

A caminho de um Dicionário do Português Medieval

Maria Francisca XAVIER
(Universidade Nova de Lisboa)

Foi em 1992-93 que um grupo de linguistas da Universidade Nova de Lisboa definiu como objectivo desenvolver estudos linguísticos diacrónicos tendo por base textos dos primeiros tempos da História da Língua Portuguesa. Procurava-se desse modo dar início à realização tanto de um *Dicionário* como de uma *Gramática do Português Medieval*¹.

A preparação para o desenvolvimento do projecto de elaboração de um *Dicionário* e de outros estudos linguísticos envolveu, em primeira instância, a criação de um *Corpus* Informatizado de Textos Portugueses Medievais – o *CIPM*. Este inclui actualmente textos latinos datados entre o século IX e o século XII e textos portugueses desde o século XII até ao século XVI².

Pretendemos, fundamentalmente, dar um contributo para a explicitação do modo como o léxico e a gramática variavam nos textos daquele período e de que modo se foram verificando mudanças, ao longo daqueles séculos, no território que, actualmente, corresponde a Portugal³.

Os estudos linguísticos sobre o Português Medieval que estão a ser desenvolvidos no âmbito deste projecto incidem principalmente sobre a estrutura da frase e, em particular, sobre a sintaxe do verbo, numa perspectiva comparada e diacrónica. Os tópicos de investigação que têm estado a ser estudados pelos membros da equipa, em particular, por estudantes de pós-graduação, são os seguintes:

- (i) desaparecimento gradual das construções com o verbo finito em final de frase a par da generalização das construções com o verbo finito em posição medial;
- (ii) desaparecimento da ordem OV a favor da ordem VO;
- (iii) desenvolvimento da sintaxe dos clíticos;
- (iv) construções com alteração da ordem dos constituintes e com interpolação de constituintes na sequência "clítico... verbo".

Estes estudos envolvem a análise sistemática das orações infinitivas complemento de verbos modais, de controlo, de elevação, de percepção e causativos, e ainda de orações finitas transitivas, intransitivas e inacusativas.

Pretende-se verificar se existem diferenças estruturais diacrónicas que estejam relacionadas, por um lado, com propriedades morfossintáticas do(s) sistema(s) gramaticais e, por outro lado, com propriedades lexicais das classes verbais em questão.⁴

A parte do projecto relativa à construção do *Dicionário* começou pelo estudo dos verbos dos textos portugueses do século XIII que inicialmente integravam o *CIPM*. Foram então estabelecidos os critérios para a elaboração das respectivas entradas lexicais. Um primeiro volume intitulado *Dicionário de Verbos Portugueses do Século 13* foi publicado em 1999 por M.F. Xavier; G. Vicente e M.L. Crispim *orgs.* O projecto prevê, no entanto, que seja feito o estudo dos verbos encontrados em *corpora* textuais desde o período pré-literário até ao final do século XV, seguindo essencialmente a metodologia desenvolvida para o trabalho anteriormente realizado.

No primeiro volume publicado do *Dicionário de Verbos do Português Medieval - DVPM*, foram apresentadas 578 fichas de verbos, contendo estas tanto informação sobre a selecção semântica como sobre a selecção sintáctica de cada verbo. Nas entradas lexicais foram incluídas igualmente todas as variantes gráficas das formas verbais que se encontram atestadas no *corpus*

de textos do século XIII constituído para este estudo. A descrição das propriedades semânticas e sintáticas de cada verbo está ilustrada pela transcrição dos exemplos mais antigos que ocorrem no *corpus*. Sempre que os exemplos mais antigos não são completos relativamente à informação semântica e sintáctica descrita são incluídos, ainda, outros exemplos também do século XIII.

O *corpus* textual do século XIII de onde foram extraídos os verbos publicados em 1999 foi constituído pelos mais antigos textos não literários portugueses que até então se conheciam, nomeadamente: *Notícia de Torto*; *Testamento de D. Afonso II*; *Chancelaria de D. Afonso III*; *Tempos dos Preitos*; todos os textos do século XIII publicados na *História do Galego-Português* e no Appendix de *Clíticos na História do Português*. Mais recentemente outros textos dos séculos XII e XIII/XIV foram já objecto de estudo⁵.

A primeira fase de construção do *DVPM* permitiu desenvolver e testar uma metodologia que provou ser capaz de satisfazer o objectivo de extrair informação dos textos antigos com a finalidade de construir as entradas do *Dicionário*.

Dando continuidade ao projecto, foi publicado em 2002 um outro volume do *DVPM* que inclui 790 verbos, dos quais 212 correspondem a novas entradas de verbos do segundo conjunto de textos dos séculos XII e XIII/XIV, o que levou, simultaneamente, a rever e a aumentar a informação das entradas incluídas no volume anteriormente publicado⁶.

O próximo volume deverá conter as entradas de verbos encontrados em textos portugueses medievais do século XIII ao século XV que não estavam atestados nos textos anteriormente estudados. Paralelamente, estão a ser elaboradas as entradas lexicais de verbos de textos latinos do século XII.

As possibilidades oferecidas pelas novas tecnologias a este tipo de projectos são particularmente significativas em dois momentos sequenciais. Em primeiro lugar, a extracção da informação dos textos já informatizados pode ser facilitada com o auxílio de ferramentas várias, não devendo ser esquecidos os próprios *corpora* informatizados, o 'software' para preparação de concordâncias e a simples operação de busca do "Microsoft Word" para o "Windows". Podem também ser utilizados com bastante sucesso outros sistemas de processamento mais sofisticados, nomeadamente um etiquetador que associa as palavras à respectiva classificação gramatical e um analisador sintáctico que segmenta os constituintes da frase e também os classifica⁷. Na fase seguinte, a informação é introduzida em bases de dados do "Access", tornando-se possível realizar buscas múltiplas e aceder muito rapidamente aos dados dos diferentes campos⁸. Os verbos publicados nos dois volumes podem ser consultados 'on-line'⁹.

A metodologia seguida no trabalho já realizado e que continua a ser utilizada no desenvolvimento do *Dicionário* inclui os passos seguintes:

(i) selecção de textos do *CIPM* a serem estudados; (ii) identificação de verbos por meio do agrupamento e quantificação das formas verbais e respectivas variantes atestadas nos textos; (iii) codificação das formas verbais com informação sobre as fontes textuais; (iv) realização automática das concordâncias; (v) estudo dos contextos em que os verbos ocorrem; (vi) identificação da acepção em Português Medieval e do(s) equivalente(s) em Português Contemporâneo; (vii) descrição da selecção semântica; (viii) descrição da selecção sintáctica; (ix) exemplificação; (x) estudo da etimologia; (xi) transcrição fonémica; (xi) indicação de verbos relacionados; (xii) decisão sobre a vedeta da entrada lexical

As entradas do *Dicionário* deverão incluir:

(a) vedeta (seguida das vedetas dos verbos relacionados); (b) transcrição fonémica; (c) etimologia; (d) formas verbais (e variantes) morfológicamente classificadas, quantificadas e

codificadas com informação sobre as fontes textuais; (e) acepção (sinónimos ou perífrases) em Português Contemporâneo; (f) estrutura argumental (ou selecção semântica); (g) estrutura de subcategorização (ou selecção sintáctica); (h) exemplo(s).

A vedeta da entrada do *Dicionário* é apresentada, como habitualmente, no Infinitivo. No entanto, algumas vedetas correspondem a formas do Particípio Passado funcionando como adjectivo, quando estas são as únicas formas atestadas no *corpus*. Nestes casos são associadas, apenas, à respectiva acepção. Por exemplo: VELADA tem unicamente a informação relativa à sua acepção de "casada".

São também registadas no *Dicionário* algumas locuções verbais encontradas no *corpus*. Estas são apresentadas no final da informação da ficha do verbo a que estão associadas e são numeradas a partir de 101. Por exemplo: **POER** acepção 1 – pôr, colocar; 2 – estabelecer, fixar, acordar; 3 – apresentar; 101 – pôr-se (o sol), desaparecer na linha do horizonte.

Com o intuito de facilitar a consulta feita por utilizadores que desconhecem a variação gráfica dos textos antigos, a forma gráfica da vedeta das entradas do *DVPM* é, em geral, muito próxima da forma moderna. No entanto, por vezes, para se obter a forma moderna são aplicadas algumas regras de substituição sistemática de grafemas (cf. Xavier; Vicente; Crispim *orgs.* (1999), p. iv):

- <h> com função de marcação de hiato é eliminado
- grafemas duplos, iniciais ou internos, com o mesmo valor fonético dos correspondentes simples → grafemas simples
- <j> ou <y> com valor de vogal → <i>
- <i>, <gi> ou <y> com valor de sibilante sonora palatal → <j>
- <u> com valor consonântico → <v>
- <ll> ou <l> com valor de líquida palatal → <lh>
- <n> ou <gn> com valor de nasal palatal → <nh>
- <n> ou <~> representando a nasalação da vogal antes de oclusiva labial → <m>
- <~> representando a nasalação da vogal antes de consoante não labial → <n>
- <ç> antes de <e> ou <i> → <c>
- <s> intervocálico com valor de surda → <ss>
- <r> intervocálico com valor de múltipla → <rr>

A escolha da vedeta dos verbos é feita tendo em consideração os três critérios hierarquizados, que se seguem: (i) Uma forma verbal de Infinitivo atestada no *corpus* idêntica à sua equivalente em Português Contemporâneo, sempre que esta exista actualmente. Por exemplo: a vedeta AMPARAR (proteger) foi escolhida apesar de não corresponder à forma gráfica mais frequente que ocorre no *corpus*. As formas gráficas mais frequentes deste verbo começam com um "e" seguido da alternância da consoante nasal, que pode ser tanto "m" como "n" – "emparar", "enparar". A forma ALÇAR (levantar) foi escolhida para ser a vedeta de entre as três formas gráficas que coocorrem no *corpus*: "alçar", "álçar" e "allzár". ALÇAR é, simultaneamente, a forma gráfica moderna do Infinitivo daquele verbo e também a mais frequente no *corpus*. (ii) Uma forma verbal de Infinitivo reconstruída. Por exemplo: ESTABELEECER é uma vedeta que não corresponde a uma forma gráfica atestada. A forma atestada no Futuro do Conjuntivo, que é frequentemente idêntica à do Infinitivo, é "estabelescer". Contudo, outras formas que se encontram atestadas no *corpus*, como: "estabeleçe" e "estabelecemos", serviram de base para a decisão de apresentar aquela forma como vedeta. (iii) A forma com maior número de ocorrências no *corpus*.

Este critério raramente é aplicado e só se justifica quando, por um lado, não há equivalente à forma gráfica contemporânea e, por outro lado, a frequência elevada de uma forma o impõe. De facto, o verbo TER poderia corresponder a esta última situação¹⁰, uma vez que a forma gráfica de Infinitivo "teer" se encontra atestada no *corpus* textual que esteve na base do levantamento dos verbos já disponíveis 'on-line' e a forma com apenas um "e" não ocorre naqueles textos. Em contrapartida, a vedeta SER foi seleccionada, apesar da forma "seer" ter uma frequência mais elevada, porque a forma de Infinitivo "ser" se encontra atestada no *corpus*.

A questão sobre quais deverão ser as formas das vedetas num *Dicionário* histórico que abrange vários séculos continua em aberto e sujeita a ser repensada em diferentes fases da sua elaboração.

Os verbos que são semanticamente e morfossintacticamente semelhantes são agrupados de modo a poderem ser facilmente relacionados entre si. Deste modo, as vedetas que são relacionáveis com uma outra são apresentadas entre parênteses curvos, à direita daquela que encabeça a entrada do *Dicionário*. Por exemplo: APLAZAR (cf. EMPRAZAR) e EMPRAZAR (cf. APLAZAR). Estes dois verbos estão, assim, naturalmente associados no *Dicionário*, embora distantes na ordenação alfabética.

A classificação morfológica das formas do paradigma verbal é feita de acordo com o sistema verbal do Português, como se segue:

IP = Indicativo Presente; IPI = Indicativo Pretérito Imperfeito; IPP = Indicativo Pretérito Perfeito; IP+ = Indicativo Pretérito-mais-que-Perfeito; IF = Indicativo Futuro; C = Condicional; I = Imperativo; G = Gerúndio; PR = Particípio Presente; CP = Conjuntivo Presente; CPI = Conjuntivo Pretérito Imperfeito; CF = Conjuntivo Futuro; INFL = Infinitivo Flexionado; INF = Infinitivo; PP = Particípio Passado.

Relativamente à análise semântica dos verbos, procura-se definir o seu conteúdo proposicional e identificar a sua aceção. No entanto, e como seria de esperar, verifica-se que um verbo pode ter mais de uma aceção, dependendo do contexto em que ocorre. Por exemplo, o verbo ABRIR selecciona dois argumentos e é apresentado no *Dicionário* associado a duas aceções, nomeadamente 1. abrir e 2. apresentar, porque o verbo aparece em dois contextos situacionais bem distintos, tornando a sua interpretação bastante diferente em cada uma das situações, como se ilustra através dos exemplos seguintes:

1 abrir

alguém abre algo

[— SN]

1272 CA13 Pes - e se ly quiserem **abrir** a porta da adegã

2 apresentar

alguém abre alg

[— SN]

1280? FR PBA - Depoys que as prouas das testimonhyas forem **ab(er)tas** ant' o alcajde enaq(ue)l preyto

Observa-se também que alguns verbos seleccionam mais de uma proposição ou estrutura argumental. O verbo ACABAR, por exemplo, selecciona duas proposições, uma com um argumento apenas e a outra com dois argumentos. Contudo, na entrada lexical deste verbo encontra-se uma única aceção: 1 acabar, como se transcreve a seguir:

1 acabar

alg acaba algo

[— SN]

1280? FR PBA - *E se o p(re)yto for **acabado** p(er) este acusador*

acaba algo

[— SN]

1280? FG3 PAo - *Ata q(ue) **Acabe** o tempo*

Como se pode ver nos exemplos anteriores, a descrição semântica dos verbos apresentada no *Dicionário* contempla o número e o tipo de argumentos seleccionados por cada verbo, explicitando-os numa sequência linear, que inclui o próprio verbo. Esta descrição é feita tendo por base a classificação tripartida dos verbos geralmente adoptada no quadro da Gramática Generativa. De acordo com esta classificação, os verbos principais ou temáticos são divididos em transitivos, intransitivos e inacusativos. Os verbos transitivos têm dois argumentos obrigatórios e podem ainda ter mais argumentos, mas o aspecto que melhor caracteriza os verbos desta classe é a sua capacidade de seleccionarem, simultaneamente, um argumento agente e um argumento tema, os quais são instanciados, em Português, na sequência seguinte: x-agente verbo y-tema (a/de z-alvo/fonte), que corresponde à ordem básica do Português – Sujeito Verbo Objecto directo (Objecto indirecto). Saliente-se que dentro dos parênteses se encontram argumentos que não são definidores da classe dos verbos transitivos. Por essa razão, a estrutura de subcategorização (ou selecção sintáctica) dos verbos desta classe mostra, necessariamente, um sintagma nominal objecto directo como complemento do verbo e omite os outros argumentos-complementos inerentes que possam existir, como se pode ver na representação seguinte: [— SN]. Os verbos intransitivos caracterizam-se por seleccionarem um argumento obrigatório, que é tipicamente interpretado como agente, e não seleccionarem, obrigatoriamente, um objecto directo, embora possam, em muitos casos, realizar um, para além de poderem também seleccionar outro tipo de complementos, como se apresenta na sequência seguinte: x-agente verbo (y-alvo). E, uma vez que os verbos desta classe não seleccionam, necessariamente, complementos, a estrutura de subcategorização característica dos verbos intransitivos é representada apenas pelo travessão que corresponde ao próprio verbo [—]. Os verbos inacusativos seleccionam obrigatoriamente um argumento que é tipicamente um tema: verbo x-tema (y-lugar). Embora possam também seleccionar outros complementos, a sua estrutura de subcategorização mínima é idêntica à dos verbos transitivos: [— SN]. A diferença fundamental existente entre os verbos transitivos e os inacusativos reside no facto de seleccionarem estruturas argumentais diferentes, o que dá origem a que o argumento tema dos verbos de cada classe seja legitimado em posições estruturais também diferentes. Assim, os verbos transitivos legitimam o argumento tema como complemento directo, enquanto o tema dos inacusativos é legitimado como sujeito.

Na descrição do conteúdo proposicional dos verbos que é feita para o *DVPM*, o argumento agente é apresentado à esquerda do verbo, representado pela palavra "alguém"; o argumento tema é apresentado à direita do verbo, representado por "algo". Por exemplo: "alguém abre algo". São também explicitados os argumentos adicionais, sempre que há evidência da sua existência, e quando o contexto textual é insuficiente recorremos ao nosso conhecimento lexical para descrever as estruturas argumentais completas de alguns verbos.

Na descrição dos argumentos entra-se ainda em consideração com alguma informação sobre restrições de selecção que os verbos podem impor aos seus argumentos. Deste

modo, a palavra "alguém" refere argumentos animados, "algo" representa argumentos não animados, e "alg" é utilizado para argumentos que não parecem ser semanticamente restringidos. São também adoptados alguns termos genéricos para descrever argumentos de outras proposições seleccionadas por alguns verbos, nomeadamente "função", "situação", "lugar", "fazer", "acontecer", etc.¹¹.

Podem também estar associadas a um verbo tanto diversas proposições como várias estruturas de subcategorização. As estruturas de subcategorização dos verbos do DVPM correspondem a 15 esquemas: - O esquema A: [—], exemplificado por CHORAR, descreve a selecção sintáctica de um verbo intransitivo que não tem qualquer argumento-complemento. Aquele verbo selecciona apenas um argumento agente – alguém chora. Os esquemas de B a F representam as selecções sintácticas associadas a verbos que têm argumentos-complementos simples – SN ou SP. - O esquema B: [— SN] apresenta apenas um argumento-complemento SN, estando acima exemplificado por ACABAR, um verbo que pode ter um funcionamento tanto transitivo como inacusativo – alguém acaba algo ou acaba algo –, respectivamente. - O esquema C: [— SN SP] inclui dois argumentos-complementos, que podem ser seleccionados, por exemplo, por ACUSAR e por AFORAR, embora estes verbos apresentem diferenças significativas entre eles. Cada um destes verbos impõe restrições de selecção aos seus argumentos e realiza os complementos de maneiras diferentes, como mostram as sequências seguintes: alguém acusa (a) alg em/de/sobre/per/por algo ou alguém afora algo a/de alguém. - O esquema D: [— SN SP SP] corresponde à estrutura de complementação mais longa possível. Este esquema é exemplificado com o par de verbos COMPRAR e VENDER – alguém compra/vende algo de/a alguém por algo. - O esquema E: [— SP], exemplificado por ACORRER – alguém ocorre a alguém –, mostra que o verbo selecciona apenas um argumento-complemento SP. - O esquema F: [— SP SP] representa a subcategorização de verbos intransitivos da classe de FALAR, que seleccionam obrigatoriamente um argumento agente e, opcionalmente, dois argumentos-complementos preposicionados – alguém fala em/sobre algo a/com alguém. Os esquemas de G a O representam as estruturas de subcategorização de verbos que seleccionam complementos oracionais, tanto orações pequenas, ou seja, relações predicativas sem tempo independente, como orações plenas. Este conjunto de estruturas de subcategorização engloba um conjunto de verbos sintacticamente interessante que merece ser descrito mais detalhadamente de modo a caracterizar as diferenças existentes entre eles. Os esquemas de G a I podem ser ilustrados pelo verbo ACHAR. Este verbo tem, essencialmente, duas acepções: "considerar" e "encontrar". Correspondem a este verbo, no entanto, várias proposições, uma vez que pode seleccionar conjuntos de argumentos diversos. De facto, ACHAR encontra-se no *corpus* a seleccionar um argumento com a interpretação de experienciador (lugar psicológico) e quatro estruturas de complementação diferentes: - O esquema B: [— SN] corresponde à acepção de ACHAR como "encontrar" – alguém acha alg, estando nos textos frequentemente explicitado o lugar onde o objecto é achado. - Os esquemas G: [— [SN SX]] e H: [— que Find] estão associados à acepção de "considerar". O verbo com esta acepção selecciona quer uma oração pequena, cuja estrutura simplificada se encontra descrita em G, à qual corresponde a proposição – alguém acha alg em situação, quer uma oração plena no modo indicativo, introduzida pela conjunção declarativa "que", representada em H: [— que Find], que tem a correspondente proposição – alguém acha que... - O esquema I: [— a SN Vinf] é também associado a ACHAR significando "encontrar" – alguém acha a alguém fazer... Aqui, o complemento do verbo é uma oração infinitiva cujo sujeito é regido pela preposição "a". - O esquema J: [— (a) Vinf] pode ser ilustrado pelo verbo ATENDER. Este selecciona tanto um argumento experienciador, que funciona como sujeito, como um

complemento infinitivo, antecedido, por vezes, pela preposição *a* – alguém atende (*a*) acontecer. - O esquema K: [— *a/de* Vinf] pode ser exemplificado pelo verbo CONVIR, significando “concordar”, o qual selecciona um argumento externo e um complemento infinitivo precedido de uma das duas preposições *de/a* – alguém conuem *de/a* fazer. - O esquema L: [— (SP) *que/se* Find/conj] descreve a selecção sintáctica de uma oração finita no modo indicativo ou no conjuntivo, e esta pode ser introduzida por qualquer um dos dois complementadores – “*que*” ou “*se*”. O verbo DIZER pode servir de exemplo para este esquema de subcategorização, associado à proposição – alguém diz a alguém *que/se*... - O esquema M: [— Q Find] representa a selecção de uma oração complemento introduzida por um quantificador ou por um operador interrogativo. O verbo DESCOBRIR, por exemplo, selecciona, para além de um argumento individual que funciona como sujeito, um argumento complemento que corresponde a uma proposição introduzida por um operador – alguém descobre Q... - O esquema N: [— Vger/inf] corresponde à complementação de um verbo que selecciona ou uma oração gerundiva ou uma infinitiva, e pode ser exemplificado pelo verbo IR – vai acontecendo/fazendo; vai acontecer/fazer. - O esquema O: [— Vppflex] representa o complemento participial do verbo SER, o qual flexiona conforme os traços nominais de pessoa e número da concordância com o sujeito – é acontecido/feito. Como alternativa, a simples introdução de flex entre parênteses curvos - [— Vpp(flex)] descreve a possibilidade de serem flexionados os participios complementos dos verbos HAVER e TER.

A descrição sintáctica e semântica dos verbos incluídos no *Dicionário* já publicado foi preparada, essencialmente, tendo por base as atestações encontradas nos textos do CIPM. Contudo, como já atrás foi referido, sempre que os dados são insuficientes ou levantam dúvidas, a análise dos verbos é feita na base de hipóteses sugeridas pelo nosso conhecimento lexical do Português Contemporâneo. Este é, certamente, um aspecto a reter e que poderá levar a futuras alterações à medida que novos dados recolhidos de outros textos, de outros períodos ou estilos, estejam disponíveis.

Consideremos, ainda, quatro exemplos de entradas lexicais de verbos do século XIII, que são morfológica, semântica e sintacticamente relacionadas entre si – APENHORAR, EMPENHORAR, PENHORAR e SOPENHORAR. Estes quatro verbos têm a mesma base - “penhorar”. Um deles não tem qualquer prefixo e os outros três exibem prefixos diferentes. Todos, com a excepção de SOPENHORAR, estão atestados em diversos textos de diferentes proveniências, este só foi encontrado em textos da Galiza. Verifica-se que as formas mais antigas atestadas datam de 1255 e têm todas um prefixo, correspondendo aos verbos APENHORAR, EMPENHORAR e SOPENHORAR. A primeira ocorrência de PENHORAR é de 1277. As quatro entradas lexicais têm essencialmente a mesma informação semântica e sintáctica, embora revelem algumas pequenas diferenças, que resultam directamente dos contextos textuais em que as formas verbais se encontram. As diferenças resumem-se, por um lado, ao facto de estar ou não realizado o argumento inerente alvo/fonte destes verbos, o que é equivalente ao que acontece com as estruturas argumentais de verbos como DAR/ROUBAR, em que há sempre “algo” para ser “dado/roubado a alguém”. Por outro lado, como acontece também com o verbo ROUBAR, é possível com aqueles verbos realizar o argumento alvo/fonte como objecto directo – alguém penhora alguém – ficando, assim, implícito no verbo o argumento tema - “algo” que foi penhorado ou roubado. Acontece, porém, que o argumento alvo/fonte se encontra realizado em algumas atestações daqueles verbos, enquanto noutras situações este argumento permanece implícito, à excepção do que se verifica com o verbo SOPENHORAR, que nas suas atestações tem sempre presente o argumento alvo/fonte.

Uma questão que se nos coloca é a de saber se aquelas quatro entradas lexicais serão redutíveis a uma única, uma vez que a análise linguística mostra que não existem diferenças semânticas e sintáticas importantes entre elas, e as formas com prefixos distintos podem, eventualmente, ser consideradas variantes sem valor linguístico relevante. A resposta a esta questão terá de ser adiada até à altura em que tenhamos já conseguido esclarecer outros aspectos. Interessa-nos verificar até que ponto poderão estas variantes ser estudadas. Teremos de observar, tanto em diacronia como em sincronia, outros grupos de verbos exibindo este tipo de relação entre eles. Gostaríamos de compreender o que motivará o aparecimento destas variantes. Como, onde e quando são estas variantes realizadas? Estas são algumas questões que temos em aberto para investigação a desenvolver sobre o assunto, no futuro próximo.

A informação do *DVPM*, necessariamente, muito sucinta poderá ser completada pela consulta das próprias fontes textuais e de outros documentos. Encontra-se em estudo a possibilidade de estabelecer ligações 'on-line' entre as formas verbais atestadas e os textos de onde estas foram retiradas, bem como com as concordâncias preparadas a partir das mesmas formas ou, ainda, com outros objectos escritos, nomeadamente glossários e outros estudos linguísticos e culturais realizados tendo por base os textos que integram o *CIPM*. Finalmente, os resultados já conseguidos através da investigação multidisciplinar em curso estão a ser gradualmente disponibilizados na "Internet", dando, assim, a possibilidade de serem partilhados por todos os interessados, o que constituirá, ainda, um estímulo não só para continuar a melhorar e a desenvolver o *CIPM* e o *DVPM* mas, igualmente, para prosseguir diversos estudos linguísticos e finalmente realizar um Dicionário e uma Gramática do Português Medieval.

Notas

¹ Este projecto foi financiado inicialmente pela JNICT e pelo PRAXIS XXI e mais recentemente pelo MCES/FCT e POCTI.

² Ver constituição do CIPM em <http://cipm.fcsh.unl.pt>.

³ Tanto a lista dos textos que constituem actualmente o CIPM bem como muitos dos textos podem ser consultados na Internet, em <http://cipm.fcsh.unl.pt>. A sua preparação tem envolvido para além de linguistas, filólogos e informáticos também historiadores, investigadores de literatura e estudantes.

⁴ A investigação acima referida está em grande parte incluída nos programas de doutoramento de M.Cristina V. Silva, M. Alexandra Fiéis, Maria Lobo e Ana Castro.

Nesta fase de desenvolvimento do projecto de Dicionário está planeado o seu alargamento, simultaneamente, quanto à cronologia e quanto à diversidade das fontes textuais a cobrir. Tal procedimento implicou, necessariamente, uma reflexão nova e aprofundada relativamente às decisões anteriormente tomadas sobre os critérios para a preparação das fichas dos verbos.

⁵ As referências completas dos textos estudados até 2002 encontram-se em <http://cipm.fcsh.unl.pt>.

⁶ Cf. M.F. Xavier, G. Vicente e M.L. Crispim orgs. (2002).

⁷ Cf. Rocio; Alves; Lopes; Xavier; Vicente (2003).

⁸ A base de dados do *Dicionário de Verbos do Português Medieval* foi desenvolvida pelo engenheiro informático Luís Reis.

⁹ Em <http://cipm.fcsh.unl.pt>

¹⁰ No volume do *DVPM* publicado em 1999 a vedeta aparece como TEER, porque aí foi aplicado este critério e no *corpus* ocorre um número elevado da forma gráfica de Infinitivo "teer", ao passo que a forma "ter" não está atestada. No entanto, no volume de 2002 a vedeta é TER e uma nota de rodapé indica a representação do que poderá ter sido a oralização correspondente: t/e~e/r, e a respectiva etimologia: do latim *tene~re*.

¹¹ Cf. a descrição destes casos em Xavier; Vicente (1997), assim como as entradas lexicais apresentadas em Xavier; Vicente; Crispim orgs. (1999); (2002) e em <http://cipm.unl.pt>.

Referências

- Rocio, V., M. A. Alves, J. G. Lopes, M. F. Xavier & G. Vicente (2003) Automated Creation of a Medieval Portuguese Treebank, in Abeillé, A. *ed. Treebanks - Building and Using Parsed Corpora*, Dordrecht, Kluwer Academic Publishers, Chapter 12, 211-227.
- Xavier, M.F.; M.G. Vicente (1997) "A Problemática de um Dicionário de Verbos do Século XIII" in *Sentido que a Vida Faz. Estudos para Óscar Lopes*, Porto, Campo das Letras, 897-904.
- Xavier, M.F.; M.L. Crispim; M.G. Vicente (1999) "A Língua Portuguesa Medieval e o Diálogo Homem-Máquina" in I.H. Faria *org.*, *Lindley Cintra. Homenagem ao Mestre e ao Cidadão*. Lisboa, Edições Cosmos e Faculdade de Letras da Universidade de Lisboa (749-760).
- Xavier, M.F.; M.G. Vicente; M.L. 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, M. F., M.G. Vicente, M. L. Crispim *orgs.* (2002) *Dicionário de Verbos do Português Medieval - Séculos 12 e 13/14*, Lisboa, Linha de Investigação 1 - Centro de Linguística da Universidade Nova de Lisboa.