

## Léxico e Sintaxe do Verbo no Português Medieval\*

*M. Francisca Xavier, M. de Lourdes Crispim,  
M. da Graça Vicente, M. Alexandra Fiéis,  
M. Cristina Silva, Maria Lobo, Ana Castro,  
M. de Fátima Martins, João Loureiro,  
Sandro Dias, Ana G. Reis*  
Centro de Linguística da Universidade Nova de Lisboa

O projecto tem como objectivo contribuir para um melhor conhecimento da origem e da história da língua portuguesa, desde o período pré-histórico até aos finais do século XV, através do estudo do léxico e da estrutura da frase, em particular, da sintaxe do verbo, numa perspectiva diacrónica e comparada.

Esta investigação desenvolve-se em três domínios complementares: (i) a construção de um extenso *corpus* informatizado de textos latinos dos séculos IX-XII e portugueses dos séculos XII-XVI, o CIPM; (ii) a adaptação e aplicação de ferramentas informáticas para extracção de informação relevante para a análise linguística<sup>1</sup>; (iii) os estudos linguísticos diacrónicos, incluindo a realização de um dicionário histórico, o *Dicionário do Português Medieval*.

Apresenta-se uma síntese dos resultados obtidos através de estudos diacrónicos das diferentes ordens de constituintes da frase, os quais têm vindo a ser desenvolvidos com base na análise sistemática das estruturas em que ocorrem verbos de diferentes classes. Mostrar-se-á, em particular, a cronologia respeitante ao desaparecimento gradual, nos textos antigos, das construções com a ordem Vinfinitivo Vfinito bem como das construções em que o pronome clítico em próclise está separado do verbo por um ou mais constituintes.

No que respeita ao dicionário, apresenta-se a metodologia adoptada para a sua elaboração e ilustram-se as fichas lexicográficas das entradas de cada domínio (verbos, nomes e termos) que está a ser desenvolvido nas respectivas bases de dados – DVPM, DNPM e Jurix<sup>2</sup>.

---

\* O texto procura transcrever a apresentação feita em poster ao XVII Encontro Nacional da Associação Portuguesa de Linguística.

Este projecto é financiado pela Fundação para a Ciência e a Tecnologia – MCT e POCTI.

<sup>1</sup> Os projectos para adaptação e utilização de ferramentas informáticas são desenvolvidos em colaboração com Gabriel Lopes, Nuno Marques, Vitor Rocio e Joaquim Ferreira da Silva do Centro de Inteligência Artificial da FCT-UNL.

<sup>2</sup> As bases de dados e a *homepage* do projecto foram concebidas por Luís Reis, com *design* gráfico de Miguel Chaves e encontram-se disponíveis em <http://cipm.fcsh.unl.pt>.

## • CIPM – *Corpus Informatizado do Português Medieval*

A criação de um Corpus Informatizado de Textos Portugueses Medievais – o CIPM – teve início em 1993. Este corpus inclui textos latinos do século IX ao século XII e textos portugueses dos séculos XII a XVI. Integram o CIPM textos de diferentes tipos – documentos notariais, particulares, régios, administrativos, textos didáctico-moralísticos, narrativos e poéticos, já publicados ou fornecidos em suporte digital pelos próprios editores.

Os textos do CIPM foram sujeitos a um tratamento segundo uma metodologia criada com vista a uniformizar as anotações relativas a referências e comentários e a normas de transcrição.

### Metodologia

A metodologia seguida na preparação dos textos que integram o CIPM compreende essencialmente três fases:

- (i) digitalização e correcção;
- (ii) introdução de anotações de acordo com o conjunto de referências e comentários do CIPM;
- (iii) adaptação das normas de transcrição dos editores às normas do CIPM.

### Constituição do CIPM

Textos latinos dos Séculos IX a XII	Nº. Palavras
Textos notariais in <i>Portugaliae Monumenta Historica a Saeculo Octavo post Christum unque ad Quintum Decimum – Diplomata et Chartae</i>	
<b>TOTAL</b>	<b>478 552</b>

Textos portugueses dos Séculos XII e XIII	Nº. Palavras
NT-Notícia de Torto (1214?)	778
TL e TT -Testamento de D. Afonso II (1214): Ms L e Ms T	2 868
CA-Chancelaria de D. Afonso III (1255)	17 937
DN-Textos notariais (sd,1243-1300) in <i>Docs. Notariais dos Séculos XII a XVI</i>	27 302
CHP-Textos notariais in <i>Clíticos na História do Português</i>	922
HGP-Textos notariais (1262–1300) in <i>História do Galego-Português</i>	28 221
TOX-Textos notariais do Arquivo de Textos do Português Antigo (Oxford)	7 203
TP-Tempos dos Preitos (1280?)	1 594
FR-Foro Real (1280?)	49 736
FG-Foros de Garvão (1280?)	6 603
VS-Vidas de Santos (sd) (cópias tardias de um ms. alcobacense – século XV)	29 566
<b>TOTAL</b>	<b>172 730</b>

## CEM-Cantigas de Escarneo e Maldizer dos séculos XIII e XIV (em preparação)

Textos portugueses do Século XIV	Nº. Palavras
HGP-Textos notariais (1301-1399) in <i>História do Galego-Português</i>	33 617
DN-Textos notariais (1304-1397) in <i>Docs. Notariais dos Séculos XII a XVI</i>	44 746
FG-Foros de Garvão (sd)	1 159
PP-Afonso X, <i>Primeyra Partida</i> (1350?)	170 300
CS-Dos Costumes de Santarém (1340-1360)	30 923
TOX-Textos Notariais do Arquivo de Textos do Português Antigo (Oxford)	4 265
CGE-Crónica Geral de Espanha de 1344 (sd)	426 027
CAXL-Crónica de Afonso X (Ms L) (sd) in <i>Crónica Geral de Espanha de 1344</i>	5 404
CAXP-Crónica de Afonso X (MsPP) (sd) in <i>Crónica Geral de Espanha de 1344</i>	13 923
<b>TOTAL</b>	<b>730 364</b>

## DSG-A Demanda do Santo Graal (sd) (cópia do século XV) (em preparação)

Textos portugueses do Século XV	Nº. Palavras
HGP-Textos notariais (1401-1500) in <i>História do Galego-Português</i>	29 964
DN-Textos notariais (1402-1498) in <i>Docs. Notariais dos Séculos XII a XVI</i>	39 647
ZPM-Crónica do Conde D. Pedro de Meneses (sd)	136 021
HRP-História dos Reis de Portugal (sd) in <i>Crónica Geral de Espanha de 1344</i>	4 996
LTV-O Livro das Tres Vertudes (1447/1453?)	56 272
OE-Orto do Esposo (sd)	105 276
CP-Castelo Perigoso (sd)	28 355
LE-Livro da Ensinança de Bem Cavalgar Toda Sela (1437/1438?)	38 784
LC-Leal Conselheiro (1433/1438?)	105 762
CD-Chancelarias Portuguesas: D. Duarte (1433-1438)	326 046
<b>TOTAL</b>	<b>871 123</b>

Textos portugueses do Século XVI	Nº. Palavras
HGP-Textos notariais (1502-1516) in <i>História do Galego-Português</i>	3 086
DN-Textos notariais (1504-1548) in <i>Docs. Notariais dos Séculos XII a XVI</i>	29 151
CRB-Crónica dos Reis de Bisnaga (sd)	43 900
<b>TOTAL</b>	<b>76 137</b>

### • O Dicionário do Português Medieval

Os dados do CIPM, extraídos com o auxílio de ferramentas automáticas, são divididos em três domínios – verbos, nomes e terminologias – são analisados e inseridos nas bases de dados respectivas: DVPM, DNPM e Jurix.

*O Dicionário de Verbos do Português Medieval – DVPM* – está a ser desenvolvido de acordo com a metodologia e os critérios estabelecidos para a 1ª. edição publicada por Xavier, Vicente e Crispim, *orgs.* 1999 – *Dicionário de Verbos Portugueses do Século 13*, que inclui 578 fichas lexicográficas. Uma edição revista e alargada incluirá verbos de textos do século XII e mais textos do século XIII. Decidiu-se também incluir os verbos das *Cantigas de Escarneo e Maldizer*, porque nelas se encontram muitos verbos de uso coloquial que não estão atestados nos documentos técnicos, régios ou nas vidas de santos.

Alguns dos critérios da edição anterior serão necessariamente alterados e outros novos serão introduzidos. Nesta edição serão incluídas também as **etimologias** dos verbos, serão apresentados **mais verbos, substituídas vedetas, estabelecidas novas correspondências** e acrescentadas **outras formas morfológicas e gráficas, novas acepções, proposições, subcategorizações e atestações**. Veja-se, por exemplo, a revisão de uma entrada que na 1ª. edição aparecia com um particípio passado porque no corpus estudado só havia uma atestação da forma correspondente, enquanto na nova edição já se encontram registadas mais quatro formas verbais que permitem introduzir a vedeta no infinitivo:

### 1ª edição

#### AMOESTADO

ocorrências : 1

PPmp amoestados 1 (HGP)

#### 1 repreendido, advertido

1285 HGP056 GOr – *sse alguma neglee~za y ouuer deuem(os) a séer amoestados ata duas uegadas q(ue) o corregam(os)*

### Nova edição

#### AMOESTAR

De *\*admonestare* de *admonere*

ocorrências: 5

IP1 amoesto 1 (VS)

IP3 amoesta 2 (VS)

PP3 amoestou 1 (VS)

PPmp amoestados 1 (HGP)

#### 1 admoestar, advertir

alguém amoesta alguém

[ — SN ]

sd VS3 – *E por em te rogo e amoesto que de todo e~ todo nõ desprezes nem abaixes minha hu(m)ildade*

*O Dicionário de Nomes do Português Medieval – DNPM* – contempla subclases semânticas para os **Nomes próprios (NP)** e para os **Nomes comuns (NC)**, simples e complexos:

- NP – *Deus, Santa Trindade*  
 NPA – antropónimo: *dona Marina Perez, São João*  
 NPT – topónimo: *Monte Mayor o Velho*  
 [...]
   
 NC – *água, amor, terra...*  
 Nca – animais: *porco galinha...*  
 NCm – metrológicos: *moio, spadoa...*  
 [...]

*O Dicionário de Terminologia Portuguesa Medieval – Jurix*. O estudo da terminologia presente em textos técnicos do CIPM constituirá este Dicionário. Em Xavier, Castro e Gonçalves (2000) apresentou-se o estudo cronológico e a história de alguns termos bem como a constituição da base de dados concebida para a terminologia antiga – *Jurix*. Esta apresenta os seguintes campos: vedeta, etimologia, definição e fonte da definição, variantes gráficas e respectivas fontes textuais, definição de temas associados aos termos e estabelecimento de termos relacionados – sinónimos/antónimos, termos de subdomínios. A inserção de dados nestes campos torna possível realizar múltiplas consultas cruzadas: obter um conjunto de termos pertencentes a um campo, seleccionar um termo e procurar os textos onde ocorre, etc.

### • Os Estudos Linguísticos Diacrónicos

Tendo por base os textos do CIPM, estão em curso diversos estudos linguísticos, visando particularmente o estudo da estrutura da frase e da sintaxe do verbo numa perspectiva diacrónica e comparada, através da análise sistemática das estruturas em que ocorrem os verbos das diferentes classes. As ordens de constituintes e as estruturas das orações infinitivas complementos de verbos de controlo, de elevação, modais, causativos e perceptivos estão a ser estudadas por M. Cristina V. da Silva; as orações transitivas, inergativas e inacusativas são objecto de análise por M. Alexandra Fiéis. Numa fase mais avançada das suas investigações, Maria Lobo observará nos textos antigos as adjuntas não finitas com verbos flexionados e não flexionados; Ana Castro deverá também fazer um estudo diacrónico das possessivas com os verbos “ter”, “haver”, “possuir”, “ser” e “estar”; Fátima Martins iniciou o estudo histórico dos advérbios, simultaneamente à análise dos verbos que vão sendo estudados diariamente com João Loureiro e Sandro Dias.

A ordem VinfinitivoVfinito em textos latinos e portugueses medievais tem estado a ser observada sistematicamente e quantificada tendo em conta dois grupos de textos – latino-romances e portugueses medievais.

A diferença do número de frases com as sequências V<sub>f</sub>in<sub>f</sub> V<sub>i</sub>n<sub>f</sub>in<sub>f</sub> e V<sub>i</sub>n<sub>f</sub>in<sub>f</sub> V<sub>f</sub>in<sub>f</sub> encontrada no *corpus* de textos latinos e no outro de textos portugueses medievais, analisados por M. Cristina Silva, é considerável. No conjunto de frases latinas dos séculos XI e XII, em que ocorrem um verbo finito e outro infinitivo – V<sub>f</sub>IV ou VV<sub>f</sub>I –, verifica-se que aquelas ordens de constituintes são quantitativamente semelhantes:

(1)	Século XI	Século XII
V <sub>f</sub> IV	271	407
VV <sub>f</sub> I	<u>202</u>	<u>471</u>
Total	473	878

Pelo contrário, o *corpus* de frases equivalentes extraídas de textos portugueses dos séculos XIII e XIV mostra uma diferença muito significativa no número de realizações de cada uma daquelas sequências verbais, em ambos os séculos:

(2)	Século XIII	Século XIV
V <sub>f</sub> IV	1570	1234
VV <sub>f</sub> I	<u>45</u>	<u>77</u>
Total	1615	1311

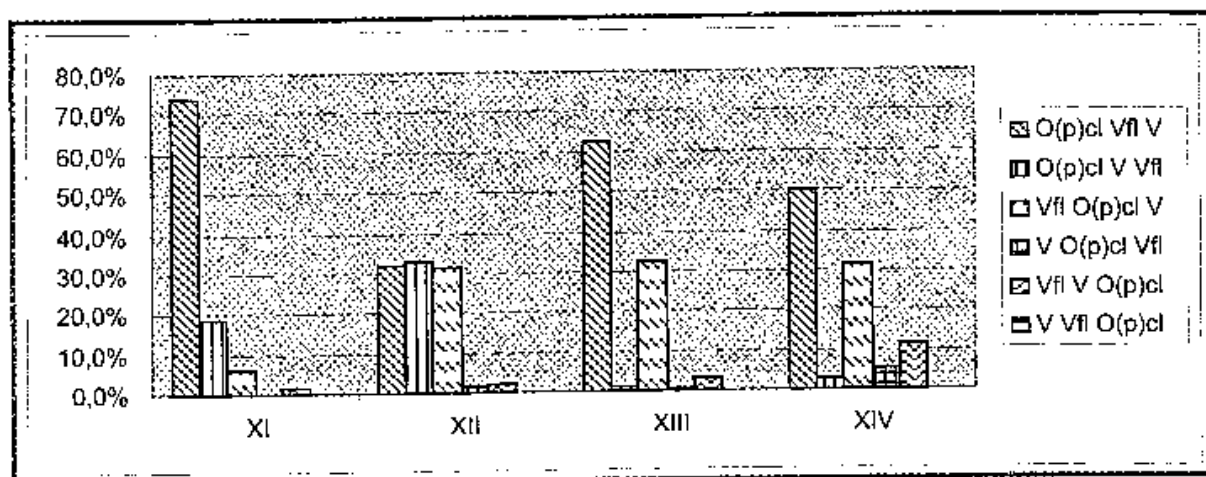
Esta importante diferença revela uma lacuna de fontes que impede a percepção clara do percurso de mudança linguístico-discursiva que o enfraquecimento de alternância daquelas duas ordens indicia. O afastamento da língua latina escrita da língua romance falada na época é revelado por esta notória diferença entre os valores das ordens de constituintes encontrados nos textos escritos em latim medieval e nos textos escritos em português antigo, estes certamente com uma sintaxe mais próxima da língua oral.

As sequências de dois verbos combinadas com as ordens VO e OV possibilitam seis alternativas:

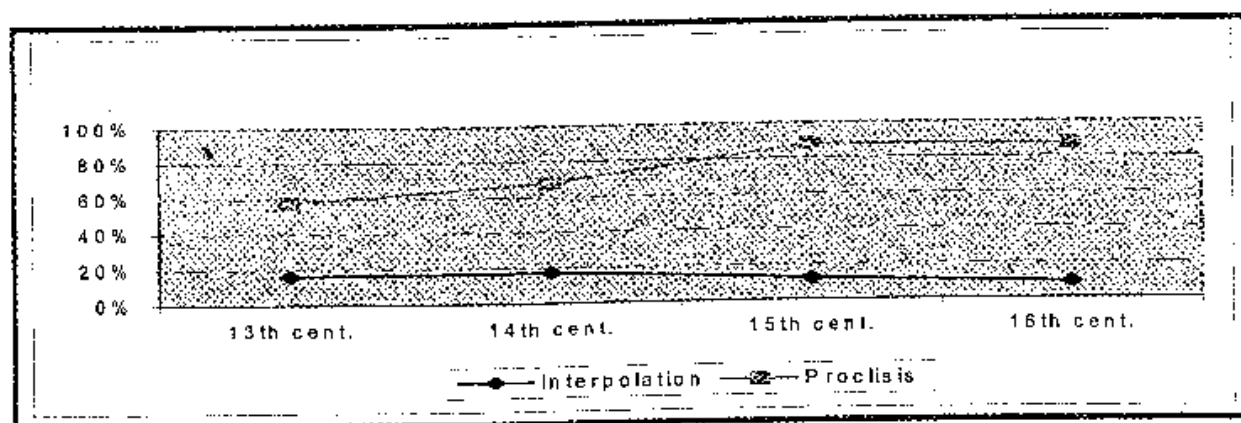
(i) OV <sub>f</sub> IV	(iv) OVV <sub>f</sub> I
(ii) V <sub>f</sub> IOV	(v) VOV <sub>f</sub> I
(iii) V <sub>f</sub> IVO	(vi) VV <sub>f</sub> IO

As seis ordens de constituintes são quantitativa e qualitativamente distintas em cada século, tanto nos textos em Latim Medieval como em Português Antigo.

Os objectos verbais realizados como sintagmas plenos ou como pronomes têm distribuição e frequência diferentes. Por ex., não se encontram pronomes – O(p)cI – na sequência (vi); a ordem (i) em que o objecto verbal é um pronome – O(p)cIV<sub>f</sub>IV – é a mais frequente em três dos quatro séculos, como mostra o gráfico seguinte:



A interpolação de constituintes na sequência *cl ...V* em textos em Latim Medieval dos séculos XI e XII e em textos do Português Antigo foi objecto de estudo recente. O estudo da interpolação desenvolvido por M. Cristina Silva (2001) a partir de textos notariais latinos dos séculos XI e XII e de textos portugueses dos séculos XIII e XIV, bem como o de M. Alexandra Fiéis (2000) tendo por base dois grupos de textos, tipologicamente diferentes, dos séculos XIII a XVI, mostram o decréscimo diacrónico da interpolação e a redução do número e da variedade de categorias interpoláveis. Os valores percentuais, sintetizados no gráfico abaixo, de Fiéis (2000), revelam que, nos séculos XIII a XVI, se verifica um decréscimo significativo do já reduzido número de frases com interpolação a par do aumento progressivo de frases com próclise.



No estudo, em curso, sobre a relação entre a ordem *VinfinitivoVfinito* e interpolação assume-se que a ordem básica de constituintes é *SVfl(S)VO* e que as diferentes ordens resultam de movimentos sintácticos e discursivos, por substituição ou adjunção, em derivações com o esqueleto *C T+fin vV T-fin vP* (Chomsky 1993-2001).

O confronto de diferentes teorias e propostas avançadas na explicação de fenómenos lexicais e morfossintáticos de outras línguas, dialectos e estados de língua, como os sistemas de auxiliar, as diversas realizações de caso, inversões e adjunções, posicionamento dos pronomes clíticos, entre outros, asseguram que os trabalhos em curso podem contribuir para uma melhor compreensão dos processos implicados na mudança e na variação linguísticas.

### Referências bibliográficas

- Chomsky, Noam (1995) *The Minimalist Program*. Cambridge MA, MIT Press.
- (2001) “Beyond Explanatory Adequacy”. Cambridge MA, MIT ms.
- Fiéis, Maria Alexandra (1997) *Clíticos num Corpus do Século XIII*. Dissertação de Mestrado apresentada à Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa.
- (2001) “Interpolação em Português Medieval como Adjunção a XP”. In *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, Colibri, pp. 197-211.
- Silva, Maria Cristina (1998) “Entre o Latim Medieval e o Português Medieval: a Complementação Infinitiva”. In *Actas do XIV Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, Associação Portuguesa de Linguística, pp. 431-441.
- (2001) “Ordens de Constituintes no Latim Medieval e Português Antigo: Scrambling e Interpolação”. In *Actas do XVII Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, Associação Portuguesa de Linguística (no prelo).
- Xavier, Maria Francisca e Maria de Lourdes Crispim (1999) “*Corpus* Informatizado do Português Medieval – CIPM. Constituição e Processamento Automáticos”. In *Actas do 6.º Congresso da Associação Internacional de Lusitanistas*, Rio de Janeiro (no prelo).
- Xavier, Maria Francisca, Graça Vicente e Maria de Lourdes Crispim orgs. (1999) *Dicionário de Verbos Portugueses do Século 13*, Lisboa, Linha de Investigação 1 do Centro de Linguística da Universidade Nova de Lisboa.
- Xavier, Maria Francisca, Ana Castro e Ana Gonçalves (2000) “A mais Antiga Terminologia Notarial Portuguesa”. In *Actas do Congresso Internacional 500 Anos da Língua Portuguesa no Brasil*, Évora (no prelo).