*-ARCs in CEP also obeys the *Interpretative Principle* given in (263):

(263) *Interpretative Principle*

The antecedent of extraposed *o qual*-ARCs must occur in a position non-ambiguously interpreted as non-topic (in Kuroda's 2005 sense).

In (263), the concept of topic is understood in a semantic sense, referring to a constituent that expresses an *aboutness relation* (see Kuroda 2005); it can be familiar, recognizable, presupposed, or part of the common ground, but it need not be old information.

The rationale behind the Interpretative Principle in (263) is found in word-order patterns in CEP. Several authors (see, e.g., Duarte 1997, Martins in prep.) have already observed that word order in CEP reflects not only information structure but also the contrast between categorical andthetic judgments (in the sense of Kuroda 1965, 1972, 2005). For further details and examples, see Chapters 1 (Section 4.3.2) and 3 (Section 5.4).

C1. *Restriction on extraposition from pre-verbal subjects*

The Interpretative Principle in (263) derives the restriction on extraposition from pre-verbal subjects as follows. [Spec, IP] is an ambiguous position in CEP; it can be filled by topic elements (i.e., the subject of predication in sentences expressing categorical judgments) and by non-topic elements (i.e., the subject of a sentence expressingthetic/descriptive judgments). In contrast, the Specifier and adjunct positions of VP are non-ambiguous positions; they can only be filled by non-topic elements. Hence, in accordance with the Interpretative Principle in (263), the antecedent of an extraposed *o qual*-ARC cannot be a pre-verbal subject in [Spec, IP] because this position can be filled by topic and non-topic elements. In contrast, a post-verbal subject left-adjoined to VP satisfies the Interpretative Principle in (263) and, therefore, can be taken as the antecedent of an extraposed *o qual*-ARC. For further details, see Chapter 3 (Section 5).

C2. Restriction on extraposition from other pre-verbal constituents

The Interpretative Principle in (263) explains why extraposition cannot take place from topics. Assuming, along with Rizzi (1997) and much related cartographic work, a split-CP approach, according to which there are multiplied functional projections especially dedicated to single discourse functions (such as Topic and Focus), the position occupied by topic constituents is non-ambiguously interpreted as topic; hence, it is ruled out by the Interpretative Principle in (263).

On the other hand, the position occupied by preposed foci and *wh*-constituents is non-ambiguously interpreted as non-topic; hence, the possibility of extraposition from these constituents is straightforwardly derived.

D. Restriction on the intervening material

In CEP, if the antecedent occurs in a post-verbal position, only PPs and adverbs can break the adjacency between the antecedent and the *o qual*-ARC. These constituents can be either adjuncts or complements of the verb.

Consider first the contexts involving an intervening adjunct, as in (264) (repeated from (97)):

- (264) Vi um filme ontem, do qual nunca mais me esquecerei.
saw:1SG a movie yesterday of.the which never more forget:FUT.1SG
 ‘Yesterday I saw a movie, which I will never forget.’

Assuming that adverbs and PP-adjuncts are left-adjoined to VP (Barbiers 1995), I claim that a sentence such as (264) involves the base-order adverb-object and that extraposition is derived by leftward movement of the antecedent (*um filme* ‘a movie’ in (264)) past the position of the adverb and by stranding the relative clause in its base-position; see (265):

- (265) a. [_{IP} V [_{VP} DO [_{VP} Adjunct [_{VP} t_v t_{do} *o qual*-ARC]]]] [object antecedent]
 b. [_{IP} V [_{VP} S [_{VP} Adjunct [_{VP} t_s *o qual*-ARC t_v]]]] [subject antecedent]

Consider now the contexts involving an intervening complement, as in the double complement constructions given in (266) (repeated from (98)).

- (266) Foi oferecido um jantar aos congressistas, no qual estava
was offered a dinner to.the congresspersons in.the which was
 presente o Presidente da República.
present the President of.the Republic
 ‘A dinner, at which the President of the Republic was present, was offered to the congresspersons.’

Following Costa (2004a), I assume that, under certain circumstances, the PP-DP order can be base-generated in CEP. The factor that induces the PP-DP base order in syntactic environments such as (266) is the heaviness of the DP *um jantar no qual estava presente o Presidente da República* ‘a dinner, at which the President of the Republic was present.’ The extraposition is then derived by leftward movement of the antecedent *um jantar* ‘a dinner’ past the position of the intervening PP complement *aos congressistas* ‘to the congresspersons.’ The *o qual*-ARC is stranded in its base position, as depicted in (267).

- (267) [IP V [VP DO [VP PP [V' t_v t_{do} *o qual*-ARC]]]]

When extraposition takes place from constituents in a pre-verbal position, other elements may intervene between the antecedent and the *o qual*-ARC, such as the subject and the verb. This can be explained under the assumption that preposed foci and *wh*-constituents undergo leftward movement to the left periphery, stranding the ARC *in situ*. As a consequence, the elements located in the IP domain (such as the verb and the subject) may show up as intervening material.

Let us consider now the extraposition in earlier stages of Portuguese. The major challenge here is to explain why the restrictions on contemporary extraposition do not hold for earlier stages of Portuguese. Specifically, it is necessary to provide an answer to the following questions.

- Why can *o qual*-ARCs in earlier stages of Portuguese take pre-verbal subjects as antecedent?
- Why can *o qual*-ARCs take strong noun phrases as antecedents?
- Why do there seem to be no restrictions with respect to the type and number of elements that may appear as intervening material?

- Furthermore, why can *o qual*-ARCs relate to an antecedent across discourse?

Assuming a strict correlation between the syntactic structure of relative clauses and the type of structure that derives extraposition, I propose that both the extraposed *o qual*-ARCs and the non-extraposed *o qual*-ARCs are derived from specifying coordination. The difference concerns the level of attachment of the second conjunct: if $XP = DP$, adjacency between the antecedent and the *o qual*-ARC is derived; if $XP \neq DP$ extraposition is derived.

(268) $[_{CoP} XP Co [_{DP} D [_{CP} [o qual \text{ (internal head)}]_k C [_{IP} \dots t_k \dots]]]]$

With this background in mind, let us consider how to derive the properties of *o qual*-ARCs in earlier stages of Portuguese discussed in Section 3.2.1.2.

A. *Definiteness restriction*/B. *Restriction on extraposition from PPs*

Under the raising analysis, strong noun phrases and PPs are not constituents and consequently cannot undergo leftward movement, stranding the relative clause in its base position. Under the specifying coordination analysis, however, strong noun phrases and PPs are detached from the relative clause and base-generated in the first conjunct of the coordinate structure. There is simply no movement chain between the antecedent and the position of the gap inside the relative CP; thus, no restriction on movement applies. This is illustrated in the simplified structure given in (269), where the extraposed *o qual*-ARC takes a strong noun phrase as antecedent (*o dicto herdamento* ‘the aforementioned land’).⁶⁰

⁶⁰ It may be asked why the relative is not simply conjoined with VP, after which the verb might raise to I, as in (i). The answer is that (i) would violate the Coordinate Structure Constraint (CSC), according to which “In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.” (Ross 1967: 98-99)

(i) $V [_{CoP} [_{VP} S t_V O] [Co [_{DP} o qual\text{-ARC}]]]$

- (269) [_{CoP} [_{IP} ficou [o dicto herdamento] [ao dicto Moesteyro
stayed the mentioned land to.the mentioned monastery
de suso nomeado]] [Co [_{DP} o qual herdamento est ...]]]
of above mentioned the which land is ...

C. Restriction on extraposition from pre-verbal positions

Given the possibility of attaching the relative clause at different levels of projection, the specifying coordination analysis predicts that an extraposed *o qual*-ARC can take any constituent as antecedent. This would be derived as follows: when the antecedent is a subject in a pre-verbal position, the second conjunct would be attached at the IP level projection; when the antecedent is a topic, the second conjunct would be attached at the Topic-level projection (assuming a split CP system). However, the actuality is not so simple because in the historical data considered thus far, extraposed *o qual*-ARCs can take pre-verbal subjects as antecedents but not topics (as antecedents).

C1. Restriction on extraposition from pre-verbal subjects

The analysis of extraposition of *o qual*-ARCs from a pre-verbal subject is schematically represented in (270):

- (270) [_{CoP} [_{IP} toda a outra cidade era devassa] [Co
all the other city was opened
 [_{DP} na qual moravam muitas gentes]]]
in.the which lived many people

Note, however, that according to the Interpretative Principle in (263), extraposition from pre-verbal subjects should not be allowed because a constituent in [Spec, IP] can be semantically interpreted as topic or non-topic. In line with the proposal put forward for Chapter 3, I tentatively submit that CEP and earlier stages of Portuguese may resort to different strategies to resolve the ambiguity expressed by the Interpretative Principle in (263). Although in CEP, the ambiguity associated with [Spec, IP] is resolved syntactically and prosodically (through subject inversion), in earlier stages of Portuguese, it may be resolved by prosody alone. In this case, a constituent in [Spec, IP] may be unambiguously interpreted as non-topic if it is prosodically marked by

pitch accent. Further research is necessary in this domain to warrant the validity of this hypothesis.

C2. Restriction on extraposition from other pre-verbal constituents

In the corpus of historical Portuguese inspected thus far, *o qual*-ARC extraposition can take place from preposed foci. In this case, the coordinate structure involves coordination of a dedicated functional projection (say, Focus) of the left periphery, as shown in (271). Such a configuration satisfies the Interpretative Principle presented in (263) because the position occupied by the preposed constituent is non-ambiguously interpreted as non-topic.

- (271) [_{CoP} [_{FocusP} a de Aristrato determinava _____ de quebrar] [Co
the of A. determined:3SG of break:INF
 [_{DP} a qual pintura era nobre á maravilha]]]
the which painting was noble to.the wonder

In turn, extraposed *o qual*-ARCs with a topic as antecedent are not present in the corpus under consideration. Such a restriction follows from the Interpretative Principle in (263): under a split-CP system, a constituent in [Spec, TopicP] is non-ambiguously interpreted as topic.

D. Restriction on the intervening material

In the analysis developed thus far, extraposed *o qual*-ARCs involve attachment at different levels of projection; see (272). Such a structure is instantiated, for example, in sentences (270) and (271) above, where the second conjunct is attached, respectively, at the IP and FocusP levels of projection.

- (272) [_{CoP} XP Co [_{DP} D [_{CP} [*o qual* (internal head)]_kC [_{IP} ... t_k ...]]]]

There are, however, more extreme cases in which the clause introduced by *o qual* refers to an antecedent across the discourse, as in (273), repeated from (123);⁶¹ this

⁶¹ The anaphoric elements that refer to the antecedent are marked in bold. For ease of exposition, the null subject of \emptyset *diz* ‘he says’ is not represented.

construction has been referred to in the literature by different labels, e.g., *relatif de liaison*, *connecting relative* and *relative junction*.

(273) - Senhor, chegou ally *o allmocadẽ*, e parece-me que diz que *the* he neçessario de vos fallar llogo amte que amanheça.

‘- Sir, the Moorish captain arrived there, and it seems to me he is saying he needs to speak to you immediately, before dawn.’

O qual o comde mamdou que viesse.
the which the count ordered that come:IMPERF.SUBJ:3SG

‘The count ordered the Moorish captain to come.’

(Brocardo 1997; Gomes Eanes de Zurara, *Crónica do Conde D. Pedro de Meneses*, 15th century)

In (273), the purported antecedent is not contained in the same utterance as the clause introduced by *o qual*. Instead, it appears in a first-person direct speech, whereas the clause introduced by *o qual* occurs in the third-person narration.

It is not completely clear whether the clause **O qual** *o comde mamdou que viesse* is syntactically connected. Bianchi (1999: 152) suggests that in these constructions there is simply no relative construction involved. Under that view, *o qual* is used as an anaphoric pronoun or determiner, and the clause is either coordinate to the main clause or parenthetical (see Section 2.2.3, for further details).

De Vries (2002), commenting on the sentence from German displayed in (274), emphasizes the apparently ambiguous status of this construction. On the one hand, the second sentence in (274) is verb-final, which is the clause structure of subordinate clauses in German. However, its intonation pattern differs from that in appositive relative constructions, and perhaps may equal the one found in main clauses. Equating these properties, De Vries (2002: 66) concludes that “The relative junction is a special case of a more general pattern whereby, for stylistic reasons, the junction between a main clause and a subordinate clause looks like one between main clauses.”

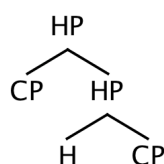
(274) Dieser Wagen ist nicht mehr verbesserungsfähig.
this car is not anymore improvable.

Weshalb wir ihn unverändert weiterbauen.
for.which.reason we it unchanged further.build

(Lehmann 1984, cited in De Vries 2002: 66)

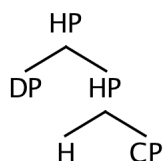
In turn, Cinque (2008: 117-119) claims that the fact that *non-integrated* ARCs may be separated from the head across the discourse follows from ARCs being an independent sentence at the discourse level (see Section 2.2.4 above). Assuming Kayne's LCA to hold for Discourse Grammar as well, Cinque (2008) claims that linear precedence in a discourse must also reflect asymmetric c-command. Under this view, a linearly preceding main sentence is placed in the Specifier of an (empty) head, which, in turn, takes the following main sentence as its complement.⁶²

(275)



Another hypothesis is that a DP is placed in the Specifier of an (empty) head, taking a sentence as its complement.⁶³

(276)



Cinque (2008) takes the configurations in (275) and (276) to underlie the non-integrated ARCs; (275) for the contexts of the *relatif de liaison* and (276) for the anaphoric relations within a sentence.

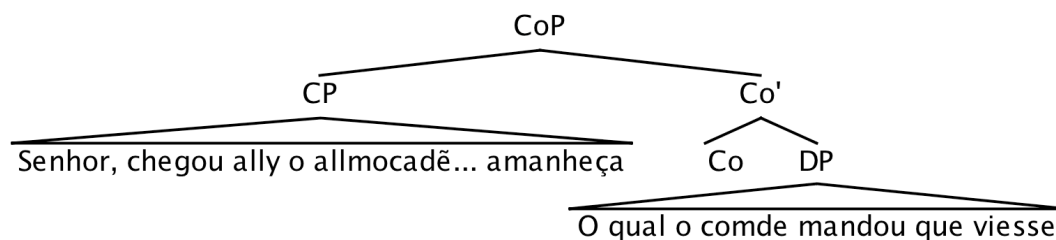
In the present study, I propose, inspired by Cinque (2008), that the so-called *relatif de liaison* introduces, in fact, an ARC that involves coordination at the discourse level (as opposed to the sentence level). Note that the same is true of regular coordination (see Matos 2003: 576), as in *She said: "- Aren't you even curious?" And he looked at her with a strange expression on his face.*

⁶² The structure represented in (275) is instantiated in a sequence such as *John is no longer here. He left at noon.* (Cinque 2008: 118).

⁶³ Recall from Section 2.2.4 that this configuration is found in sequences such as *A pink shirt? I will never wear any such thing in my life!* (Cinque 2008: 118).

This idea can be implemented by assuming that the discourse unit that contains the antecedent surfaces in the first conjunct of a specifying coordinate structure, while the *o qual*-ARCs surfaces in the second conjunct. As an example, in (273), the sentence - *Senhor, chegou ally o allmocadẽ ... amanheça* is merged in the first conjunct, whereas the *o qual*-ARCs is merged in the second conjunct, as illustrated in (277).⁶⁴

(277)



Observe that, given the E-type character of the referential link between the ARC and the antecedent, no adjacency requirement holds between the antecedent and the ARC (see Section 4.1.1). The abstract pronoun heading the second conjunct is able to pick up the right antecedent in the first conjunct, similar to how definite anaphoric or demonstrative pronouns refer to a phrase across discourse.

5.3. Pied-piping

In Section 3.2.1.3, I have shown that pied-piping in contemporary *o qual*-ARCs is subject to constraints that appear not to hold in earlier stages of Portuguese. The contrasts are summarized in Table 9 (repeated from Table 4 above).

⁶⁴ In the structural representation given in (277), I assume, following Moro (2001), that vocative phrases (such as *Senhor* ‘Sir’ in (277)) are located in the CP domain. Under Moro’s (2001) proposal, vocative phrases are hosted in the Specifier of the head projected by a Voc° feature governing Force $^\circ$. Hence, the split Comp field is expanded as follows:

(i) $C^\circ = \dots \text{Voc}^\circ > \text{Force}^\circ > (\text{Top}^\circ > \text{Foc}^\circ > \text{Top}^\circ) > \text{Fin}^\circ \dots$

Given that the split CP has no direct bearing on the analysis at hand, in (277) I adopt a non-split representation, labeling the first conjunct simply as CP.

Table 9. Restrictions on pied-piping (CEP vs. earlier stages of Portuguese)

	DPs	APs	CPs	AdvPs	PPs	Partitive constructions
CEP	-	-	-	+	+	+
Earlier stages of Port.	+	?	+	+	+	+

The explanation I will provide for these contrasts relies on the assumption that the restrictions on pied-piping found in relative constructions result from restrictions on NP movement.⁶⁵

Let us first consider *o qual*-ARCs in CEP, which are derived from the raising analysis. Recall that one of the basic tenets of the raising analysis is that the head NP is generated inside the relative clause as a complement of D_{rel} . Imagine now that D_{rel} is embedded in a PP, as in (278) below:

$$(278) \quad [_{DP} D [_{CP} [_{PP} NP_i [_{PP} P [_{DP_{rel}} DP_{rel} t_i]]]] [C [_{IP} \dots t_j \dots]]]]$$

In this case, the whole PP rises to [Spec, CP] to check the *wh*-feature on C (see Section 4.2.1). However, some explanation is needed for why the *wh*-movement does not only affect the constituent that bears the *wh*-feature (D_{rel}) but instead targets a phrase that properly contains the maximal projection of that item (the PP in (278)). In other words, some mechanism must explain why pied-piping exists.

The standard answer to this question is that there is a mechanism, called *feature percolation*, that spreads the *wh*-features of the *wh*-word up to higher phrases. This proposal refers back to Chomsky (1973) and has been revived by many authors, such as Webelhuth (1992) and Grimshaw (2000). In addition to explaining the nature of this mechanism, these studies are concerned with identifying and explaining the restrictions on percolation. Some of the questions that arise in this respect are what prevents *wh*-feature percolation from occurring freely? Why is it sensitive to the category of the phrases involved?

Note, however, that this line of research does not provide any clue to explain the contrast found in the history of Portuguese: if the same phrasal categories are

⁶⁵ The explanation for the pied-piping found in other constructions (e.g., *wh*-exclamatives, questions) is beyond the scope of this dissertation. Note, however, that the rationale behind my proposal is that the mechanism of *feature percolation* (see, e.g., Chomsky 1973, Webelhuth 1992, and Grimshaw 2000) exists and applies irrespective of the categories involved. Different restrictions on pied-piping result from the different syntactic environment in which pied-piping takes place.

involved, the same restrictions on percolation should hold in CEP and in earlier stages of Portuguese, which is contrary to the actual situation.

Additional evidence for the idea that feature percolation cannot be the whole story is provided by the fact that pied-piping exhibits construction-specific variation. For instance, pied-piping of DPs and CPs is possible in English ARCs but not in RRCs, as shown in (279)-(280):

- (279) a. Most students are interested in Prof. Rotestern, [the security file on whom] the government won't release.
 b. *Most students are interested in any professor [a security file on whom] the government won't release.
- (Emonds 1985: 304)

- (280) a. Egbert, [α to hire whom] would be a real scoop, . . .
 b. *four consultants [to hire whom] would be a real scoop ...
- (Heck 2008: 168)

Such contrasts seem to suggest that the restrictions on pied-piping cannot be simply derived by the restrictions on percolation. Pied-piping appears to be sensitive to the type of relative construction involved; therefore, the syntax of relativization might play an important role in this story.

The hypothesis that I want to put forward here is that percolation exists (or some equivalent of it, as is the case of *feature movement* proposed by De Vries 2006a) and that it applies irrespective of the categories/distance involved. Restrictions on pied-piping are, then, derived not from the restrictions on percolation but rather from the syntactic environment in which pied-piping occurs.

With these ideas in mind, let us consider how the raising analysis can explain the restrictions on pied-piping found in CEP. As already mentioned in Section 4.2.1, the raising analysis involves two basic movement steps: movement of the operator phrase DP_{rel} to the CP domain and subsequent DP movement of the head NP to the left of D_{rel} . The latter movement usually targets [Spec, DP_{rel}]. However, when DP_{rel} is embedded in another constituent, the head NP targets the highest Specifier position within the pied-pied constituent.

The hypothesis I would like to suggest is that the movement of the head within the pied-piped constituent is subject to the following condition:

(281) *The LP-Intervention Condition on pied-piped constituents*⁶⁶

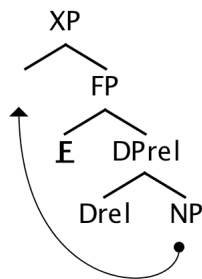
Within pied-piped constituents, NP movement to the highest Specifier position cannot cross lexical projections (LP).

This is illustrated in (282). In (282)a, NP movement is allowed because the head does not cross any lexical projection on its path to the highest Specifier position; in contrast, in (282)b, NP movement is blocked by an intervening lexical projection.

(282)

a.

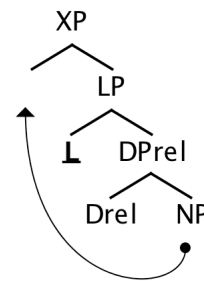
Intervening Functional Projection (FP)



NP movement: *OK*
(no intervening LP)

b.

Intervening Lexical Projection (LP)



NP movement: *BLOCKED*
(by an intervening LP)

Under the standard assumption that N, A and V are lexical projections, the restrictions found in CEP can be derived from the LP-Intervention Condition in (281). Observe the schematic representation given in (283) below, where the constituents to be

⁶⁶ The idea that the intervention of lexical heads is relevant for constraining pied-piping has already been put forth in the literature by different authors; see, e.g., Grimshaw (2000). Here, I am inspired by the LP-Intervention condition proposed by Cable (2007). Although I do not wish to review that proposal here, the basic idea is that *wh*-words are rendered interpretable through the help of a Q-particle, which heads its own projection – QP – and c-commands the *wh*-word. Hence, restrictions on pied-piping result from the fact that in some languages, an agreement relationship must be established between a Q-particle and the *wh*-word. According to Cable, languages showing more limited pied-piping structures are the ones that show Q/Wh-agreement. In technical terms, the Q-interpretable, unvalued instance of Q undergoes agreement with the *wh*-word, which has a valued instance of Q. The most important constraint that holds in these languages is the LP (Lexical Projection)-Intervention Principle, whereby agreement only holds between Q/Wh if no lexical head intervenes between them.

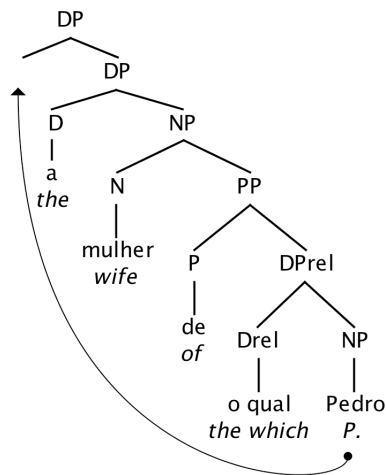
pied-piped is a DP ((283)a), an AP ((283)b), and a CP ((283)c).⁶⁷ The examples are from CEP.

(283)

a.

Pied-piping of DP

*(O) Pedro, a mulher do qual ...
lit. (the) P. the wife of.the which

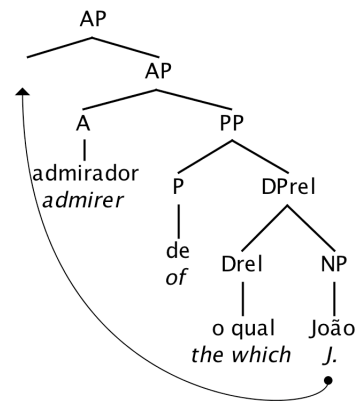


NP movement: *BLOCKED*
(by an intervening LP: N)

b.

Pied-piping of AP

*(O) João, admirador do qual ...
lit. (the) J. admirer of.the which



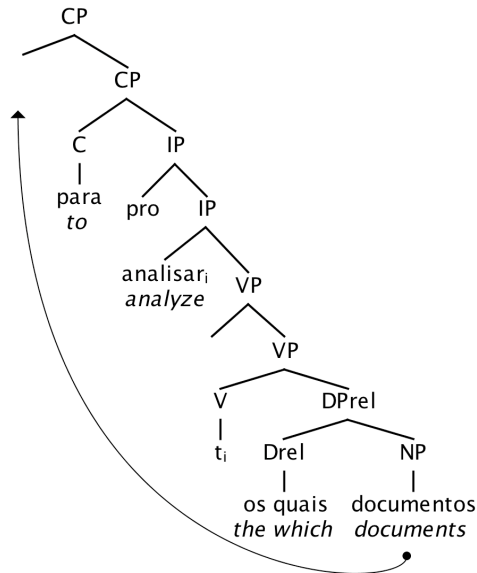
NP movement: *BLOCKED*
(by an intervening LP: A)

⁶⁷ To keep the representation simple, in (283) I abstract away from movement of the head NP to possible intermediate landing sites.

c.

Pied-piping of CP

* (os) documentos, para analisar os quais ...
lit. (the) documents, to analyze the which ...



NP movement: *BLOCKED*
(by an intervening LP: V)

This approach explains the restrictions on pied-piping found in CEP. As shown in (283), pied-piping of DPs, APs and CPs in CEP is blocked by the LP-Intervention condition in (281) because the head crosses a lexical projection (N, A, V) on its path to the highest Specifier position.

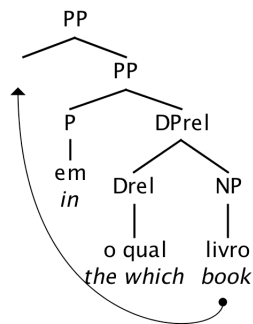
Consider now the pied-piping of PPs and AdvPs. The LP-Intervention condition straightforwardly derives the pied-piping of these phrasal categories: the head NP on its path to the highest Specifier position only crosses functional projections. If a PP is involved (see (284)a), the head crosses D_{rel} and P; if an AdvP is involved (see (284)b), the head crosses D_{rel} , P and Adv.

(284)

a.

Pied-piping of PP

... (este) livro, no qual ...
lit. (this) book, in.the which

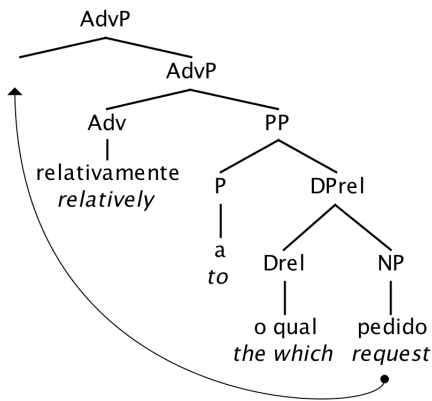


NP movement: OK
(no intervening LP)

b.

Pied-piping of AdvP

... (o) pedido relativamente ao qual
lit. (the) request relatively to which



NP movement: OK
(no intervening LP)

Additional evidence for this theory of pied-piping comes from the unexpected pied-piping of partitive constructions found in CEP. Recall from Section 3.2.1.3 that, unlike DPs, partitive construction may get pied-piped in *o qual*-ARCs, as shown in (285) (repeated from (130)):

(285) Este acto terá levado o industrial a disparar três tiros,
this act have:FUT led the industrialist to fire:INF three shots
[DP dois dos quais] terão atingido o filho no abdómen.
two of.the which have.FUT hit the son in.the stomach

‘This act might have led the industrialist to fire three shots, two of which might have hit his son in the stomach.’

(CRPC [jornal_annotado_RL, AT-0334])

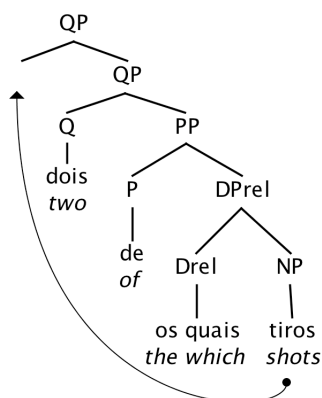
Again, this possibility is derived from the approach adopted here: if pied-piping of partitive constructions is involved, the head only crosses functional projections on its way to the highest Specifier position. See (286), where the head crosses D_{rel} , P and Q.⁶⁸

⁶⁸ In (286), I assume, along with López (2000), that partitive constructions: (i) do not involve an empty nominal head; and (ii) involve a syntactic structure in which the quantifier directly selects a (to be continued)

(286)

Pied-piping of Partitive Constructions

... (três) tiros, dois dos quais ...
lit. (three) shots, two of.the which



NP movement: OK
(no intervening LP)

In sum, the restrictions imposed by the LP-intervention condition explain why PPs and AdvPs can get pied-piped in CEP, whereas DPs, APs and clausal constituents cannot. Note that ultimately, this amounts to saying that the limitations on relative-clause pied-piping follow from a restriction on NP movement.

Consider now the situation in historical Portuguese. Recall from Section 3.2.1.3 that pied-piping of DPs, PPs, AdvPs, partitive constructions and clausal constituents occurs in earlier stages of Portuguese. Let us consider how the broader possibilities for pied-piping can be derived by the specifying coordination analysis.

Under the raising analysis, the antecedent is base-generated as the complement of D_{rel} and then moves to the highest position within the pied-piped constituent. Recall from the previous discussion that this movement is subject to the LP-Intervention Condition. In the specifying coordination analysis, however, such a movement does not have to occur because the overt antecedent is base-generated in

PP. One of the most convincing arguments provided by López (2000) in favor of this analysis is that it is not possible to find a counterpart of the null nominal head in partitive constructions (several (*ones/units) of the students); by contrast, elided pro-forms can always have an overt counterpart ('several (pictures) made in Canada'). It should be mentioned, however, that the earlier hypotheses proposing a structure of the type: [NP [QP several] [N' \emptyset [PP of the students]]] (see, e.g., Jackendoff 1977 and Milner 1978) are also compatible with the approach developed here; in this case, it would only be necessary to assume that empty N is not a blocker for NP movement.

the first conjunct.⁶⁹ As a result, the LP-Intervention Condition is not active and there is no restriction on pied-piping.

5.4. Clausal antecedent

I have noted above (see Section 3.2.2.1) that *o qual*-ARCs in earlier stages of Portuguese, in contrast to *o qual*-ARCs in CEP, can take clausal antecedents.

This contrast can be explained straightforwardly by the dual approach adopted here. Consider the representations in (287):

- (287) a. $[_{DP} D [_{CP} \underline{NP}_i [o\ qual\ t_i]_k C [_{IP} \dots t_k \dots]]]$ [raising analysis]
 b. $[_{CoP} \underline{XP} Co [_{DP} [_{D} D^\circ [_{CP} [_{DPrel} o\ qual\ NP]_j C^\circ [_{IP} \dots t_j \dots]]]]]$ [specif. coord.]

On the raising structure, the NP head (which surfaces as the external antecedent) originates as the complement of the relative determiner D_{rel} ; consequently, it has to be a nominal projection.

On the specifying coordination account, there is no such restriction because the visible antecedent (i.e., XP in (287)b) originates in the first conjunct. Hence, the second conjunct containing the ARC may be attached at different levels (including AP, VP, IP, CP, PP) simply because coordination at any structural level is independently allowed.⁷⁰

5.5. Split antecedents

In Section 3.2.2.2, I have noted that *o qual*-ARCs in earlier stages of Portuguese can take split antecedents, in contrast to the situation found in CEP. Under the dual

⁶⁹ For the sake of concreteness, I assume that in the specifying coordination configuration, a null head always stays in the complement position of D_{rel} (see (i)) and checks the phi-features of the external D (and DP_{rel}) via Agree.

(i) $[_{CoP} DP Co [_{DP} [_{D} D^\circ [_{CP} [_{DPrel} D_{rel} NP]_j C^\circ [_{IP} \dots t_j \dots]]]]]$

⁷⁰ Notice that if $XP = CP$, the coordination is syntactically unbalanced. However, De Vries (2006b) argues that this is permitted if the abstract D° element that heads the second conjunct (possibly associated with the head) refers to CP, such that the two conjuncts are functionally equivalent. According to De Vries, this is possible because a pronoun, in principle, can refer to any syntactic category.

approach advocated here, the question that arises is how the raising analysis can block split antecedents and how the specifying coordination analysis can account for them.

On the raising analysis, the head of the relative clause is considered to originate inside the relative clause. Hence, when split antecedents are involved, two different hypotheses can be formulated. The first one supposes that the antecedents are generated inside the relative clause as a conjoined noun phrase and are subsequently split and moved to different positions, as in (288).⁷¹

(288) [_{CoP} [_{CP} [A man]_k entered the room] and [_{CP} [a woman]_j went out [_{DP} D [_{CP} [_{DPrel} who [_{CoP} t_k [_{Co} Co t_j]]]_i C [_{IP} t_i were quite similar]]]]]]

This hypothesis would explain the plural agreement found in the relative pronoun and the verb (in the languages that can morphologically manifest it). However, the movement of the two conjuncts in (288) violates the CSC (see fn. 60) and, more precisely, the Conjunct Constraint (see Grosu 1973), which bars the movement of whole conjuncts of coordinate structures, ruling out sentences such as **This is the magazine which John bought the book and.*⁷² It is noteworthy that the Across-the-Board raising cannot rescue this CSC violation either because (i) it only applies to movement of constituents contained within a conjunct (as opposed to the conjunct itself); and (ii) it involves extraction of the same element from all the conjuncts (and not extraction of two different constituents).

Another hypothesis states (in line with Suñer 2001) that two identical relative clauses modify each noun phrase, with the subsequent deletion of the first one, as in (289):⁷³

⁷¹ Note that the hypothetical structure given in (288) involves extraposition, as none of the antecedents are adjacent to the ARC. Following Kayne (1994), in this representation I assume that extraposition is derived by VP-internal stranding and that weak determiners are located not in the external determiner but within the NP (see Section 5.2). For ease of exposition, in (288) I abstract away from eventual intermediate landing sites of the conjoined heads.

⁷² For a different interpretation of the CSC, see Zhang (2007).

⁷³ Again, observe that the structure given in (289) involves extraposition of the relative clause modifying *a woman*. See fn. 71 for more details of implementation.

- (289) [_{CoP} [_{CP} [_{DP} D [_{CP} [a man]_i [_{DPrel} who t_i]_k C [_{IP} t_k was/were quite similar]]]] entered the room] and [_{CP} [a woman]_i went out [_{DP} D [_{CP} t_i [_{DPrel} who t_i]_k C [_{IP} t_k were quite similar]]]]]

However, this analysis fails to explain the plural forms found in the relative clause (see, e.g., Andrews 1975). Consider first the plural agreement of the verb. In a relative clause taking split antecedents, the verb in the relative clause is plural (*who were quite similar*). However, the plural agreement is not derived from the structure given in (289); under this structure, the verb should be singular (*was*) because it agrees with a singular head.

The pluralization of the relative pronoun (and of the adjective) raises a similar problem. This can be seen in languages such as Portuguese, where the plural is morphologically visible in relative pronouns and in adjectives (this is evident in the translation *os quais eram bastante parecidos* lit. ‘the:PL which:PL were:PL quite similar:PL’). Again, the structure in (289) cannot explain the pluralization of the relative pronoun and of the adjective because these elements are supposed to agree with a singular head.

Given these facts, it is reasonable to conclude that split antecedents stand out as an obstacle to the raising analysis. Interestingly, this obstacle is highly desirable for the dual approach proposed here because it explains that *o qual*-ARCs cannot take split antecedents in CEP because they are derived from the raising analysis.

Regarding *o qual*-ARCs in earlier stages of Portuguese, a different scenario emerges. Under the approach adopted here, the specifying coordination analysis is expected to allow for ARCs with split antecedents. This is indeed the case. It is simply necessary to assume that split antecedents appear in the first conjunct, whereas the ARC occurs in the second conjunct. Observe that, in this case, the second conjunct is attached not to a noun phrase but to a higher level. Schematically,

- (290) [_{CoP} [_{XP} ... [_i ... [_j]]] [_{Co'} Co [_{DP} D^o _{i+j} [_{CP} ... D_{rel} ...]]]] (XP=IP, CP, ...)

As already mentioned in Section 4.1.1, the abstract D^o heading the second conjunct can be considered a pronoun that behaves as an E-type pronoun requiring co-reference with some objects. For this reason, it can be interpretatively linked to the two parts of

the antecedent (see the referential indexes in (290)), similar to how a pronoun can refer to split antecedents across the discourse, as in (291):

- (291) *A man_i entered the room and a woman_j went out. They_{i+j} were quite similar.*
 (Demirdache 1991: 166)

5.6. Coordination of the *wh*-pronoun with another DP

O qual-ARCs in CEP differ from the ones in earlier stages of Portuguese by not allowing coordination of the *wh*-pronoun with another DP. In accordance with the dual approach adopted here, the question that arises is how the raising analysis blocks the coordination of the *wh*-pronoun with another DP and how the specifying coordination analysis accounts for it.

One of the basic tenets of the raising analysis is that the antecedent is generated inside the relative clause as a complement of D_{rel} . As shown in Section 4.2.1, there are two movement steps: movement of the operator phrase DP_{rel} to the CP domain and subsequent movement of the head NP to the left of D_{rel} . Usually, the head NP targets [Spec, DP_{rel}]. However, when pied-piping is involved, a larger constituent is dragged along with D_{rel} to the CP domain and the head NP moves to the highest position within the pied-piped constituent.

Let us now see how this works in the case of coordination of the *wh*-pronoun with another DP, as in (292), from CEP (repeated from (185) above).

- (292) *O presidente elogiou o João, o qual e a sua mulher
the president praised the J. the which and the his wife
 têm desenvolvido um ótimo trabalho naquela instituição.
have developed a great work in.that institution
 ‘The president praised João; he and his wife have been developing a great work in that institution.’

In this construction, DP_{rel} is conjoined with the DP *a sua mulher* ‘his wife.’ According to the first movement step mentioned above, the whole coordinate structure (CoP) is pied-piped to the CP domain. Then, the head NP undergoes movement to the highest Specifier position within the pied-piped constituent, which in this case corresponds to the Specifier of CoP, as shown in (293).

(293) [_{CoP} João_i [_{CoP} [_{DPreI} o qual t_i] e [_{DP} a sua mulher]]]
 J. the which and the his wife

However, note that this step constitutes a violation of the Coordinate Structure Constraint (CSC),⁷⁴ because it has to postulate the viability of movement of one conjunct alone. Crucially, this violation explains, as desired, why ARCs generated by the raising analysis fail to allow the property at hand.

Turning now to the specifying coordination analysis, the relevant contrasting fact is that there is no movement chain between the antecedent and the position of the gap inside the relative CP. Hence, the coordination of a *wh*-pronoun with another DP would involve the structure in (294):

(294) [_{CoP} antecedent Co [_{DP} D [_{CP} [_{CoP} [_{DPreI} D_{rel} (internal head)] Co [_{DP} D NP]]_k C [_{IP} ... t_k ...]]]]

Because in (294) there is no asymmetric extraction of the antecedent, the CSC is not violated, and the possibility of having a *wh*-pronoun conjoined with another DP follows. According to this approach, the sentence in (186) above, from a fifteenth-century Portuguese text, has the (simplified) structure in (295):

(295) [_{CoP} DP Co [_{DP} D [_{CP} [_{CoP} [_{DPreI} D_{rel} NP] [_{Co} Co DP]]_k C [_{IP} ... t_k ...]]]]
 Dom Henrrique ... o qual dom Henrrique e hum seu tio
 D. H. the which D. H. and a his uncle

5.7. Illocutionary force

O qual-ARCs in earlier stages of Portuguese and *o qual*-ARCs in CEP behave differently with respect to the system of basic clause types: the former allow different clause types, whereas the latter do not allow any clause types beyond the declarative.

⁷⁴ Ross's (1967: 98-99) definition of the CSC is given in fn. 60 and repeated here for ease of exposition: "In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct."

In what follows, I will show that this divergent behavior with respect to the basic clause types can be interpreted as reflecting a functional difference between coordinate and subordinate constructions in terms of illocutionary force.

Several formal criteria have traditionally been used in the literature to distinguish coordinate structures from subordinate ones, e.g., verb second, the possibility of topicalization/preposing, the occurrence of certain adverbs, scopal independence, and illocutionary force. Regarding the illocutionary force, the basic claim is that coordinate constructions have independent illocutionary force, whereas subordinate clauses do not. As Verstraete (2005) puts it,

The basic idea is that coordinate constructions are characterized by the presence of illocutionary force in both clauses in the construction, either separately or shared, whereas subordinate constructions are characterized by absence of illocutionary force in the subordinate clause. In the case of coordination, the presence of illocutionary force in both clauses reflects the ‘equality’ and ‘independence’ that has traditionally been associated with coordinate constructions: the clauses are equal and independent in that each constitutes a speech act just like independent main clauses. In the case of subordination, the absence of illocutionary force in the subordinate clause reflects its status as a discursively presupposed or backgrounded proposition relative to the main clause which does have illocutionary force.

(Verstraete 2005: 613)

Let us consider what predictions these ideas make for the raising/specifying coordination analyses. Under the specifying coordination analysis, the ARC surfaces in the second conjunct of a coordinate structure. Recall that coordinate structures have independent illocutionary force, as illustrated in (296) below, taken from Verstraete (2005: 614).⁷⁵

- (296) a. John was imprisoned, but did he really rob the bank?
 b. John was imprisoned, but don't forget that he robbed the bank!

⁷⁵ Note, however, that not all of the coordinate constructions allow differing illocutionary types (see, for instance, (i) below, taken from Ross 1967/1986: 114). To account for these examples, I assume, following Verstraete (2005), that all coordinate constructions have an independent illocutionary force, and that the fact that not all of them allow the same range of illocutionary force types can be explained by the semantics of the interclausal relation.

- (i) a. *Sally's sick and what did you bring me?
 b. *(You) make yourself comfortable and I got sick.

This fact straightforwardly captures the possibility of having *o qual*-ARCs with (non-declarative) illocutionary force in earlier stages of Portuguese: the second conjunct of the specifying coordination (just like the second conjuncts of the traditional types of coordination) have independent illocutionary force, which is expressible in terms of different clause types.

Let us now consider what the raising analysis predicts. Under the raising analysis, the relative clause is a complement of the external determiner; consequently, it is syntactically a subordinate clause. Given that no coordinate structure is involved, ARCs unambiguously pair with the subordinate constructions and, consequently, are characterized by the absence of illocutionary force. This explains why *o qual*-ARCs in CEP do not allow any clause types beyond the declarative.

There is, however, one possible complication that I wish to make explicit. As mentioned in Section 3.2.2.4, ARCs in CEP are declarative even if the matrix is interrogative or imperative. This is illustrated in (297)-(298) (repeated from (192)-(193) above, for ease of exposition).

- (297) Será que o João, com o qual pudemos sempre contar,
be:FUT.3SG that the J. with the which could:3PL always count:INF
 estará disponível desta vez?
be: FUT.3SG available of.this time
 ‘Will João, on whom we have always counted, be available this time?’

- (298) Telefona aos teus pais, os quais estarão certamente
phone: :IMP.2SG to.the your parents the which be:FUT.3PL certainly
 disponíveis para te ajudar.
available to you.CL help:INF
 ‘Phone your parents, who will certainly be available to help you.’

This might lead us to the assumption that *o qual*-ARCs in CEP are characterized by the presence of illocutionary force. However, as also mentioned in Section 3.2.2.4, ARCs in CEP do not allow any clause type beyond the declarative type found in (297)-(298). This is shown by the unacceptability of (299) and (300) (repeated from (194) and (195) above), where the matrix is declarative and the ARC is, respectively, interrogative or imperative.

- (299) *O único que te apoiou foi o João, ao qual já
the only that you:CL supported was the J. to the which already
 agradeceste devidamente por tudo o que te fez?
thanked:2SG properly by all the which to you:CL made:3SG
 ‘The only person who supported you was John; have you already thanked him properly for everything he made for you?’
- (300) *Acabou de chegar o João, ao qual vai já oferecer
has just of arrive:INF the J. to the which go:IMP.2SG now offer:INF
 uma bebida.
a drink
 ‘João has just arrived; offer him a drink now.’

The divergent behavior of (297)-(298) and (299)-(300) with respect to the different clause types casts some doubt on the presence of illocutionary force in *o qual*-ARCs in CEP. According to Verstraete (2005), if the *o qual*-ARCs in (297)-(298) were genuinely assertive, we would also expect this assertive force to be expressible with non-declarative clause types, such as interrogatives and imperatives, which is contrary to fact. Therefore, following Verstraete (2005), I submit that the declarative in (297)-(298) should be regarded not as a marker of assertive illocutionary force but rather as the unmarked option that emerges in contexts of neutralization of the illocutionary force.⁷⁶

⁷⁶ This line of reasoning is put forward by Verstraete (2005) to account for the contrast between the coordinate construction in (i) and the subordinate construction in (ii). Note that the but-clause in (i) structurally allows different clause types beyond the declarative in (i)a, such as the interrogative in (i)b and the imperative in (i)c. In contrast, the after-clause in (ii) does not allow any clause types beyond the declarative in (ii)a.

- (i) a. John was imprisoned, but he didn't rob the bank.
 b. John was imprisoned, but did he really rob the bank?
 c. John was imprisoned, but don't forget that he robbed the bank!
- (ii) a. John was imprisoned after he robbed the bank.
 b. *John was imprisoned after didn't he rob the bank?
 c. *John was imprisoned after do keep in mind that he robbed the bank!

(Verstraete 2005: 614)

Given these contrasts, Verstraete concludes that the divergent behavior of (i) and (ii) can be interpreted as reflecting a functional difference in terms of illocutionary force: the declarative in (i)a functions as a marker of illocutionary force, whereas the declarative in (ii)a should be analyzed in terms of “a typical instance of a switch to the unmarked option of a paradigm in contexts of neutralization, in this case neutralization of illocutionary force (comparable to the switch to the unmarked member of the paradigm in contexts of phonological neutralization, as discussed by Trubetzkoy, 1939: 77–79, 81).” (Verstraete 2005: 614).

In sum, the theoretical apparatus adopted here derives the divergent behavior of *o qual*-ARCs from the coordinate/subordinate dichotomy; *o qual*-ARCs in earlier stages of Portuguese involve a coordinate structure and, consequently, are characterized by the presence of illocutionary force; *o qual*-ARCs in CEP do not involve a coordinate structure and, consequently, are characterized by the absence of illocutionary force.

5.8. Coordinator

As mentioned in Section 3.2.2.5, *o qual*-ARCs in CEP differ from the ones in earlier stages of Portuguese by not allowing a coordinator preceding the relative clause.

This contrast can be easily explained by the dual approach adopted here. The presence of a coordinator in earlier stages of Portuguese is straightforwardly derived by the specifying coordination analysis; it corresponds to the spelling out of the specifying coordination position Co, as shown in (301):

- (301) [CoP [DP antecedent] [Co [DP D [CP ARC]]]]
 e.g. *bãa raçom e mâtijmêto de pam aluo boroa. carne e viho e o qual mâtijmêto...*
 good ration and provisions of bread white corn.bread meat and wine and the which provisions

On the other hand, the impossibility of having such an element in CEP is straightforwardly explained by the raising analysis; if the relative CP is the complement of the external determiner, and there is no coordinate structure involved, there is simply no room for a coordinative head in the structure.

5.9. Summary

The major goal of this section was to show that the contrasting properties of *o qual*-ARCs in CEP and in earlier stages of Portuguese can be explained by the dual approach to ARCs advocated in this chapter. In particular, it was argued that the properties of *o qual*-ARCs in CEP are derived from the raising analysis of relative clauses, whereas the properties of *o qual*-ARCs in earlier stages of Portuguese are derived by means of the specifying coordination analysis.

The comparison between different stages of the same language proved to be precious empirical grounds for testing the syntax of ARCs. By controlling important

variables (e.g., ARCs introduced by the same relativizer, attested in different periods of the same language), this study offers challenging evidence for the idea that ARCs do not constitute a uniform syntactic phenomenon.

In the next section, I will show that the non-uniform approach to ARCs can be independently confirmed by synchronic evidence, within a single language and across languages.

6. Some comparative remarks

This section is devoted to some comparative remarks on the syntax of ARCs. In light of the dual approach to ARCs advocated in this chapter, it offers an integrated account of the diachronic and synchronic variation found within the same language and across languages.

The diachronic change in the syntax of *o qual*-ARCs is discussed in Section 6.1. The synchronic variation that may be found within the same language and across languages is considered in Sections 6.2 and 6.3.

6.1. Diachronic change in the syntax of *o qual*-ARCs

This section investigates the diachronic change that occurred in the syntax of *o qual*-ARCs. It aims to show that the contrasting properties of *o qual*-ARCs discussed throughout this chapter can be ascribed to a shift from a specifying coordination structure to a raising structure.

I propose that *o qual*-ARCs in earlier stages of Portuguese are generated by the specifying coordination analysis. As already noted in Section 5, this explains why *o qual*-ARCs in earlier periods of Portuguese can have an additional internal head, allow for generalized extraposition and pied-piping, take clausal and split antecedents, allow for the coordination of the *wh*-pronoun with another DP, have illocutionary independence, and be preceded by a coordinator.

In contrast, *o qual*-ARCs in CEP are generated by the raising analysis. As shown in Section 5, this explains why *o qual*-ARCs in CEP cannot have an additional internal head, do not allow for generalized extraposition or generalized pied-piping, cannot take clausal or split antecedents, do not allow for the coordination of the

wh-pronoun with another DP, cannot have illocutionary independence, and cannot be preceded by a coordinator.

I submit that the contrasting properties of *o qual*-ARCs in CEP and earlier stages of Portuguese result from a change in the syntactic structure of *o qual*-ARCs. Specifically, *o qual*-ARCs ceased to be generated by the specifying coordination analysis and started to be generated by the raising analysis.

This diachronic development requires that the starting point was a structure such as (302)a, where the visible antecedent occurs in the first conjunct of a coordinate structure. Then, at a certain period in the history of Portuguese, *o qual*-ARCs started to be generated by a different structure (see (302)b), where the antecedent is internally generated. For the sake of clarity, the visible antecedent is underlined in both structures:

- (302) a. [_{CoP} DP Co [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} ... t_k ...]]]] [spec. coord. analysis]
 b. [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} ... t_k ...]]] [raising analysis]

The change that occurred in the history of Portuguese can therefore be reduced to the loss of the coordinate structure. The antecedent that was externally generated in the first conjunct of a coordinate structure (see (302)a) is analyzed as being generated in an RRC internal position. Hence, (302)b dispenses with the coordinate part of the structure and generates ARCs through a raising structure that is independently available within the second conjunct of (302).

Adopting Lightfoot's (1991, 1999) insights on the relationship between language change and language acquisition, this hypothesis implies that positive evidence triggering the acquisition of a specifying coordination structure ceased to be available to the learners. In the case at hand, I would like to submit that such evidence was found in the context of ARC extraposition involving a strong noun phrase or the object of a preposition as antecedent. The cue for grammars with the specifying coordination analysis might then be an abstract structure such as (303), with a strong noun phrase or the object of a preposition in the antecedent position.

- (303) [antecedent] XP [*o qual*-ARC]

In earlier stages of Portuguese, children knew that the antecedent was externally generated because no movement chain could be established between the visible antecedent and a position inside the relative clause in the extraposition context. As already mentioned in Section 5.2, extraposed *o qual*-ARCs taking a strong noun phrase or the object of a preposition as the antecedent cannot be derived from the raising analysis because such a derivation would require movement of a non-constituent.

If this hypothesis is correct, then the scenario that emerges is that in earlier periods of Portuguese, the cue (303) occurred robustly in the primary linguistic data. Then, for some reason, the expression of the cue decreased. The language learners heard contexts of extraposition less frequently than required, losing evidence of the specifying coordination structure. The question that arises, then, is why the contexts of extraposition involving *o qual*-ARCs decreased in frequency over the history of Portuguese.

I would like to suggest that this decrease is related to independent changes that took place within the history of Portuguese that led to a reduction of the functional positions available in the clausal architecture. Hence, in line with Martins (2002), I submit that earlier stages of Portuguese had a richer clausal structure than CEP, making available more syntactic positions, which, in turn, induced more displacement operations with specific interpretative effects. This is the case for the multiple Specifier positions that were available in the IP domain, which were responsible for the IP scrambling attested in earlier stages of Portuguese.

According to Martins (2002), the IP scrambling found in the history of Portuguese may involve constituents with different categorial properties (e.g., noun phrases, PPs), as illustrated in (304) and (305) below (repeated from Chapter 3, Section 4.3). Moreover, such scrambling is prosodically/discourse-driven, as it allows constituents to escape the default focus stress (and the information focus interpretation).

- (304) sse pela uêtujra uos alguẽ **a dita** **vÿa** enbargar
if by chance you:CL someone the mentioned vineyard blocks:IMPERF.SUBJ
 ‘and if by chance someone blocks the vineyard from you’

- (305) de quẽ lhe **sobre elle** embargo poser
from whoever him:CL over it obstruction put:FUT.SUBJ.3SG

‘[protecting him] from whoever tries to block it [the land] from him.’

(Martins 2001, *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1509,
 cited in Martins 2002: 244)

Not surprisingly, the antecedent of extraposed *o qual*-ARCs (occurring in a embedded or non-embedded position) can undergo IP scrambling. This is illustrated in (306) (repeated from (119)), where the oblique complement *nesta carta* ‘in this letter’ undergoes IP scrambling, stranding the relative clause *in situ*. Observe that the scrambling of the *nesta carta* ‘in this letter’ is confirmed by the fact that this constituent occurs between the proclitic *se* and the verb *contem* (see fn. 31 for more details).

- (306) que este emprazamento valha e se cumpra
that this emphyteusis be.valid:PRES.SUBJ. and SE:CL carry.out:PRES.SUBJ.
 como se nesta carta contem haa qual dou minha auctoridade
as SE:CL in.this letter contains to which give:1SG my authority

‘I want this emphyteusis to be valid and to be carried out as it is written in this letter, to which I give my authority’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1538)

IP scrambling can also generate extraposition from a (strong) noun phrase, as illustrated in (307) (repeated from (115)), which involves scrambling of the direct object *meu sinal* ‘my sign’.

- (307) aqieste prazo fizi. é én testemoyo destas cousas meu
this ±document made:1SG and in testimony of.these things my
 sinal pusi en elle o qual sinal tal este.
sign put:1SG on it the which sign such is

‘I made this document and, as testimony of these things, I put my sign on it; the sign is as follows’

(Martins 2001; *Doc. Portugueses do Noroeste e da Região de Lisboa*; year 1279)

In these examples, the constituents that undergo scrambling are the antecedents of the *o qual*-ARC. Therefore, it is easy to see that the scrambling provides the clue for the specifying coordination structure. For instance, in (306), if scrambling did not apply,

the order V O RC would be derived and, consequently, no extraposition would be obtained.

With the loss of IP scrambling after the 16th century (see Martins 2002), there was a decrease in the frequency of extraposition contexts and, consequently, an important and frequently occurring trigger of the specifying coordination analysis was lost.⁷⁷ Given that the contexts of extraposition became less and less robust in the primary linguistic data, the linguistic learners started to be exposed to more contexts of adjacency between the antecedent and the *o qual*-ARC that were equally compatible with either the raising analysis or the specifying coordination analysis. Of course, the contexts of extraposition still existed (for instance, by movement of the antecedent to a topic position or to [Spec, IP]), but these instances of the cue were not robust enough to trigger the specifying coordination analysis.

Given the higher frequency of the contexts involving adjacency between the antecedent and the *o qual*-ARC, the antecedent is analyzed as being internally generated, and the *o qual*-ARCs start to involve the raising of the visible antecedent, dispensing with the coordinate part of the structure:

- (308) a. [_{CoP} DP Co [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} ... t_k ...]]]] ⇒
 b. [_{DP} D [_{CP} NP_i [D_{rel} t_i]_k C [_{IP} ... t_k ...]]]

Observe that the approach offered here is consonant with the general assumption that “(...) syntactic change is by definition a failure in the transmission across-time of linguistic features.” (Kroch 2001: 699). In this particular case, it results from the ‘imperfect’ acquisition of the structure in (308)a.

A potential problem with this account is that it does not explain why learners did not take utterances expressing other unambiguous cues for the specifying coordination analysis (not affected by the ‘domino effects’ just mentioned) as evidence for the acquisition of this structure. One possible unambiguous cue for the specifying coordination analysis is the structure *o qual N*, which is found in *o*

⁷⁷ In this respect, it is worth noticing that there are probably other independent changes that have not yet been properly studied but that contributed to the decrease in frequency of the extraposition contexts. A possible candidate is a focus position in the left periphery of embedded clauses, which apparently was available in earlier stages of Portuguese, but is not available in CEP, at least in some clause types (see Chapter 2).

qual-ARCs with an additional internal head. As already mentioned in Section 5.1, this structure can only be generated by the specifying coordination analysis. In the raising structure displayed in (309) (repeated from (254)), there is simply no room for an additional internal head because the only NP position available is already occupied by the antecedent of the relative clause. Such a position is, however, available in the specifying coordination analysis in (310). As the antecedent is base-generated in the first conjunct, the complement of D_{rel} may be spelled out as an additional internal head.

(309) $[_{DP} D [_{CP} \underline{NP}_i [_{DP_{rel}} o\ qual\ t_i]_k C [_{IP} \dots t_k \dots]]]$

(310) $[_{CoP} [_{DP} antecedent] Co [_{DP} D [_{CP} [_{DP_{rel}} o\ qual [_{NP} \underline{internal\ head}]_k C [_{IP} \dots t_k \dots]]]]$

Hence, at this point, the question that arises is why were utterances manifesting the cue in (303) more relevant for the learners than utterances exhibiting, for instance, the cue *o qual N*?

Under a cue-based model of acquisition, the most likely answer to this question is that it depends on the robustness of the cue, i.e., on the frequency of utterances that unambiguously express the different cues (see Lightfoot 1999).

However, quantification of the degree to which these two cues are expressed in earlier texts does not permit the drawing of any firm conclusions. In the corpus edited by Martins (2001), the total number of *o qual*-ARCs is 489 but, as illustrated in Table 10, no significant contrast is found in the frequency of relative clauses with and without an internal head.⁷⁸

Table 10. Frequency of *o qual (N)*-ARCs

<i>o qual</i>	<i>o qual N</i>	total
207 (42.3%)	282 (57.7%)	489

Moreover, no significant contrast is found between different periods. The frequency of *o qual (N)*-ARCs (broken down by century) is displayed in Table 11:

⁷⁸ Only *o qual*-ARCs with a nominal antecedent are considered in these figures.

Table 11. Frequency of *o qual* (N)-ARCs (broken down by century)

	<i>o qual</i>	<i>o qual N</i>	total
13 th	30 (39.5%)	46 (60.5%)	76
14 th	73 (39.2%)	113 (60.8%)	186
15 th	68 (47.6%)	75 (52.4%)	143
16 th (first half)	36 (42.9%)	48 (57.1%)	84

Additionally, in the corpus edited by Martins (2001), the total number of *o qual*-ARCs with nominal antecedents is 446. As illustrated in Table 12, the cases of extraposed *o qual*-ARCs correspond to 36.5% of the total instances of *o qual*-ARCs.

Table 12. Frequency of (non) extraposed *o qual* (N)-ARCs

<i>extraposed o qual</i> -ARCs	<i>non-extraposed o qual N</i> -ARCs	total
163 (36.5%)	283 (63.5%)	446

Again, no significant contrast was found across the 13th-16th centuries. The frequency of *o qual*-ARCs (broken down by century) is displayed in Table 13.

Table 13. Frequency of (non) extraposed *o qual* (N)-ARCs (broken down by century)

	<i>extraposed o qual</i> -ARCs	<i>non-extraposed o qual</i> -ARCs	total
13 th	19 (32.8%)	39 (67.2%)	58
14 th	58 (35.4%)	106 (64.6%)	164
15 th	44 (32.8%)	90 (67.2%)	134
16 th (first half)	39 (48.1%)	42 (51.9%)	81

Although more texts must be inspected to confirm these tendencies, I would like to tentatively suggest that the explanation might rely upon the different types of registers in which the different cues are expressed. As mentioned in Section 3.2.1.1, the presence of an additional internal head was mainly used as a strategy to avoid ambiguity when the relative and the antecedent were non-adjacent, as a way of conferring more precision to the utterance. Recent studies have shown that there are good reasons for assuming that ambiguity avoidance determines syntactic choices (see Temperley 2003), and that this might happen more frequently in written language than in spoken language (because writing allows more time for such considerations to

be brought to bear). If this is so, then the explanation for the non-relevance of the cue *o qual N* may rely upon the low frequency of *o qual*-ARCs with an internal head in the spoken language that a child is exposed to during the process of language acquisition (as opposed to non-extraposed *o qual*-ARCs, which occur robustly in the primary linguistic data).

In sum, the basic claim made in this section is that the change that occurred in the syntax of *o qual*-ARCs can be taken as a by-product of other changes independently occurring in the history of Portuguese (the so-called ‘domino effects’, see Lightfoot 1999) that removed the evidence available to language learners of the specifying coordination analysis. Specifically, the decrease in the displacement operations available in the grammar (e.g., IP scrambling) caused a decrease in utterances manifesting extraposition. Given that the available triggering experiences changed in a critical way, children converged on a new grammar and started to generate the *o qual*-ARC with a raising structure.

6.2. Synchronic variation in the syntax of *o qual*-ARCs

In Section 3.2, while describing the syntactic properties of *o qual*-ARCs, I pointed out that there is variation across speakers with respect to some of the syntactic properties of *o qual*-ARCs in CEP. Given that such remarks are scattered in different footnotes throughout the chapter, I will put them together here to provide a comprehensive view of such variation.

With regard to the presence of an additional internal head, I pointed out in footnote 17 (Section 3.2.1.1) that, according to Brito (1991: 133), an internal head is (marginally) possible in CEP, as in (311):

- (311) A falta de monitores na Faculdade de Direito de Lisboa não permitiu
the lack of tutors in.the Faculty of Law of L. not allowed
 ainda que começassem as aulas das subturmas, as quais aulas
yet that start:IMPERF.SUBJ the lessons of.the subclasses the which lessons
 funcionam em regime de avaliação contínua de conhecimentos.
function in regime of evaluation continuous of knowledge

‘The lack of tutors in the Faculty of Law of Lisboa did not yet allow the lessons of the subclasses to start; the lessons function in a system of continuous evaluation of knowledge.’

Bechara (1961/2001: 488) refers to this possibility in sentences such as (312):

- (312) Ao livro ninguém fez referência, o qual livro merece a maior
to.thebook nobody made mention the which book deserves the best
 consideração, no meu entender.
consideration in.the my understand:INF

‘Nobody made any mention of the book, which deserves the best consideration, in my opinion.’

In this respect, it is worth reiterating here that I do not share this judgment, and that the same is true of the people I consulted. To confirm these introspective judgments, I also inspected a large written corpus of European Portuguese: CETEMPúblico (primeiro milhão). This corpus contains some one million words in European Portuguese that were taken from the daily newspaper PÚBLICO. In this corpus, no occurrence of *o qual*-ARCs with an additional internal head is attested.

For the extraposition, I have shown, in Section 3.2.1.2, that in written corpora of CEP it is possible to find extraposed *o qual*-ARCs that take strong noun phrases as antecedents (in a pre-verbal position), as in (313). Again, according to my judgment and informants I consulted, this sentence is ungrammatical.

- (313) Na região da Trofa, dos quatro fogos registados, o mais difícil
in.theregion of.the T. of.the four fires registered the most difficult
 de combater ocorreu em S.Mamede do Coronado, o qual implicou
of fight:INF occurred in S.M. d. C. the which required
 ainda a ajuda dos bombeiros da Maia, Matosinhos e Santo Tirso.
also the help of.the firemen from.the M. M. and S. T.

‘In the region of Trofa, four fires took place. The most difficult fire to fight occurred in S. Mamede do Coronado, which also required the help of the firemen from Maia, Matosinhos, and Santo Tirso.’

(CETEMPúblico 1.7 v. 4.0)

Regarding pied-piping, Brito (1991) and Peres and Mória (1995) claim that pied-piping of DPs (see (314),(315)) and clausal constituents (see (316),(317)) is possible in CEP.⁷⁹ Again, for me and for most informants I consulted, the examples (314)-(317) are ungrammatical.

⁷⁹ I leave aside here the special case of pied-piping of participial clauses. For more details, see fn. 40, Section 3.2.1.3.

- (314) O João, [DP a amiga do qual] tu conheces, telefonou agora mesmo.
the J. the friend of.the which you know called now right
 ‘João, the friend of whom you know, called right now.’

(Brito 1991: 132)

- (315) Foram apresentados vários filmes portugueses muito interessantes,
were presented various movies Portuguese very interesting
 [DP os realizadores dos quais] o Estado deveria apoiar.
the directors of.the which the state should support
 ‘Various interesting Portuguese movies were presented; the state should support the directors of those movies.’

(Peres and Mória 1995: 278)

- (316) Foram descobertas novas provas, [CP para analisar as quais]
were found new proofs to analyze:INF the which
 o tribunal precisa de muito tempo.
the court needs of much time

‘New proofs were found; to analyze them the court will need much time.’

(Peres and Mória 1995: 279)

- (317) Foram descobertas novas provas, [CP considerando as quais] o
were found new proofs consider:GER the which the
 tribunal mudou de opinião.
court changed of opinion

‘New proofs were found; considering these proofs the court changed its opinion.’

(Peres and Mória 1995: 279)

As mentioned in Section 3.2.2.1, the use of *o qual* with a clausal antecedent is found in the *Syntactic Annotated Corpus of Portuguese Dialects*; see ((318)). However, it is completely excluded from the standard variety.

- (318) e era tudo pregado com cravetes, o qual desta
and was everything nailed with ±metal.slivers the which of.this
 forma é mais fácil, com menos despesa
way is more easy with less expense

‘And everything was nailed with metal slivers, which was the easiest and less expensive way to do it.’

(CORDIAL-SIN - Graciosa - GRC12)

For split antecedents, Brito (1991) asserts that *o qual*-ARCs in CEP can take split antecedents, as in (319) and (320) (Section 3.2.2.2). However, for me and the informants I consulted, the ungrammaticality of these sentences is sharp.

- (319) Como a Maria_i não se estava a dar muito bem com o
as the M. not SE:CL was get:INF along with the
 António_j, os quais_{i+j} de facto não têm muito em comum, ele resolveu
A. the which in fact not have much in common he decided
 aceitar o emprego em Lisboa.
accept:INF the job in L.

‘Since Maria was not getting along with António, who in fact do not have much in common, he decided to accept the job in Lisboa.’

- (320) Como a Maria_i veio ao Porto com o Henrique_j, com os
as the M. came to.the P. with the H. with the
 quais_{i+j} eu já não estava há muito tempo, fui jantar
which I already not was has much time went:1SG dine:INF
 com eles.
with them

‘As Maria came with Henrique to Porto, with whom I was not for long time, I had dinner with them.’

These examples provide a great amount of information. One point that I would like to highlight here is that they manifest theoretically inconvenient variation (in the sense of Lightfoot 1991: 98). As the reader may have already noticed, in these sentences *o qual*-ARCs display a range of syntactic properties that are unexpected if *o qual*-ARCs are generated by the raising analysis. As shown in Section 5, the raising analysis cannot derive *o qual*-ARCs with an additional internal head or extraposed *o qual*-ARCs with strong noun phrases as antecedents; it also fails to derive generalized pied-piping and *o qual*-ARCs with clausal or split antecedents.

There is another aspect that is worth mentioning: the sentences outlined above (with the exception of (318)) have a prestigious flavor, in the sense that they would never be used in ‘normal’ CEP. This means that they are somewhat artificial and unnatural, even for people who accept/produce them.

The synchronic variation reported above obviously raises non-trivial questions, such as: What structure is involved in these *o qual*-ARCs? What is the source of the synchronic variation?

Clearly, more research is needed to answer these questions. For example, it is important to determine whether the speakers who accept/produce the prestige *o qual*-ARCs also accept/produce all of the other structures that can be generated by the specifying coordination analysis.

Nevertheless, by capitalizing on the investigation already developed in the domain of language change and variation, at least two hypotheses may be raised to explain the variation under scrutiny. A possible line of explanation is to assume the *competing-grammars hypothesis* (or *syntactic diglossia*) put forth by Kroch (1989, 1994, 2001). This approach rests on the assumption that individuals may synchronically instantiate several grammars in a kind of internalized diglossia. The competing grammars emerge when individuals are exposed to linguistic data that lead to incompatible analysis; a case in point is the competition between a vernacular language and a superposed prestige language. I quote:

(...) it could easily be the case that the forms in competition in syntactic diglossia represent an opposition between an innovative vernacular and a conservative literary language. Since the former would have both a psycholinguistic advantage and the advantage of numbers, it should win out over time, even in written texts.

(Kroch 2001: 723)

Crucially, the competing grammars do not have the same acquisitional status. The vernacular grammar is subject to L1 acquisition, whereas the prestige language is learned a bit later in life for the purpose of reading and writing.

In light of this view, let us suppose that Portuguese children have a grammar that generates *o qual*-ARCs with a raising structure. Later in life, upon exposure to a wider range of language, children may be exposed to *o qual*-ARCs of the type illustrated in (311)-(320), which for them have the status of a prestige construction. Given that this sequence has a structure grammatically incompatible with the one generated by their own grammar, the children may develop a diglossic capacity, becoming able to interpret and eventually use the new construction in their own writing.

From a structural perspective, this hypothesis entails that those individuals who replicate the prestige *o qual*-ARCs resort to a grammar that generates *o qual*-ARCs with a specifying coordination structure. As shown in Section 5, it is only this

structure that can generate *o qual*-ARCs with an internal head, extraposed *o qual*-ARCs with strong noun phrases as antecedents, *o qual*-ARCs involving generalized pied-piping, and *o qual*-ARCs with clausal and split antecedents.

Such a hypothesis is not as surprising as it may first seem, given that the structures involved in specifying coordination are independently available in the grammar. The specifying coordination is independently available in regular appositions (see Section 4.1.2). The raising structure (contained in the second conjunct) is available in RRCs, in non-prestige *o qual*-ARCs, and in the second conjunct of attributive appositional constructions that involve an implicit clausal structure (see Section 4.1.2).

Finally, the competing-grammars hypothesis predicts that the individuals who were not exposed to prestige *o qual*-ARCs or who were exposed to them without sufficient linguistic evidence to develop diglossic grammars would not produce or accept prestige *o qual*-ARCs. Most of the speakers I consulted, including myself, belong to this latter group.

Another possible line of explanation is to assume that prestige *o qual*-ARCs are not part of the core Portuguese grammar but rather the result of extra-grammatical rules for producing prestige forms, which may be cataloged among what Sobin (1997) and Lasnik and Sobin (2000) have termed *grammatical viruses*. One of the central ideas of these proposals is that sometimes speakers use forms that are not generated by the grammar they acquired during the process of language acquisition. The motive for going against the initial system is the desire to employ (or the need to interpret) prestigious forms. Thus, forms licensed by grammatical viruses have a prestige status and are not typical of child language, and the intuitions about their use are strikingly different from intuitions about the use of other grammatical forms.

For the formal implementation of this idea, Lasnik and Sobin (2000) propose that prestige forms are derived from a set of extra-grammatical rules that apply externally to the central computational system in a post-syntactic component. Thus, these rules may change the output only at a very superficial level: they may take into

possible, cross-linguistic evidence was also considered, showing that *o qual*-ARCs in earlier stages of Portuguese systematically pattern with ARCs in other languages. In this comparison, particular attention was given to *il quale*-ARCs in contemporary Italian and to ARCs in English. In this section, I will tentatively provide an integrated account of the facts of cross-linguistic variation considered in this chapter. In doing so, I will show how the findings of this study can be integrated into the typological approach to ARCs put forth by Cinque (2008).

As already mentioned in Section 2.2, Cinque (1982, 2008) proposes a dual approach to ARCs, according to which there are two different types of structures that can generate ARCs (integrated and non-integrated types). These two structures are not, however, instantiated in all languages. From a cross-linguistic perspective, three different types of languages can be identified:

- Type I. Languages that display both constructions, e.g., Italian, French.
- Type II. Languages that display only one construction, e.g., English, Romanian (for the non-integrated type); Northern Italian dialects and, possibly, Chinese (for the integrated type).
- Type III. Languages that lack ARCs, e.g., Gungbe, Bunun, Mixtecan.

Assuming this tripartite classification, I would like to suggest that CEP may belong to Type I or to Type II. The present research has demonstrated that *o qual*-ARCs belong to what Cinque (2008) calls the integrated type but has not confirmed that all ARCs found in CEP are derived the same way. Further research is necessary in this domain to clarify this point. Earlier stages of Portuguese presumably belonged to Type II, as all of the ARCs were apparently generated by specifying coordination.

The dual approach adopted here, combined with Cinque's typological proposal, also provides us with a useful insight to understand the cross-linguistic variation reported throughout this chapter. First, it straightforwardly explains that *o qual*-ARCs in earlier stages of Portuguese exhibit the same syntactic properties as English ARCs and Italian *il quale*-ARCs because all of them are generated by specifying coordination. Second, it explains that *o qual*-ARCs in CEP differ from *il quale*-ARCs in contemporary Italian because *o qual*-ARCs have undergone a syntactic change that apparently did not affect its Italian counterpart.

I believe, however, that the dual approach developed in this chapter, which can be catalogued as a generalized constituency approach, is empirically superior to the approach put forward by Cinque (2008). Whereas, according to the Cinque approach, ‘integrated’ and ‘non-integrated’ *o qual*-ARCs have two completely different derivational stories, under the approach developed here both constructions basically involve one structure: the raising structure. Differences among languages and within the same language result from the possibility of having this raising structure in the second conjunct of a coordinate structure.

From a theoretical point of view, the similarity between the two constructions is highly desirable because it shows that it is still possible to pursue the ideal goal of linguistic theory, according to which the variation across languages and within the same language can be reduced to some different parametric choices. According to this view, an interesting line of research may be to assume that in the languages of the world the concept of apposition can be syntactically expressed in two different ways: complementation and/or coordination. Languages will then differ in the ways they instantiate these two options.

Importantly, the findings of this chapter also show that a generalized constituency analysis (although not uniform) can account for the dual behavior of ARCs across languages without resorting to a version of the orphanage approach. This is highly desirable because, as De Vries (2006b) notes, proposals involving radical orphanage (e.g., Fabb 1990) or the attachment at some grammatical level beyond LF (Safir 1986) cannot be easily accommodated in the standard assumptions about the organization of the grammar. ARCs are interpreted and pronounced; therefore, they must be present at the LF interface and the PF interface. The only way to get at these interfaces is via the overt syntax; if ARCs were to be added at or after the LF interface (i.e., after spell out in Chomsky’s terms), they would not be pronounced.

7. Conclusion

This chapter investigates a case of micro-variation in the syntax of ARCs. It shows that different stages of the same language turn out to be precious empirical grounds

for testing the syntax of ARCs. In particular, it is possible to demonstrate that within the same language, when dealing with the same construction introduced by the same relativizer, it is still necessary to adopt a dual approach to ARCs. Given the constrained nature of diachronic variation, such an approach is particularly attractive, because it allows the control of important variables that may incidentally interfere with the results obtained in other studies that involve, for instance, the comparison of languages historically and typologically quite distant from each other.

Specifically, this study focuses on Portuguese ARCs introduced by the relativizer *o qual*. The main claim is that *o qual*-ARCs have undergone a change from one syntactic type to another at some point in the history of Portuguese.

From an empirical point of view, I show that *o qual*-ARCs in CEP differ from *o qual*-ARCs in earlier stages of Portuguese with respect to a number of syntactic properties. The contrasting properties discussed in this chapter are as follows: (i) the possibility of having an additional internal head; (ii) restrictions on extraposition; (iii) restrictions on pied-piping; (iv) the possibility of taking clausal antecedents and (v) split antecedents; (vi) coordination of the *wh*-pronoun with another DP; (vii) illocutionary force; and (viii) the presence of coordinator.

From a theoretical point of view, I show that a single syntactic analysis cannot account for the contrasts found in the history of Portuguese. For this reason, I argue for a dual approach to ARCs, according to which *o qual*-ARCs in CEP use the raising structure, proposed by Kayne (1994) and Bianchi (1999), whereas *o qual*-ARCs in earlier stages of Portuguese use the specifying coordination structure, proposed by De Vries (2006b).

The dual approach adopted here provides us with a useful insight to understand the variation in the syntax of ARCs found within a language and across languages, both in the synchronic and diachronic dimensions. In this respect, I have shown that (i) ARCs may undergo a change from one syntactic type to another in the diachronic dimension; (ii) two different syntactic structures for ARCs may coexist synchronically within the same language; and (iii) languages may differ synchronically with respect to the syntactic types of ARCs they display.

This chapter was not intended to establish the chronology of the change affecting ARCs in earlier stages of Portuguese; instead, it focuses on the identification of the change (not yet properly identified in the literature) and on the exploitation of its empirical and theoretical consequences. One important task for future research is

to identify the chronology of the change and investigate whether the proposal put forth for *o qual*-ARCs in earlier stages of Portuguese can be extended to the other ARCs.

Chapter 5

Conclusion

This dissertation sheds light on language variation and change from a generative syntactic perspective, based on a case study of relative clause constructions in Portuguese and other languages. Furthermore, it contributes to the theoretical debate on the structural analysis of restrictive relative clauses (RRCs), appositive relative clauses (ARCs), and extraposition. Two important findings are (i) that competing theoretical analyses need not be either false or true universally, but could be instrumental in explaining language variation (both diachronically and synchronically); and (ii) simple lexical changes concerning the availability of (abstract) functional items can have dramatic consequences in the behavior of certain 'construction types' in a particular language.

The research methodology adopted involves comparative syntax (see, e.g., Cinque and Kayne 2005), both in the diachronic and the synchronic dimensions: Contemporary European Portuguese (CEP) is systematically compared with earlier stages of Portuguese; moreover, Portuguese is compared with other languages, in particular Latin, English, Dutch, and Italian.

Such methodology provided precious insights into the diachronic contrasts found in Portuguese. Of particular interest is the finding that earlier stages of Portuguese, contrary to CEP, are to a large extent Germanic-like, at least with respect to the linguistic phenomena scrutinized. The comparative approach also proved to be an invaluable way of overcoming the limitations of historical inquiry. In this respect, it was shown that studying the behavior of other contemporary languages might provide the means to overcome the difficulties posed by the limited nature of written sources and the impossibility of manipulating data.

The linguistic facts are analyzed in the light of the Minimalist version of the Principles and Parameters framework (Chomsky 1981, Chomsky 1993, 1995 and subsequent work). The interpretation and explanation of grammatical changes is

developed within the model proposed by Lightfoot (see Lightfoot 1991, 1999 and subsequent work), which associates diachronic change with language acquisition.

The benefits of using theoretical linguistics in studying diachronic (and synchronic) phenomena are substantial. To single out but a few, theoretical linguistics provided important tools to organize, describe and explain the data. It also oriented the inspection of large-scale corpora in an advanced phase of the research: with the predictions made by the theory, it was possible to search corpora for specific and theoretically informed purposes.

It may be the case that some readers have certain reservations about the methodological option of combining rich empirical documentation (from contemporary and old languages) with the insights of theoretical linguistics. As Devine and Stephens (2006) note, those with a primarily philological background may not appreciate the technical details of the discussion and ‘pure’ syntacticians may become impatient with the rich philological documentation. However, note that the subject of this dissertation does not permit to choose between philology and linguistics. Each discipline makes its own contribution and the present research demonstrates, I hope, that our understanding of language can benefit of this association. To recall the epigraph of this dissertation: *If there are no data, there cannot be any theory. If there is no theory, there can hardly be any understanding.*

The present dissertation is organized around three main linguistic phenomena: remnant-internal relativization, RRC extraposition and ARCs. The selection of these phenomena was determined by the following criteria: (i) the contrasting properties of the relevant structures in earlier stages of Portuguese with respect to CEP; (ii) their novelty (i.e., constructions/properties not yet reported in the literature) and (iii) the theoretical relevance of the facts uncovered.

The first phenomenon addressed is remnant-internal relativization, a term that covers the contexts in which the head noun and a modifier/complement related to it appear discontinuously, in a relative clause internal position; see (1):

(1) [_{S-matrix} ... [_N [_{RC} ... modifier/complement ...]] ...]

On the basis of empirical data from earlier stages of Portuguese, two distributional patterns of remnant-internal relativization were identified:

- Remnant-internal relativization (with a modifier/complement in the rightmost clausal position), as in (2):

(2) os livros que eu compus da philosophia
the books that I wrote of.the philosophy

- Remnant-internal relativization (with a modifier/complement in the left periphery of the RRC), as in (3):

(3) os livros que da philosophia eu compus
the books that of.the philosophy I wrote

These patterns coexist with a non-remnant variant, which involves adjacency between the head and the modifier/complement, as illustrated in (4):

(4) os livros da philosophia que eu compus
the books of.the philosophy that I wrote

The main goal of this initial study is to show that the phenomenon of remnant-internal relativization provides important new evidence for the raising analysis (as in Kayne 1994).

The first step of this inquiry was to demonstrate that the adjunction analysis of RRCs cannot account for the properties of remnant-internal relativization. In a second step, I have shown how the distributional patterns exemplified in (2)-(4) can be derived from the raising analysis of RRCs, combined with a PF deletion mechanism. In a nutshell, the idea is that (2)-(4) are derived from a structure like (5), implemented in terms of the copy theory of movement:

(5) [os [livros da philosophia]ⁱ [que eu compus [livros da philosophia]ⁱ]]
the books of.the philosophy that I wrote books of.the philosophy

The most economical derivation generates regular (i.e., non-remnant) relativization, as in (6). This is due to an economy principle, which prefers fewer applications of deletion in later computations of the phonological component.

(6) [os [livros da philosophia]ⁱ [que eu compus [livros da ~~philosophia~~]ⁱ]]
the books of the philosophy that I wrote books of the philosophy

Less economical derivations are permitted if a PF requirement needs to be satisfied. This is instantiated in (7) and (8), where scattered deletion applies to satisfy various PF requirements, namely discourse requirements on focus and the satisfaction of EPP selectional requirements at PF.

(7) [os [livros ~~da philosophia~~]ⁱ [que eu compus [livros da philosophia]^j]]

(8) [os [livros ~~da philosophia~~]ⁱ [que [da philosophia]^j eu compus [livros ~~da philosophia~~]^j]]

Given the success of this analysis in accounting for the facts under (2)-(4), the central thesis that emerged from the study offered in Chapter 2 is that remnant-internal relativization can figure among the arguments commonly adduced in the literature in favor of the raising analysis of RRCs.

The second study in the thesis deals with RRC extraposition. From a descriptive point of view, I show that different languages and different stages of the same language may differ with respect to the properties of RRC extraposition. The contrasting properties discussed are: (i) the definiteness effect; (ii) extraposition from embedded positions, and (iii) extraposition from pre-verbal positions. The main descriptive findings are:

- Earlier stages of Portuguese contrast sharply with CEP with respect to the properties of RRC extraposition.
- RRC extraposition in earlier stages of Portuguese is, to a large extent, Germanic-like, unlike CEP.

The overall conclusion of the discussion is that the Germanic-like pattern is more liberal than the one found in CEP (and possibly in other contemporary Romance languages).

From a theoretical point of view, I show that one and the same structural analysis cannot alone derive the contrasting properties of RRC extraposition. To account for the variation found in the diachronic and cross-linguistic dimensions, I

argue that RRC extraposition might involve two different structures, one of them derived by specifying coordination plus ellipsis (De Vries 2002), the other the result of stranding (Kayne 1994). See (9)-(10):

(9) ... [_{CoP} [_{XP1} antecedent YP] [Co [_{XP2} [~~antecedent~~ RRC] ~~YP~~]]] (specifying coordination)

(10) ... [antecedent_i YP [t_i RRC]] (stranding)

The different properties of RRC extraposition found in different languages and different stages of the same language follow from the particular structure displayed. In the diachronic dimension, I could then establish that:

- In earlier stages of Portuguese, RRC extraposition involves specifying coordination plus ellipsis.
- In CEP, RRC extraposition involves VP-internal stranding.

The change from one syntactic structure to another is attributed to the fact that the abstract restrictive specifying coordinator that is involved in extraposition in earlier stages of Portuguese is no longer available in the lexicon of CEP.

In the cross-linguistic dimension, I submit that there are at least two types of languages:

- Type I. Languages that generate RRC extraposition by stranding; e.g., CEP (and possibly Italian, Spanish and French).
- Type II. Languages that generate RRC extraposition by specifying coordination plus ellipsis; e.g., English and Dutch.

In a nutshell, the main thesis I have developed in the study offered in Chapter 3 is that the different analyses proposed in the literature for RRC extraposition might be seen not as competing analyses for a single construction but rather as complementary analyses for two distinct structures. In more general terms, this proposal entails that the concept of extraposition might be descriptively useful (in unifying a variety of apparently related constructions/structures), but lacks explanatory force, because it does not unequivocally correspond to a single syntactic type.

The third and last study in the thesis investigates a case of micro-variation in the syntax of ARCs. It focuses on the dissimilar behavior of ARCs introduced by the complex relative pronoun *o qual* in CEP and earlier stages of Portuguese.

From a descriptive point of view, eight contrasting properties are identified, relative to: (i) additional internal head; (ii) extraposition; (iii) pied-piping; (iv) clausal antecedents; (v) split antecedents; (vi) coordination of the *wh*-pronoun with another DP; (vii) illocutionary force; and (viii) coordinator.

Sticking to the comparative approach adopted in this dissertation, data from other languages (in particular, English and Italian) were inspected in the light of the same set of potentially contrasting properties. A finding of particular interest came out of this comparative scrutiny, namely that *o qual*-ARCs in earlier stages of Portuguese pattern with Italian *il quale*-ARCs and English ARCs, unlike CEP.

In order to account for the grammatical contrasts found in the diachronic dimension, I argue for a dual approach to ARCs, according to which *o qual*-ARCs in earlier stages of Portuguese are derived from specifying coordination, whereas *o qual*-ARCs in CEP are derived from raising. Additionally, I show that the dual approach can also provide a useful insight to understand why (i) two different syntactic structures for ARCs may coexist synchronically within the same language; and (ii) languages may differ with respect to the syntactic types of ARCs they display.

In line with the proposal put forward for RRC extraposition, the main thesis of the study offered in Chapter 4 is that ARCs do not correspond to a unique syntactic type and that two complementary analyses are necessary to account for the variation found in the synchronic and diachronic dimension.

I believe that it is not a coincidence that the studies on RRC extraposition and ARCs independently reached the same general conclusion. I take this to show that the apparent similarities between word strings and the traditional constructional view on grammar can be extremely misleading in linguistic inquiry. Another interesting finding is that the various analyses provided in the literature for an apparently uniform construction do not necessarily have to be taken as competing analyses for the same phenomenon. Instead, they can be evaluated as potential complementary tools to explain linguistic variation and change.

One of the main difficulties faced by the current research had to do with the well-known tension between descriptive and explanatory adequacy. The goals of the

dissertation required not only a full account of the richness and the diversity of the properties of the linguistic phenomena under consideration, but also an explanation of the variation itself in the light of the interplay between language invariant principles and conditions on linguistic variation.

Clearly, the exact nature of this interaction needs to be further explored in future research. Nevertheless, some hypotheses were advanced throughout the dissertation to explain certain variation facts:

- As for the contrasts between remnant-internal relativization (with a modifier/complement in the left periphery) in CEP and earlier stages of Portuguese, it was suggested that this change may be due to the loss of a functional projection (FocusP) in the left periphery of relative clauses. Some independent evidence (from apparently unrelated phenomena) has been adduced for this change.
- As for the contrasts between RRC extraposition and ARCs in CEP and earlier stages of Portuguese, it was hypothesized that there are different structural analyses for these constructions in the synchronic and diachronic dimensions.
- As for the contrast found in the syntax of RRCs in earlier stages of Portuguese and CEP, it was suggested that the abstract restrictive specifying coordinator that is involved in extraposition in earlier stages of Portuguese is no longer available in the lexicon of CEP. As a result, RRC extraposition ceased to be generated by specifying coordination and started to involve stranding.

Moreover, it was shown that there is a correlation between the changes affecting ARCs and RRCs, both of which involve loss of specifying coordination. Additional evidence for this view comes from the loss of extraposition of conjuncts in regular coordination.

Crucially, these hypotheses are consistent with minimalist assumptions on language variation. The dominant position among generative grammarians is that language variation is reduced to choices of values of parameters, which are regarded as the setting of a number of ‘switches’ – on or off – for particular linguistic

properties. A classic example is the pro-drop parameter, which determines whether or not a subject must be overtly pronounced in a given language.

Despite the potentialities of the parameter's approach to language variation, it remains controversial and in need of refinement. For instance, it is not yet clear how many parameters there are, in what components of the grammar they are found and how exactly they interact with each other to derive the attested grammatical systems. Nevertheless, the so-called the Borer-Chomsky conjecture (see, e.g., Baker 2008) has become a widespread view in this domain. According to this hypothesis, all the parameters of variation are due to difference in the features of particular items in the lexicon. This line of research also led to the distinction between micro and macro-parameters. Broadly speaking, the idea is that micro-parameters define small-scale differences, whereas macro-parameters aggregate micro-parameters that act in concert for markedness reasons (see Holmberg and Roberts 2010).

Assuming a parametric view on language change, it seems to me that the diachronic facts investigated in the present dissertation may be the result of a macro-parametric change. This means that the loss of remnant-internal relativization (with a modifier/complement in the left periphery), the loss of generalized RRC extraposition and the more restricted possibilities of ARCs in CEP might be integrated in a cluster of phenomena changing at the same time in the history of Portuguese. I believe that the loss of IP scrambling investigated by Martins (2002) may also be a by-product of this change.

The most evident superficial effect of this change is the reduction of word order configurations available in the grammar. Somewhat tentatively I submit that this superficial effect may be due to:

- The amount of functional structure projected (e.g., the FocusP projection ceased to be available in the left periphery of some embedded clauses; the abstract restrictive specifying coordinator used in RRC extraposition ceased to be available in the lexicon).
- The more restrictive nature of the EPP-feature associated with some functional categories (e.g., the loss of IP scrambling is due to a change in the selectional properties of the AgrS functional head; AgrS ceased to allow multiple Specifiers, i.e., it lost the option of being associated with an Attract-all-F EPP-feature; see Martins 2002).

In order to square the change observed in ARCs with the preceding categorization, a deeper investigation is necessary on the typology of specifying coordinators. The opposition between non-restrictive (parenthetical) specifying coordinators (available, for example, in ARCs and regular appositions) and restrictive specifying coordinators (available in extraposition) casts light on the loss of RRC extraposition generated by specifying coordination, since this loss can be simply attributed to the unavailability of a restrictive specifying coordinator in the lexicon. However, the referred opposition does not explain why ARCs ceased to involve specifying coordination, whereas regular appositions still involve this structure. A finer-grained typology of specifying coordinators may possibly illuminate this puzzle.

In conclusion, this dissertation examines synchronic and diachronic contrasts, which, unfortunately, are often consigned to ad hoc and language-particular processes. I hope to have demonstrated that focusing on cross-linguistic and intra-linguistic contrasts offer precious means of enhancing our understanding of various linguistic phenomena, thereby contributing also to our understanding of the interaction between principles and parameters.

References

- Abney, Steven (1987). *The English Noun Phrase in its Sentential Aspect*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Alexandre, Nélia (2000). *A Estratégia Resumptiva em Relativas Restritivas do Português Europeu*. Master Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Alexiadou, Artemis (1997). *Adverb Placement: A Case Study in Antisymmetric Syntax*. Amsterdam: John Benjamins.
- Alexiadou, Artemis and Anagnostopoulou, Elena (2001). The Subject-in-Situ Generalization and the Role of Case Driving Computations. *Linguistic Inquiry*, 32(2), 193-231.
- Alexiadou, Artemis, Law, Paul, Meinunger, André and Wilder, Chris (2000). *The Syntax of Relative Clauses*. Amsterdam/Philadelphia: John Benjamins.
- Ambar, Manuela (1992). *Para uma Sintaxe da Inversão Sujeito-Verbo em Português*. Lisboa: Edições Colibri.
- Ambar, Manuela (1999). Aspects of the Syntax of Focus in Portuguese. In Georges Rebuschi and Laurice Tuller (Eds.), *The Grammar of Focus* (pp. 23-53). Amsterdam/Philadelphia: John Benjamins.
- Ambar, Manuela, Obenauer, Hans, Pereira, Iris, Tapazdi, Judit and Veloso, Rita (1998). From Wh-Questions to Wh-Exclamatives: the Internal Structure of Wh-Phrases and the Left Periphery. Evidence from Portuguese, French and Hungarian. Posted at Chomsky Celebration's site for his 70th anniversary in December 1998.
- Ambar, Manuela and Veloso, Rita (2001). On the Nature of Wh-phrases – Word Order and Wh-in-Situ: Evidence from Portuguese, French, Hungarian and Tetum. In Yves D'Hulst, Johan Rooryck and Jan Schroten (Eds.), *Romance Languages and Linguistic Theory 1999: Selected Papers from 'Going Romance' 1999* (pp. 1-37). Amsterdam/Philadelphia: John Benjamins.
- Andrews, Avery (1975). *Studies in the Syntax of Relative and Comparative Clauses*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Androutsopoulou, Antonia (1997). Reduced Relatives in DPs: Evidence from Adjective Extraction in Modern Greek. *MIT Working Papers in Linguistics*, 31, 19-40.
- Arnold, Doug (2007). Non-restrictive Relatives are not Orphans. *Journal of Linguistics*, 43(2), 271-309.
- Avesani, Cinzia and Vayra, Mario (2003). Broad, Narrow, and Contrastive Focus in Florentine Italian. In Maria Josep Solé, Daniel Recasens and Joaquín Romero (Eds.), *Proceedings of the XVth International Congress of Phonetic Sciences. Volume 2* (pp. 1803-1806). Barcelona: Futurgraphic.
- Baker, Mark (2003). *Lexical Categories: Verbs, Nouns and Adjectives*. Cambridge: Cambridge University Press.
- Baker, Mark (2008). The Macroparameter in a Microparameter World. In Theresa Biberauer (Ed.), *The Limits of Syntactic Variation* (pp. 351-337). Amsterdam/Philadelphia: John Benjamins.
- Baltin, Mark (1984). Extraposition rules and discontinuous constituents. *Linguistic Inquiry*, 15, 157-163.
- Baltin, Mark (2006). Extraposition. In Martin Everaert and Henk van Riemsdijk (Eds.), *The Blackwell Companion to Syntax. Volume 2* (pp. 237-271). Malden, Massachusetts: Blackwell Publishing.
- Barbiers, Sjef (1995). *The Syntax of Interpretation*. The Hague: HAG.

- Barbosa, Pilar (1995). *Null Subjects*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Barbosa, Pilar (2000). Clitics: a Window into the Null Subject Property. In João Costa (Ed.), *Portuguese Syntax. New Comparative Studies* (pp. 31–93). Oxford, New York: Oxford University Press.
- Barbosa, Pilar (2009). Two Kinds of Subject pro. *Studia Linguistica*, 63(1), 2-58.
- Barbosa, Pilar, Duarte, Maria and Kato, Mary (2005). Null Subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, 11-52.
- Barreto, Mário (1911). *Novos Estudos da Língua Portuguesa*. Rio de Janeiro: Livraria Francisco Alves (1980, 3rd ed., Rio de Janeiro: Presença Edições).
- Bašić, Monika (2004). *Nominal Subextractions and the Structure of NPs in Serbian and English*. Master Dissertation, Universitetet i Tromsø.
- Basilico, David (1996). Head Position and Internally Headed Relative Clauses. *Language*, 72, 498-533.
- Bassols de Climent, Mariano (1967). *Sintaxis Latina*. Madrid: Consejo Superior de Investigaciones Cientificas.
- Bastos, Ana Cláudia (2001). *Fazer, eu faço! Topicalização de Constituintes Verbais em Português Brasileiro*. Master Dissertation, Instituto da Linguagem, Universidade Estadual de Campinas.
- Bechara, Evanildo (1961). *Moderna Gramática Portuguesa*. Rio de Janeiro: Companhia Editora Nacional/Editora Lucerna (2001, 37th ed.).
- Bhatt, Rajesh (2002). The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification. *Natural Language Semantics*, 10, 43-90.
- Bianchi, Valentina (1999). *Consequences of Antisymmetry: Headed Relative Clauses*. Berlin: Mouton de Gruyter.
- Bianchi, Valentina (2000). The Raising Analysis of Relative Clauses: A Reply to Borsley. *Linguistic Inquiry*, 31(1), 123-140.
- Bianchi, Valentina (2002). Headed Relative Clauses in Generative Syntax - Part II. *GLOT International*, 6(8), 235-247.
- Borsley, Robert (1997). Relative Clauses and the Theory of Phrase Structure. *Linguistic Inquiry*, 28(4), 629-647.
- Bošković, Željko (2001). *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Elsevier. Amsterdam: Elsevier Science.
- Bošković, Željko (2002). On Multiple Wh-Fronting. *Linguistic Inquiry*, 33, 351-383.
- Bošković, Željko (2004a). On the Clitic Switch in Greek Imperatives. In Olga Tomić (Ed.), *Balkan Syntax and Semantics* (pp. 269-291). Amsterdam: John Benjamins.
- Bošković, Željko (2004b). PF Merger in Stylistic Fronting and Object Shift. In Arthur Stepanov, Gisbert Fanselow and Ralf Vogel (Eds.), *Minimality Effects in Syntax* (pp. 37-71). Berlin: Mouton de Gruyter.
- Bošković, Željko (2005). On the Locality of Left Branch Extraction and the Structure of NP. *Studia Linguistica*, 59(1), 1-45.
- Bošković, Željko and Nunes, Jairo (2007). The Copy Theory of Movement: a View from PF. In Norbert Corver and Jairo Nunes (Eds.), *The Copy Theory of Movement* (pp. 13-74). Amsterdam/Philadelphia: John Benjamins.
- Bowers, John (1988). Extended X-Bar Theory, the ECP, and the Left Branch Condition. In *Proceedings of the West Coast Conference on Formal Linguistics* (pp. 47-62). Stanford Linguistics Association, Stanford University.

- Brame, Michael (1968). A New Analysis of the Relative Clause: Evidence for an Interpretive Theory. Unpublished manuscript, Massachusetts Institute of Technology.
- Brito, Ana Maria (1991). *A Sintaxe das Orações Relativas em Português. Estrutura, Mecanismos Interpretativos e Condições sobre a Distribuição dos Morfemas Relativos*. Porto: Instituto Nacional de Investigação Científica/Centro de Linguística da Universidade do Porto.
- Brito, Ana Maria (1995). As Orações Relativas Restritivas nas Variantes Culta e Oral em Quatro Línguas Românicas, com Especial Incidência em Português. *Lusorama*, 70-81.
- Brito, Ana Maria and Duarte, Inês (2003). Orações Relativas e Construções Aparentadas. In Maria Helena Mateus, Ana Maria Brito, Inês Duarte, Isabel Hub Faria, Sónia Frota, Gabriela Matos, Fátima Oliveira, Marina Vigário and Alina Villalva, *Gramática da Língua Portuguesa* (pp. 653-694). Lisboa: Caminho.
- Brito, Ana Maria, Duarte, Inês and Matos, Gabriela (2003). Estrutura da Frase Simples e Tipos de Frase. In Maria Helena Mateus, Ana Maria Brito, Inês Duarte, Isabel Hub Faria, Sónia Frota, Gabriela Matos, Fátima Oliveira, Marina Vigário and Alina Villalva, *Gramática da Língua Portuguesa*. Lisboa: Caminho.
- Brito, Ana Maria (2004). As Relativas não Restritivas com Antecedente Nominal como um Caso de Aposição. In *Actas do XX Encontro Nacional da Associação Portuguesa de Linguística* (pp. 401-419). Lisboa: APL.
- Broekhuis, Hans (2007). Object Shift and Subject Shift. *The Journal of Comparative Germanic Linguistics*, 10, 109-141.
- Brucart, José María (1999). La Estructura del Sintagma Nominal: Las Oraciones de Relativo. In Ignacio Bosque and Violeta Demonte (Eds.), *Gramática Descriptiva de la Lengua Española* (pp. 395-522). Madrid: Espasa Calpe.
- Büring, Daniel and Hartmann, Katharina (1997). Doing the Right Thing – Extraposition as a Movement Rule. *The Linguistic Review*, 14, 1-42.
- Butler, Alastair and Mathieu, Eric (2004). *The Syntax and Semantics of Split Constructions: A Comparative Study*. Basingstoke, Hampshire: Palgrave Macmillan.
- Cable, Seth (2007). *The Grammar of Q: Q-Particles and the Nature of Wh-Fronting, As Revealed by the Wh-Questions of Tlingit*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Cardinaletti, Anna (1987). Aspetti Sintattici dell'Estraposizione della Frase Relativa. *Rivista di Grammatica Generativa*, 12, 3-59.
- Cardoso, Adriana (2008). Relativas com Núcleo Interno e Relativo de Ligação na História do Português. In *Actas do XXIII Encontro Nacional da Associação Portuguesa de Linguística* (pp. 77-92). Lisboa: APL.
- Cardoso, Adriana (2009). Extraposição de Orações Relativas: Uma Abordagem Comparativa entre o Português Antigo e o Português Actual. In *XXIV Encontro Nacional da Associação Portuguesa de Linguística – Textos Seleccionados* (pp. 169-184). Associação Portuguesa de Linguística: Lisboa.
- Cardoso, Adriana and Vries, Mark de (2010). Internal and External Heads in Appositive Constructions. Ms., University of Lisbon/University of Groningen, Version 2, May 2010.
- Carlson, Greg (1977). Amount Relatives. *Language*, 53, 520-542.
- Carrilho, Ernestina (2005). *Expletive ele in European Portuguese Dialects*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.

- Cecchetto, Carlo (2005). Reconstruction in Relative Clauses and the Copy Theory of Traces. In Pierre Pica and Johan Rooryck (Eds.), *Linguistic Variation Yearbook N. 5* (pp. 73-103). Amsterdam: John Benjamin.
- Chaves, Rui (2007). *Coordinate Structures - Constraint-based Syntax-Semantics Processing*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Chomsky, Noam (1973). Conditions on Transformations. In Stephen Anderson and Paul Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 232-286). New York: Holt, Reinhart and Winston.
- Chomsky, Noam (1977). On Wh-Movement. In Peter Culicover, Thomas Wasow and Adrian Akmajian (Eds.), *Formal Syntax* (pp. 71-132). New York: Academic Press.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1993). A Minimalist Program for Linguistic Theory. In Kenneth Hale and Samuel Keyser (Eds.), *The Review from Building 20* (pp. 1-52). Cambridge, Massachusetts: The MIT Press.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, Massachusetts: The MIT Press.
- Chomsky, Noam (2000). Minimalist Inquiries: The Framework. In Roger Martin, David Michaels and Juan Uriagereka (Eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* (pp. 89-156). Cambridge, Massachusetts: The MIT Press.
- Cinque, Guglielmo (1982). On the Theory of Relative Clauses and Markedness. *The Linguistic Review*, 1, 247-294.
- Cinque, Guglielmo (1990). *Types of A'-Dependencies*. Cambridge, Massachusetts: The MIT Press.
- Cinque, Guglielmo (1994). On the Evidence for Partial N Movement in the Romance DP. In Jan Koster, Jean-Yves Pollock, Luigi Rizzi and Raffaella Zanuttini (Eds.), *Paths Towards Universal Grammar* (pp. 85-110). Georgetown: Georgetown University Press.
- Cinque, Guglielmo (1999). *Adverbs and Functional Heads. A Cross-linguistic Perspective*. New York: Oxford University Press.
- Cinque, Guglielmo (2003). The Prenominal Origin of Relative Clauses. Talk presented at the Workshop on Antisymmetry and Remnant Movement at NYU, October 31 - November 2, 2003.
- Cinque, Guglielmo (2006). *Restructuring and Functional Heads. The Cartography of Syntactic Structures*. Volume 4. Oxford and New York: Oxford University.
- Cinque, Guglielmo (2007). A Note on Linguistic Theory and Typology. *Linguistic Typology*, 11, 93-106.
- Cinque, Guglielmo (2008). Two Types of Nonrestrictive Relatives. In Olivier Bonami and Patricia Cabredo Hofherr (Eds.), *Empirical Issues in Syntax and Semantics 7* (pp. 99-137).
- Cinque, Guglielmo (2009). The Prenominal Origin of Relative Clauses. Talk presented in a Research Seminar at the University of Hong Kong School of Humanities Department of Linguistics, May 5, 2009.
- Cinque, Guglielmo and Kayne, Richard (Eds.) (2005). *The Oxford Handbook of Comparative Syntax*. New York: Oxford University Press.
- Colaço, Madalena (2006). Coordenação e Movimento Sintático: os Dados do Português Europeu. *Letras de Hoje*, 1(1), 75-97.

- Cole, Peter (1987). Internally Headed Relative Clauses. *Natural Language & Linguistic Theory*, 5, 277-302.
- Comrie, Bernard and Kuteva, Tania (2008). Relativization Strategies. In Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (Eds.), *The World Atlas of Language Structures Online* (Chapters 122-123). Munich: Max Planck Digital Library.
- Corominas, Joan and Pascual, José (1980). *Diccionario Crítico Etimológico Castellano e Hispánico*. Madrid: Gredos.
- Corver, Norbert (1990). *The Syntax of Left Branch Extractions*. Doctoral Dissertation, Katholieke Universiteit Brabant, Tilburg.
- Costa, Ana (2003). Construções de Relativização numa Partição de Chelas de 1425. Ms., Faculdade de Letras da Universidade de Lisboa.
- Costa, Ana (2004). Aspectos das Construções de Relativização no Português do Século XV. In *Actas do XIX Encontro da Associação Portuguesa de Linguística* (pp. 409-420). Lisboa: APL.
- Costa, João (1996). Adverb Positioning and V-movement in English: Some more Evidence. *Studia Linguistica*, 50(1), 22-34.
- Costa, João (1998). *Word Order Variation. A Constraint-based Approach*. The Hague: Holland Academic Graphics.
- Costa, João. (2000). Focus in situ: Evidence from Portuguese. *Probus*, 12-2, 187-228.
- Costa, João (2001). Postverbal Subjects and Agreement in Unaccusative Contexts in European Portuguese. *The Linguistic Review*, 18(1), 1-17.
- Costa, João (2004a). *Subject Positions and Interfaces: The Case of European Portuguese*. Berlin, New York: Mouton de Gruyter.
- Costa, João (2004b). A Multifactorial Approach to Adverb Placement: Assumptions, Facts, and Problems. *Lingua*, 114, 711–753.
- Costa, João and Duarte, Inês (2002). Preverbal Subjects in Null Subject Languages are not Necessarily Dislocated. *Journal of Portuguese Linguistics*, 1, 157–175.
- Cuesta, Pilar Vásquez and Luz, Maria Albertina (1971). *Gramática Portuguesa*. Madrid: Gredos (1980, Portuguese translation, *Gramática da Língua Portuguesa*. Lisboa: Edições 70).
- Culicover, Peter and Rochemont, Michael (1990). Extraposition and Logical Form. *Linguistic Inquiry*, 21(1), 23-47.
- Culicover, Peter and Winkler, Susanne (2008). English Focus Inversion Constructions. *Journal of Linguistics*, 44, 625-658.
- Culy, Christopher (1990). *The Syntax and Semantics of Internally Headed Relative Clauses*. Doctoral Dissertation, Stanford University.
- Cunha, Celso and Cintra, Luis Filipe Lindley (1984). *Nova Gramática do Português Contemporâneo*. Lisboa: Ed. J. Sá da Costa. (1997, 17th ed.).
- Del Gobbo, Francesca (2008). On the Syntax and Semantics of Appositive Relative Clauses. In Nicole Dehé and Yordanka Kavalova (Eds.), *Parentheticals* (pp. 173-201). Amsterdam: John Benjamins.
- Delorme, Evelyne and Dougherty, Ray (1972). Appositive NP Constructions. *Foundations of Language*, 8, 2-29.
- Demirdache, Hamida (1991). *Resumptive Chains in Restrictive Relatives, Appositives, and Dislocation Structures*. Doctoral Dissertation, Massachusetts Institute of Technology.

- Devine, A. M. and Stephens, Laurence (1999). *Discontinuous Syntax: Hyperbaton in Greek*. Oxford: Oxford University Press.
- Devine, A. M. and Stephens, Laurence (2006). *Latin Word Order: Structured Meaning and Information*. Oxford: Oxford University Press.
- Dias, Augusto Epifânio da Silva (1933). *Sintaxe Histórica Portuguesa*. Lisboa: Clássica Editora (1970, 5th ed.).
- Diesing, Molly (1992). *Indefinites*. Cambridge, Massachusetts: The MIT Press.
- Drozd, Kenneth (2001). Metalinguistic Sentence Negation in Child English. In Jack Hoeksema, Holtze Rullmann, Victor Sanchez-Valencia and Ton van der Wouden (Eds.), *Perspectives on Negation and Polarity Items* (pp. 49-78).
- Duarte, Inês (1987). *A Construção de Topicalização na Gramática do Português: Regência, Ligação e Condições sobre o Movimento*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Duarte, Inês (1996). A Topicalização em Português Europeu: uma Análise Comparativa. In *Actas do Congresso Internacional sobre o Português. Volume 1* (pp. 327-360). Lisboa: APL-Colibri.
- Duarte, Inês (1997). Ordem de Palavras: Sintaxe e Estrutura Discursiva. In Ana Maria Brito, Fátima Oliveira, Isabel Pires de Lima and Rosa Maria Martelo (Eds.), *Sentido que a Vida Faz: Estudos para Óscar Lopes* (pp. 581-592). Porto: Campo das Letras.
- Duarte, Inês (2000). Sobre Interrogativas-Q em Português Europeu e Português Brasileiro. Talk presented at the international conference *500 anos da Língua Portuguesa no Brasil* at Universidade de Évora, May 8-13, 2000.
- Duarte, Inês (2003). A Família das Construções Inacusativas. In Maria Helena Mateus, Ana Maria Brito, Inês Duarte, Isabel Hub Faria, Sónia Frota, Gabriela Matos, Fátima Oliveira, Marina Vigário and Alina Villalva (Eds.), *Gramática da Língua Portuguesa* (pp. 507-548). Lisboa: Caminho.
- Duarte, Inês and Oliveira, Fátima (2003). Referência Nominal. In Maria Helena Mateus, Ana Maria Brito, Inês Duarte, Isabel Hub Faria, Sónia Frota, Gabriela Matos, Fátima Oliveira, Marina Vigário and Alina Villalva (Eds.), *Gramática da Língua Portuguesa* (pp. 205-242). Lisboa: Caminho.
- Emonds, Joseph (1976). *A Transformational Approach to English Syntax*. New York: Academic Press.
- Emonds, Joseph (1979). Appositive Relatives Have No Properties. *Linguistic Inquiry*, 10(2), 211-243.
- Emonds, Joseph (1985). *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.
- Enkvist, Nils Erik (1980). Marked Focus: Function and Constraints. In Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (Eds.), *Studies in English Linguistics for Randolph Quirk* (pp. 134-131). London: Longman.
- Ernout, Alfred and Thomas, François (1972). *Syntaxe Latine*. Paris: Klincksieck.
- Ernst, Thomas (2002). *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.
- Erteschik-Shir, Nomi (2007). *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press.
- Espinal, M. Teresa (1991). The Representation of Disjunct Constituents. *Language*, 67, 726-762.
- Evans, Gareth (1980). Pronouns. *Linguistic Inquiry*, 11, 337-362.
- Fabb, Nigel (1990). The Difference between English Restrictive and Nonrestrictive Relative Clauses. *Journal of Linguistics*, 26(1), 57-78.

- Fanselow, Gisbert (1988). Aufspaltung von NPn und das Problem der “freien” Wortstellung. *Linguistische Berichte*, 114, 91-113.
- Fanselow, Gisbert and Ćavar, Damir (2002). Distributed Deletion. In Artemis Alexiadou (Ed.), *Theoretical Approaches to Universals* (pp. 65–107). Amsterdam: John Benjamins.
- Fanselow, Gisbert and Féry, Caroline (2006). Prosodic and Morphosyntactic Aspects of Discontinuous Noun Phrases: a Crosslinguistic Perspective. *Linguistics in Postdam*, 26.
- Fiéis, Alexandra and Lobo, Maria (2010). Aspectos da Sintaxe das Orações Gerundivas em Português Medieval e em Português Contemporâneo. In *Textos Seleccionados, XXV Encontro Nacional da Associação Portuguesa de Linguística* (pp. 419-434). Porto: APL.
- Fiengo, Robert (1977). On Trace Theory. *Linguistic Inquiry*, 8, 35-61.
- Finch, James (2006). *Caesar Completely Parsed*. Mundelein, Illinois: Bolchazy-Carducci Publishers.
- Fiorentino, Giuliana (1999). *Relativa Debole. Sintassi, Uso, Storia in Italiano*. (Materiali Linguistici-Università di Pavia 24). Milano: Franco Angeli.
- Fox, Danny (2002). Antecedent-Contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry*, 33(1), 63-96.
- Fox, Danny and Nissenbaum, Jon (1999). Extraposition and Scope: A Case for Overt QR. In Sonya Bird, Andrew Carnie, Jason Haugen and Peter Norquest (Eds.), *Proceedings of the 18th West Coast Conference on Formal Linguistics* (pp. 132-144). Somerville: Cascadilla Press.
- Franks, Steven and Progovac, Ljiljana (1994). On the Displacement of Serbo-Croatian Clitics. *Indiana Linguistic Studies*, 7, 69-78.
- Frota, Sónia (1998). *Prosody and Focus in European Portuguese*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Frota, Sónia (2002). Nuclear Falls and Rises in European Portuguese: a Phonological Analysis of Declarative and Question Intonation. *Probus*, 14, 113-146.
- Giorgi, Alessandra and Longobardi, Giuseppe (1991). *The Syntax of Noun Phrases*. Cambridge: Cambridge University Press.
- Givón, Talmy (2001). *Syntax: An Introduction*. Amsterdam: John Benjamins.
- Grimshaw, Jane (2000). Locality and Extended Projections. In Peter Coopmans, Martins Everaert and Jane Grimshaw (Eds.), *Lexical Specification and Insertion*. Amsterdam: Benjamins.
- Grosu, Alexander (1973). On the Nonunitary Nature of the Coordinate Structure Constraint. *Linguistic Inquiry*, 4, 88-92.
- Grosu, Alexander and Landman, Fred (1998). Strange Relatives of the Third Kind. *Natural Language Semantics*, 6, 125-170.
- Guéron, Jacqueline (1980). On the Syntax and Semantics of PP Extraposition. *Linguistic Inquiry*, 11(4), 637-678.
- Haider, Hubert (1997). Extraposition. In Dorothee Beerman, David LeBlanc and Henk van Riemsdijk (Eds.), *Rightward Movement* (pp. 115–151). Amsterdam: John Benjamins.
- Haider, Hubert (1996). Downright down to the Right. In Uli Lutz and Jürgen Pafel (Eds.), *On Extraction and Extraposition in German* (pp. 245–271). Amsterdam: John Benjamins.

- Hale, Ken (1983). Walpiri and the Grammar of Non-configurational Languages. *Natural Language & Linguistic Theory*, 1, 5-47.
- Hale, William and Buck, Carl (1966). *A Latin Grammar*. Tuscalosa: The University of Alabama.
- Halitsky, David (1974). Deep Structure Appositive and Complement NPs. *Language*, 50, 446-455.
- Halliday, Mak (1967). Notes on Transitivity and Theme in English (Part 2). *Journal of Linguistics*, 3, 199-244.
- Heck, Fabian (2008). *On Pied-Piping. Wh-Movement and Beyond*. Berlin, New York: Mouton de Gruyter.
- Heim, Irene (1987). Where Does the Definiteness Restriction Apply? Evidence from the Definiteness of Variables. In Eric Reuland and Alice ter Meulen (Eds.), *The Representation of (In)definiteness* (pp. 21-42). Cambridge, Massachusetts: The MIT Press.
- Henk, Paula (2010). *Information Structure of Estonian - Compared to Finnish and Hungarian*. Osnabrück: Publications of the Institute of Cognitive Science.
- Heringa, Herman (2007). Appositional Constructions: Coordination and Predication. In Marlies Kluck and Erik-Jan Smits (Eds.), *Proceedings of the Fifth Semantics in The Netherlands Day* (pp. 67-82). Groningen: University of Groningen.
- Heringa, Herman (2009). A Multidominance Approach to Appositional Constructions. Ms., University of Groningen.
- Hernanz, Maria Luisa and Brucart, José María (1987). *La Sintaxis: 1. Principios Teóricos. La Oración Simple*. Barcelona: Crítica.
- Heycock, Caroline (2005). On the Interaction of Adjectival Modifiers and Relative Clauses. *Natural Language Semantics*, 13(4), 359-382.
- Holmberg, Anders and Roberts, Ian (2010). Introduction. In Theresa Biberauer, Anders Holmberg, Ian Roberts and Michelle Sheehan (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 1-57). Cambridge: Cambridge University Press.
- Hoof, Hanneke van (2005). Split Topicalization. In Martin Everaet and Henk van Riemsdijk (Eds.), *The Blackwell Companion to Syntax*. London: Blackwell.
- Hoop, Helen de (1992). *Case Configuration and Noun Phrase Interpretation*. Doctoral Dissertation, University of Groningen.
- Horvath, Julia (2007). Separating "Focus Movement" from Focus. In Simin Karimi, Vida Samiiian and Wendy Wilkins (Eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation. In honor of Joseph E. Emonds* (pp. 108-145). Amsterdam: John Benjamins.
- Hualde, José (2002). Intonation in Romance. Introduction to the special issue. *Probus*, 14, 1-7
- Huber, Joseph (1933). *Altportugiesisches Elementarbuch*. Heidelberg: Carl Winters (1986, Portuguese translation, *Gramática do Português Antigo*. Lisboa: Fundação Calouste Gulbenkian).
- Huddleston, Rodney, Pullum, Geoffrey and Peterson, Peter (2002). Relative Constructions and Unbounded Dependencies. In *The Cambridge Grammar of the English Language* (pp. 1031-1096). Cambridge: Cambridge University Press.
- Inaba, Jiro (2005). Extraposition and the Directionality of Movement. In Sylvia Blaho, Luis Vicente and Erik Schoorlemmer (Eds.), *Proceedings of ConSOLEXIII* (pp. 157-169). Leiden University.
- Inada, Shun'ichiro (2007). Towards a Syntax of Two Types of Relative Clauses. *Linguistic Research: Working Papers in English Linguistics*, 23, 1-41.

- Jackendoff, Ray (1977). *X'-Syntax: A Study of Phrase Structure*. Cambridge, Massachusetts: The MIT Press.
- Jackendoff, Ray (1990). On Larson's Treatment of the Double Object Construction. *Linguistic Inquiry*, 21(3), 427-456.
- Jelinek, Eloise (1984). Empty Categories, Case, and Configurationality. *Natural Language & Linguistic Theory*, 2, 39-76.
- Jespersen, Otto (1949). *A Modern English Grammar on Historical Principles*. (Syntax. Volume 2). London: Allen and Unwin Ltd.
- Johannessen, Janne (1998). *Coordination*. Oxford: Oxford University Press.
- Kariaeva, Natalia (2009). *Radical Discontinuity: Syntax at the Interface*. Doctoral Dissertation, Rutgers University.
- Kato, Mary and Nunes, Jairo (2009). A Uniform Raising Analysis for Standard and Nonstandard Relative Clauses in Brazilian Portuguese. In Jairo Nunes (Ed.), *Minimalist Essays on Brazilian Portuguese Syntax* (pp. 93-120). Amsterdam: John Benjamins.
- Kato, Mary and Raposo, Eduardo Paiva (2007). Topicalization in European and Brazilian Portuguese. In José Camacho, Nydia Flores, Liliane Sánchez, Viviane Déprez and Maria Jose Cabrera (Eds.), *Romance Linguistics: Selected Papers from the 36th Linguistic Symposium on Romance Languages* (pp. 213-226). Amsterdam: John Benjamins.
- Kayne, Richard (1984). *Connectedness and Binary Branching*. Dordrecht: Foris.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, Massachusetts: The MIT Press.
- Kiss, Katalin (1998). Identificational Focus versus Information Focus. *Language*, 74(2), 245-273.
- Kiss, Tibor (2002). Phrasal Typology and the Interaction of Topicalization, Wh-Movement and Extraposition. In Jong-Bok Kim and Stephen Wechsler (Eds.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar* (pp. 109-128). Stanford: CSLI Publications.
- Kiss, Tibor (2005). Semantic Constraints on Relative Clause Extraposition. *Natural Language & Linguistic Theory*, 23(2), 281-334.
- Klein, Maarten (1977). *Appositionele Constructies in het Nederlands*. Doctoral Dissertation, Radboud University Nijmegen.
- Koopman, Hilda and Sportiche, Dominique (1991). The Position of Subjects. *Lingua*, 85, 211-258.
- Koster, Jan (1987). *Domains and Dynasties. The Radical Autonomy of Syntax*. Dordrecht: Foris Publication.
- Koster, Jan (1995). Extraposition as Coordination. Paper presented at Colloquium Max-Planck-Institut and Humboldt University, Berlin, February 9, 1995.
- Koster, Jan (2000). Extraposition as Parallel Construal. Ms., University of Groningen.
- Krifka, Manfred (2007). Basic Notions of Information Structure. In Caroline Féry, Gisbert Fanselow and Manfred Krifka (Eds.), *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6* (pp. 13-56). Potsdam: Universitätsverlag Potsdam.
- Kroch, Anthony (1989). Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change*, 1, 199-244.

- Kroch, Anthony (1994). Morphosyntactic Variation. In Katharine Beals, Jeannette Denton, Robert Knippen, Lynette Melmar, Hisame Suzuki and Erica Zeinfeld (Eds.), *Proceedings of the Thirtieth Annual Regional Meeting of the Chicago Linguistics Society* (pp. 180–201). Chicago: Chicago Linguistics Society.
- Kroch, Anthony (2001). Syntactic Change. In Mark Baltin and Chris Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (pp. 699-729). Malden, Massachusetts and Oxford: Blackwell Publishing.
- Kunstmann, Pierre (1991). Création et Diffusion du Relatif/Interrogative *lequel* en Ancien Français. Comparaison avec d'autres Langues Romanes. In Dieter Kremer (Ed.), *Actes du XVIIIe Congrès International de Linguistic Romane* (pp. 660-670). Tübingen: Max Niemeyer.
- Kuroda, S.-Y. (1965). *Generative Grammatical Studies in the Japanese Language*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Kuroda, S.-Y. (1972). The Categorical and the Thetic Judgments. *Foundations of Language*, 9, 153-185.
- Kuroda, S.-Y. (1992). *Japanese Syntax and Semantics: Collected Papers*. Dordrecht: Kluwer.
- Kuroda, S.-Y. (2005). Focusing on the Matter of Topic: A Study of Wa and Ga in Japanese. *Journal of East Asian Linguistics*, 14(1), 1-58.
- Lambova, Mariana (2004). On Triggers of Movement and Effects at the Interfaces. In Anne Breitbarth and Henk van Riemsdijk (Eds.), *Triggers* (pp. 231-258). Berlin: Mouton de Gruyter.
- Lambrecht, Knud (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Landau, Idan (2007). EPP Extensions. *Linguistic Inquiry*, 38(3), 485-523.
- Larson, Richard (1988). On the Double Object Construction. *Linguistic Inquiry*, 19(3), 335-391.
- Larson, Richard (1990). Double Objects Revisited: Reply to Jackendoff. *Linguistic Inquiry*, 21(4), 589-632.
- Lasnik, Howard and Sobin, Nicholas (2000). The who/whom Puzzle: on the Preservation of an Archaic Feature. *Natural Language & Linguistic Theory*, 18, 343–371.
- Lausberg, Heinrich (1967). *Elemente der Literarischen Rhetorik*. München: Max Hueber (1972, 2nd ed., Portuguese translation, *Elementos de Retórica Literária*. Lisboa: Fundação Calouste Gulbenkian).
- Lee, Seong-yong (2007). Sideward Movement Approach to Extraposition from DP in English. *Korean Journal of English Language and Linguistics*, 7(1), 27-49.
- Lehmann, Christian (1984). *Der Relativsatz*. Tübingen: Gunter Narr Verlag.
- Lightfoot, David (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, Massachusetts: The MIT Press.
- Lightfoot, David (1999). *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.
- Lobo, Maria (2003). *Aspectos da Sintaxe das Orações Subordinadas Adverbiais do Português*. Doctoral Dissertation, Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa.
- Loock, Rudy (2007). Appositive Relative Clauses and their Functions in Discourse. *Journal of Pragmatics*, 39(2), 336-362.
- López, Luis (2000). Ellipsis and Discourse Linking. *Lingua*, 110, 183-213.

- Lucchesi, Dante (1990). Considerações sobre a Análise das Relativas do Português Contemporâneo e algumas Incursões na História dessas Estruturas. In *Actas do VI Encontro da Associação Portuguesa de Linguística* (pp. 175-193). Porto: APL.
- Maia, Clarinda (1986). *História do Galego Português. Estado Linguístico da Galiza e do Noroeste de Portugal desde o Século XIII ao Século XVI (com referência à situação do galego moderno)*. Coimbra: Instituto Nacional de Investigação Científica.
- Martin, Larry (1972). *Appositive and Restrictive Relative Clauses in English*. Doctoral Dissertation, University of Texas at Austin.
- Martins, Ana Maria (1994). *Clíticos na História do Português*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Martins, Ana Maria (1999). Ainda ‘os mais antigos textos escritos em português’: Documentos de 1175 a 1252. In Isabel Hub Faria (Ed.) *Lindley Cintra: Homenagem ao Homem, ao Mestre e ao Cidadão*, (pp. 491-534). Lisboa: Cosmos and Faculdade de Letras da Universidade de Lisboa.
- Martins, Ana Maria (2002). The Loss of IP Scrambling in Portuguese: Clause Structure, Word Order Variation and Change. In David Lightfoot (Ed.), *Syntactic Effects of Morphological Change* (pp. 232-248). Oxford/New York: Oxford University Press.
- Martins, Ana Maria (2004). A Emergência do Português Escrito na Segunda Metade do século XII. In Rosario Álvarez and Antón Santamarina (Eds.), *(Dis)cursos da Escrita: Estudos de Filoloxía Galega Ofrecidos en Memoria de Fernando R. Tato Plaza* (pp. 491-526). A Coruña: Fundación Pedro Barrié de la Maza.
- Martins, Ana Maria (2008). Investigating Language Change in a Comparative Setting. In Maria Clotilde Almeida, Bernd Sieberg and Ana Maria Bernardo (Eds.), *Questions on Language Change* (pp. 99-116). Lisboa: Colibri/Centro de Estudos Alemães e Europeus.
- Martins, Ana Maria (in prep.). Constituent Order in Simple Declarative Sentences. Ms. to be published in: Mary Kato, Ana Maria Martins and Jairo Nunes (Eds.), *The Syntax of Portuguese*. Cambridge: Cambridge University Press.
- Mathieu, Eric (2004). Hyperbaton and Haplology. In Anne Breitbarth and Henk Van Riemsdijk (Eds.), *Triggers* (pp. 293-330). Berlin: Mouton de Gruyter.
- Mathieu, Eric and Sitaridou, Ioanna (2005). Split Wh-Constructions in Classical and Modern Greek: A Diachronic Perspective. In Montserrat Batllori and Francesc Roca (Eds.), *Grammaticalization and Parametric Change* (pp. 236-250). Oxford: Oxford University Press.
- Matos, Gabriela (1992). *Construções de Elipse do Predicado em Português. SV Nulo e Despojamento*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Matos, Gabriela (1995). Estruturas Binárias e Monocêntricas em Sintaxe — algumas Observações sobre a Coordenação de Projecções Máximas. In *Actas do X Encontro da Associação Portuguesa de Linguística* (pp. 301-315). Lisboa: APL and Colibri.
- Matos, Gabriela (2003). Estruturas de Coordenação. In Maria Helena Mateus, Ana Maria Brito, Inês Duarte, Isabel Hub Faria, Sónia Frota, Gabriela Matos, Fátima Oliveira, Marina Vigário and Alina Villalva, *Gramática da Língua Portuguesa* (pp. 549-592). Lisboa: Caminho.
- Mattos e Silva, Rosa Virgínia (1989). *Estruturas Trecentistas: Elementos para uma Gramática do Português Arcaico*. Lisboa: Imprensa Nacional - Casa da Moeda.

- Meinunger, André (2000). *Syntactic Aspects of Topic and Comment*. Amsterdam: John Benjamins.
- Middleton, Roberta (1999). *Aspects of the History of Relative Clauses in Italo Romance*. Doctoral Dissertation, University of Oxford.
- Milner, Jean-Claude (1978). *De la Syntaxe à l'Interprétation*. Paris: Seuil.
- Milsark, Gary (1974). *Existential Sentences in English. Generative Grammatical Studies in the Japanese Language*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Móia, Telmo (2000). *Identifying and Computing Temporal Locating Adverbials with a Particular Focus on Portuguese and English*. Doctoral Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Moro, Andrea (2003). Notes on Vocative Case: A Case Study in Clause Structure. In Josep Quer, Jan Schroten, Mauro Scorretti, Petra Sleeman, and Els Verheugd (Eds.), *Romance Languages and Linguistic Theory 2001: Selected Papers from 'Going Romance'*, Amsterdam, 6-8 December 2001 (pp. 247-261). Amsterdam/Philadelphia: John Benjamins.
- Munn, Alan. (1993). *Topics in the Syntax and Semantics of Coordinate Structures*. Doctoral Dissertation, University of Maryland, College Park.
- Nanni, Debbie and Stillings, Justine (1978). Three Remarks on Pied Piping. *Linguistic Inquiry*, 9, 310-318.
- Neto, Serafim da Silva (1957). *História da Língua Portuguesa*. Rio de Janeiro: Presença/MEC (1970, 2nd ed.).
- Nunes, Jairo (1999). Linearization of Chains and Phonetic Realization of Chains Links. In Samuel Epstein and Norbert Hornstein (Eds.), *Working Minimalism* (pp. 217-249). Cambridge, Massachusetts: The MIT Press.
- Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. Cambridge, Massachusetts: The MIT Press.
- O'Connor, Kathleen (2008). *Aspects de la Syntaxe et de l'Interprétation de l'Apposition à Antécédent Nominal*. Doctoral Dissertation, Université Charles de Gaulle – Lille 3.
- Ott, Denis (2009). Multiple NP-Split: a Distributed Deletion Analysis. *Groninger Arbeiten zur Germanistischen Linguistik*, 48, 65–80.
- Peres, João Andrade and Móia, Telmo (1995). *Áreas Críticas da Língua Portuguesa*. Lisboa: Caminho.
- Perzanowski, Dennis (1980). Appositive Relatives Do Have Properties. *NELS X*, 355-368.
- Pesetsky, David (1995). *Zero Syntax*. Cambridge/Mass.: MIT Press.
- Philips, Colin (1996). *Order and Structure*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Pinkster, Harm (2005). Changing Patterns of Discontinuity in Latin. Paper presented at *Latling: 13e Colloque International de Linguistique Latine*, Brussels, April 4-9, 2005.
- Platzack, Christer (2000). A Complement-of-N Account of Restrictive and Non-Restrictive Relatives: The Case of Swedish. In Artemis Alexiadou, Paul Law, André Meinunger and Chris Wilder (Eds.), *The Syntax of Relative Clauses* (pp. 265-308). Amsterdam: John Benjamins.
- Posner, Rebecca (1996). *The Romance Languages*. Cambridge: Cambridge University Press.

- Predolac, Nikola (2009). Information Structure and Serbian bipartite NPs. In *Proceedings of Formal Approaches to Slavic Linguistics 18 – FASL 18*. Ann Arbor: Michigan Slavic Publications.
- Pollock, Jean-Yves (1989). Verb Movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry*, 20, 365-424.
- Pollock, Jean-Yves (1994). Notes on Checking Theory, pro-drop, Free Variation and Economy. Paper presented at the Going Romance VIII, Utrecht.
- Quer, Josep (2001). Interpreting Mood. *Probus*, 13, 81–111.
- Quer, Josep (1998). *Mood at the Interface*. Utrecht: LOT.
- Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey and Svartvik, Jan (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Ramat, Anna (2005). Persistence and Renewal in the Relative Pronoun Paradigm: the Case of Italian. *Folia Linguistica Historica*, XXVI, 115-138.
- Raposo, Eduardo Paiva (1995). Próclise, Ênclise e a Posição do Verbo em Português Europeu. In João Camilo dos Santos and Frederick G. Williams (Eds.), *O Amor das Letras e das Gentes. In Honor of Maria de Lourdes Belchior Pontes*. Santa Barbara: Center for Portuguese Studies, University of California at Santa Barbara.
- Rebuschi, Georges (2001). Semi-Free Relative Clauses and the DP Hypothesis: Basque Evidence and Theoretical Consequences. In Alex Grosu (Ed.), *Proceedings of The Israel Association for Theoretical Linguistics 8* (pp. 55-64). Tel Aviv University.
- Reinhart, Tanya (1980). On the Position of Extraposed Clauses. *Linguistic Inquiry*, 2(3), 621-624.
- Reinhart, Tanya (1982). *Pragmatics and Linguistics: An Analysis of Sentence Topics*. Bloomington, Indiana: Indiana University Linguistics Club.
- Reinhart, Tanya (1995). Interface Strategies. In *OTS Working Papers in Linguistics* (pp. OTS-WP-TL-95-002). Utrecht: Utrecht Institute of Linguistics.
- Remberger, Eva-Maria (2010). Left Peripheral Interactions in Romance. Workshop on *Focus, Contrast and Givenness in Interaction with Extraction and Deletion*, Universitat Tübingen, March 26-27, 2010.
- Riemsdijk, Henk van (1989). Movement and Regeneration. In Paola Beninca (Ed.), *Dialectal Variation and the Study of Theory of Grammar* (pp. 105-136). Dordrecht: Foris.
- Rizzi, Luigi (1997). The Fine Structure of the Left Periphery. In Liliane Haegeman (Ed.), *Elements of Grammar. Handbook in Generative Syntax*. (pp. 281-337). Dordrecht: Kluwer.
- Rooth, Mats (1992). A Theory of Focus Interpretation. *Natural Language Semantics*, 1, 75-116.
- Ross, John (1967). *Constraints on Variables in Syntax*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Sachs, Wulf (1937). *Black Hamlet*. London: Geoffrey Bles.
- Safir, Ken (1986). Relative Clauses in a Theory of Binding and Levels. *Linguistic Inquiry*, 17, 663-689.
- Said Ali, Manuel (1931). *Gramática Histórica da Língua Portuguesa*. (2ª ed. melhorada e aumentada de *Lexeologia e Formação de Palavras* (1921) e *Syntaxe do Portuguez Historico* (1923)). S. Paulo: Melhoramentos (1971, 7th ed.).
- Salzmann, Martin (2006). *Resumptive Prolepsis: A Study in Indirect A'-Dependencies*. Utrecht: LOT.
- Schachter, Paul (1973). Focus and Relativization. *Language*, 53, 19-49.

- Schönefeld, Doris (2006). *Constructions*. SV1-1/2006. Retrieved September 1, 2010, from www.constructions-online.de, urn:nbn:de:0009-4-6678, ISSN 1860-2010.
- Sekerina (1999). The Scrambling Complexity Hypothesis and Processing of Split Scrambling Constructions in Russian. *Journal of Slavic Linguistics*, 7(2), 265-304.
- Smith, Carlota (1964). Determiners and Relative Clauses in a Generative Grammar of English. *Language*, 40, 37-52.
- Smits, Rik (1988). *The Relative and Cleft Constructions of the Germanic and Romance Languages*. Dordrecht: Foris Publications.
- Sobin, Nicholas (1997). Agreement, Default Rules, and Grammatical Viruses. *Linguistic Inquiry*, 28, 318-343.
- Sòla, Jaume (1992). *Agreement and Subjects*. Doctoral Dissertation, Universitat Autònoma de Barcelona.
- Stassen, Leon (2008). Zero Copula for Predicate Nominals. In Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (Eds.), *The World Atlas of Language Structures Online* (Chapter 120). Munich: Max Planck Digital Library.
- Stjepanović, Sandra (2007). Free Word Order and Copy Theory of Movement. In Norbert Corver and Jairo Nunes (Eds.), *The Copy Theory of Movement* (pp. 219-248). Amsterdam: John Benjamins.
- Stone, Jon R. (2005). *The Routledge Dictionary of Latin Quotations; the Illiterati's Guide to Latin Maxims, Mottoes, Proverbs, and Sayings*. New York/London: Routledge.
- Strunk, Jan (2007). Extraposition Mythbusting. Talk presented at the *Stanford Syntax Workshop*, Stanford University, March 13, 2007.
- Suñer, Margarita (2001). The Puzzle of Restrictive Relative Clauses with Conjoined DP Antecedents. In Julia Herschensohn, Enrique Mallén and Karen Zagona (Eds.), *Features and Interfaces in Spanish and French: Essays in Honor of Heles Contreras* (pp. 267-278). Amsterdam: John Benjamins.
- Takami, Ken-Ichi (1998). A Functional Constraint on Extraposition from NP. In Akio Kamio and Ken-Ichi Takami (Eds.), *Function and Structure. In Honor of Susumu Kuno (Pragmatics & Beyond. New Series 59)* (pp. 23-56). Amsterdam: John Benjamins.
- Takano, Yuji (1998). Object Shift and Scrambling. *Natural Language & Linguistic Theory*, 16(4), 817-889.
- Takano, Yuji (2003). How Antisymmetric is Syntax? *Linguistic Inquiry*, 34, 516-526
- Temperley, David (2003). Ambiguity Avoidance in English Relative Clauses. *Language*, 79(3), 464-484.
- Tognini-Bonelli, Elena (2001). *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Truckenbrodt, Hubert (1995). *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Truswell, Robert (forthcoming). Relatives with a Leftward Island in Early Modern English. *Natural Language & Linguistic Theory*, Online first.
- Vergnaud, Jean-Roger (1974). *French Relative Clauses*. Doctoral Dissertation, Massachusetts Institute of Technology.
- Vergnaud, Jean-Roger (1985). *Dépendances et Niveaux de Représentation en Syntaxe*. Amsterdam: John Benjamins.
- Verstraete, Jean-Christophe (2005). Two Types of Coordination in Clause Combining. *Lingua*, 115(4), 611-626.
- Vries, Gertrud de (1992). *On Coordination and Ellipsis*. Doctoral Dissertation, Catholic University of Brabant.

- Vries, Mark de (1999). Extraposition of Relative Clauses as Specifying Coordination. In Tina Cambier-Langeveld, Anikó Lipták, Michael Redford and Erik Jan van der Torre (Eds.), *Proceedings of ConSole VII* (pp. 293-309). Leiden: SOLE.
- Vries, Mark de (2002). *The Syntax of Relativization*. Utrecht: LOT.
- Vries, Mark de (2004). Head Internal Relative Clauses in Dutch? *Linguistics in the Netherlands*, 21, 193-204.
- Vries, Mark de (2005). Coordination and Syntactic Hierarchy. *Studia Linguistica*, 59, 83-105.
- Vries, Mark de (2006a). Possessive Relatives and (Heavy) Pied Piping. *Journal of Comparative Germanic Linguistics*, 9, 1-52.
- Vries, Mark de (2006b). The Syntax of Appositive Relativization. On Specifying Coordination, False Free Relatives and Promotion. *Linguistic Inquiry*, 37, 229-270.
- Vries, Mark de (2007). Invisible Constituents? Parentheses as B-Merged Adverbial Phrases. In Nicole Dehé and Yordanka Kavalova (Eds.), *Parentheticals* (pp. 203-234). Amsterdam: John Benjamins.
- Vries, Mark de (2009). Specifying Coordination: An Investigation into the Syntax of Dislocation, Extraposition and Parenthesis. In Cynthia R. Dreyer (Ed.), *Language and Linguistics: Emerging Trends* (pp. 37-98). New York: Nova.
- Webelhuth, Gert (1992). *Principles and Parameters of Syntactic Saturation*. Oxford/New York: Oxford University Press.
- Werth, Paul (1974). Some Thought on Non-restrictive Relatives. *Linguistics*, 142, 33-67.
- Wilder, Chris (1994). Coordination, ATB and Ellipsis. *Groninger Arbeiten zur Germanistischen Linguistik*, 37, 291-329.
- Wilder, Chris (1995). Rightward Movement as Leftward Deletion. In Uli Lutz and Jürgen Pafel (Eds.), *Extraction and Extraposition in German* (pp. 273-309). Amsterdam: John Benjamins.
- Zhang, Niina Ning (2007). The Syntactic Derivations of Split Antecedent Relative Clause Constructions. *Taiwan Journal of Linguistics*, 5, 1, 19-48.
- Zimmermann, Malte (2007). Contrastive Focus. In Caroline Féry, Gisbert Fanselow and Manfred Krifka (Eds.), *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure 6* (pp. 147-160). Potsdam: Universitätsverlag Potsdam.
- Ziv, Yael and Cole, Peter (1974). Relative Extraposition and the Scope of Definite Descriptions in Hebrew and English. In Michael La Galy, Robert Fox and Anthony Bruck (Eds.), *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society, April 19--21 Chicago Linguistic Society, Chicago* (pp. 772-786). Chicago: Chicago Linguistic Society.
- Zubizarreta, Maria Luisa (1998). *Prosody, Focus and Word Order*. Cambridge, Massachusetts: The MIT Press.
- Zubizarreta, María Luisa (1999). Las Funciones Informativas: Tema y Foco. In Ignacio Bosque and Violeta Demonte (Eds.), *Gramática Descriptiva de la Lengua Española* (pp. 4215-4244). Madrid: Espasa Calpe.
- Zumpt, Karl (1832). *A Grammar of the Latin Language*. London: B. Fellows.

Editions

- Brocardo, Maria Teresa (1997). *Crónica do Conde D. Pedro de Meneses - Gomes Eanes de Zurara / Edição e Estudo*. Lisboa: FCG/JNICT.
- Brocardo, Maria Teresa (2006). *Livro de Linhagens do Conde D. Pedro. Edição do Fragmento Manuscrito da Biblioteca da Ajuda (século XIV)*. Lisboa: Imprensa Nacional-Casa da Moeda.
- Camões, José (Coord.). (1999). *Gil Vicente: Todas as Obras* [CD-ROM] Lisboa: CNCDP.
- Coelho, António Borges (1987). *Inquisição de Évora: dos Primórdios a 1668*. Lisboa: Caminho.
- Macchi, Giuliano (1975). *Crónica de D. Fernando - Fernão Lopes*. Lisboa: Imprensa Nacional-Casa da Moeda.
- Martins, Ana Maria (2001). *Documentos Portugueses do Noroeste e da Região de Lisboa*. Lisboa: Imprensa Nacional Casa da Moeda.
- Muhana, Adma (1995). *Os Autos do Processo de Vieira na Inquisição*. São Paulo: UNESP.
- Pereira, Isaiás da Rosa (1987). *Documentos para a História da Inquisição em Portugal: Século XVI*. Lisboa: n.p.
- Piel, Joseph (Ed.) (1948). *Livro dos Ofícios de Marco Tullio Ciceram, o qual Tornou em Linguagem o Ifante D. Pedro, Duque de Coimbra. Edicao critica, segundo o MS. de Madrid, prefaciada, anotada e acompanhada de glossario, por Joseph M. Piel*. Coimbra: Universidade de Coimbra.
- Piel, Joseph and Nunes, Irene (1988). *Demanda do Santo Graal*. Lisboa: Imprensa Nacional-Casa da Moeda.
- Ramos, Emanuel (1982). *Os Lusíadas de Luís de Camões*. Porto: Porto Editora.

Corpora

- Bacelar do Nascimento, Maria Fernanda (Coord.) *Corpus de Referência do Português Contemporâneo (CRPC)*. Centro de Linguística da Universidade de Lisboa. http://www.clul.ul.pt/index.php?option=com_content&view=article&id=183%3Areference-corpus-of-contemporary-portuguese-crpc&catid=91%3Aanagrama&lang=pt.
- Bacelar do Nascimento, Maria Fernanda (Portuguese Coord.) (2005). *Corpus of Spoken Portuguese*. In Emanuela Cresti and Massimo Monegli (Eds.), *C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: John Benjamins.
- Galves, Charlotte (Coord.). *Tycho Brahe Parsed Corpus of Historical Portuguese (TBCHP)*. Universidade de Campinas (IEL-UNICAMP). <http://www.tycho.iel.unicamp.br/~tycho/corpus/en/index.html>.
- Davies, Mark and Ferreira, Michael (2006-). *Corpus do Português*. <http://www.corpusdoportugues.org>.
- Davies, Mark (2008-). *The Corpus of Contemporary American English (COCA)*. <http://www.americancorpus.org>.
- Kroch, Anthony, Santorini, Beatrice and Diertani, Ariel (2004). *Penn-Helsinki Parsed Corpus of Early Modern English*. <http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html>.

- Kroch, Anthony, Santorini, Beatrice and Diertani, Ariel (2010). *Penn Parsed Corpus of Modern British English*. <http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html>.
- Marquilhas, Rita (Coord.) *Unknown Letters (CARDS - Cartas Desconhecidas)*. Centro de Linguística da Universidade de Lisboa. <http://alfclul.clul.ul.pt/cards-fly>.
- Martins, Ana Maria (Coord.) *Syntax-oriented Corpus of Portuguese Dialects (CORDIAL-SIN)*. Centro de Linguística da Universidade de Lisboa. <http://www.clul.ul.pt/en/research-teams/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects>.
- Rocha, Paulo Alexandre and Santos, Diana (2000). CETEMPúblico: Um Corpus de Grandes Dimensões de Linguagem Jornalística Portuguesa. In Maria das Graças Volpe Nunes (Ed.), *V Encontro para o Processamento Computacional da Língua Portuguesa Escrita e Falada (PROPOR 2000)* (<http://www.linguateca.pt/CETEMPUBLICO/>), pp. 131-140). São Paulo: ICMC/USP.
- Xavier, Maria Francisca (Coord.). *Digital Corpus of Medieval Portuguese (CIPM – Corpus Informatizado do Português Medieval)*. Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa. <http://cipm.fcsh.unl.pt/>.