

Ordens de constituintes no Latim Medieval e Português Antigo: *scrambling* e interpolação

Maria Cristina Vieira da Silva
Centro de Linguística da FCSH-UNL*

0. Introdução:

Com o presente trabalho, pretendeu-se analisar a estrutura frásica subjacente a cada um dos períodos documentados textualmente pelo nosso *corpus*¹, de forma a esclarecer quais as diferenças existentes, do ponto de vista sintáctico, entre o Português Antigo e o estágio de língua que convencionámos designar de Latim Medieval.

Defenderemos que a língua retratada quer pelos nossos textos do Latim Medieval (séculos XI e XII), quer pelos do Português Antigo (séculos XIII e XIV) partilham uma mesma ordem subjacente, isto é, a ordem SVO. Nos primeiros, as ordens atestadas sugerem que *scrambling* se aplica em maior escala e que a sua aplicação vai enfraquecendo na diacronia do Português.

* Este trabalho insere-se no âmbito de um projecto de Doutoramento subsidiado pela Fundação para a Ciência e a Tecnologia, ao abrigo do Projecto "Corpora do Português Medieval, Etiquetagem e Segmentação Automática", com a referência Praxis/2/2.1/CSH/778/95.

¹ As fontes trabalhadas foram organizadas da seguinte forma:

‡ Século XI: 428 documentos seleccionados de uma edição impressa da série *Diplomata et Chartae*, integrada na publicação, promovida pela Academia Portuguesa de História, dos *Portugaliae Monumenta Historica*.

‡ Século XII: 380 documentos seleccionados a partir de *Documentos Medievais Portugueses – Documentos Régios*, volume I (Documentos dos Condes Portucalenses e de D. Afonso Henriques (a.d. 1095-1185)), edição datada de 1962, organizada pelo historiador e paleógrafo Rui Pinto de Azevedo e promovida pela Academia Portuguesa de História.

‡ Século XIII: 175 documentos correspondendo a nove diferentes edições acedidas através do *Corpus Informatizado do Português Medieval – CIPM*.

‡ Século XIV: 146 documentos correspondendo a quatro edições distintas, acedidas através do *Corpus Informatizado do Português Medieval – CIPM*.

	Identificação da edição informatizada	Total de textos	datados entre e	Total de palavras	Vocabulário
Século XI	D&C	428	1002 – 1100	126.742	13.215
Século XII	DR	380	1101 – 1185	132.660	15.346
Século XIII	CIPM13	175	1214 – 1300	135.008	16.827
Século XIV	CIPM14	146	1301 – 1399	109.625	14.123

Qualificando-se como contexto diagnóstico para a existência de *scrambling*, o facto de haver mais interpolação nos textos médio-latinos do que nos do Português Antigo vem justamente confirmar a ideia de que a produtividade da interpolação terá constituído evidência positiva para a aquisição de *scrambling*, com a posição do (proto-)clítico a servir de diagnóstico.

1. Ordem de constituintes no Português Antigo e no Latim Medieval:

Segundo Mattos e Silva (1989), encontram-se atestadas, no Português Antigo, as seis disposições matematicamente possíveis em estruturas com verbos transitivos: SVC, SCV, VSC, VCS, CVS e CSV. De todas estas, a primeira (a “ordem directa” segundo a terminologia de Pádua (1960) ou a “ordem normal” para Huber (1933)) é considerada a ordem não-marcada dos diferentes constituintes em presença.

Corroborando as descrições já existentes na literatura, Martins (2000b) define o Português Antigo como uma língua SVO de sujeito nulo, tal como o Português Europeu Contemporâneo. Segundo Martins, a ordem OV co-existe a par com a ordem VO, em variação aparentemente livre até bastante tarde, desaparecendo dos registos escritos apenas após o século XVI.

Sendo o Português Antigo uma língua (S)VO, a ordem OV é derivada, segundo Martins (2000b), mediante três tipos de mecanismo: deslocação à esquerda, exemplificada em (1), focalização (veja-se (2)) ou *scrambling* do objecto (veja-se (3)):

- (1) a. Et aos pobres mādoulles dar algo (Ogando (1980: 254))
 b. O conselho já o eu filhei (Huber (1933: §452))
 c. E todo aq(ue)lo q(ue) hj for desspeso eyo eu po(r) firm(e) (1383, CHP142)
- (2) a. Estas pallavras e outras muitas lhe disse dom egas moniz (Pádua (1960:144))
 b. E esta cousa nos manda a Regla que façamos (Pádua (1960:149))
- (3) a. (e) nos q(ue) chu prazo Mādam(os) fazer; com nosas Man(os) proprias ho reuoram(os) por reuora recebem(os) de ti. (1272, CHP003)
 b. e se nolo vos en ese dia nō derdes (1296, CHP056)
 c. (e) dou A uos a d(i)ta mh~a herdade por escanbho (e) en nomé descambho pola d(i)ta uosa h(er)dade; pela guísa q(ue) me uos A uosa dades (1333, CHP120)
 d. sse as vender, q(ui)s(er)d(e)s deuedelo ffaz(er) sab(e)r A nos; ou Aos nos-sos supçesor(e)s (1343, CHP125)
 e. lhe desse asy este estorm(ento) dos sobred(i)ctos autos q(ue) sse p(re)ssente m̄j t(a)b(alia)m pasarã t(estemunhas) (1472, CHP185)
 f. E q(ue) elle posa q(uar)tejar as t(e)rras do d(i)cto cassal este p(re)ssente ãno com os lauradores q(ue) as ssemeadas teuerẽ ou ssemearẽ daquy ã deante (1472, CHP185)

- g. quem q(ue)r q(ue) lhe **sobre elle (e) parte delle (e) sobre as ditas casas (e) q(ua)lq(ue)r cousa delas algũu embargo ou empedym(ento)** puser (1540, CHP209)

Apesar do grande número de estudos relativos ao latim medieval, poucos são os trabalhos que, incidindo sobre questões sintácticas concretas, buscam clarificar a questão da ordem de constituintes. Ao seu reduzido número acresce ainda o facto de estes estudos incidirem sobre realidades linguísticas distintas da que nos ocupa, uma vez que a transição do Latim para as diferentes línguas românicas terá tido uma diferente cronologia nos diferentes espaços. Daqui decorre que sejamos conduzidos a adoptar como ponto de partida o Latim Clássico, a partir de descrições já existentes na literatura (vejam-se, nomeadamente, os estudos dos autores clássicos ou mais recentemente, Panhuis (1982) e Pinkster (1995)). Estas (entre outras descrições das estruturas latinas) vieram demonstrar que a ordem dos constituintes sintácticos em Latim não seria arbitrária, refutando assim uma hipótese largamente difundida, a da alegada liberdade na ordem de constituintes. Ao invés, os diferentes constituintes apresentam uma ordem não-marcada, a par de um número de variantes discursiva ou estilisticamente motivadas. Como Marouzeau (1922) assinala, se o Latim não manifesta uma ordem fixa de constituintes, também é verdade que duas ordens diferentes não são necessariamente sinónimas:

“L' ordre des mots en latin est libre, il n'est pas indifférent. Libre, en ce sens que, sauf exception, il n'y a pas pour chaque terme de la phrase une place attitrée, obligatoire. Mais non pas indifférent, parce qu'en general deux ordres possibles ne sont pas synonymes.” Marouzeau (1922: 1)

Sem fazer ainda uso dos termos “ordem marcada” e “ordem não-marcada”, Marouzeau manifesta já consciência da distinção entre uma ordem subjacente e ordens derivadas, resultantes de contextos sintácticos e estilísticos particulares.

Neste sentido, assumiremos, como hipótese de trabalho a proposta de Kayne (1994) segundo a qual as línguas são universalmente VO. Note-se que a existência de ordem VO não é incompatível com o postular de um movimento do Objecto para a esquerda, para além do verbo, independentemente deste último se poder mover igualmente. Daqui resulta uma ordem OV derivada a partir de um ordem VO subjacente, sendo que as diferenças observadas quanto à ordem superficial resultariam de diversas aplicações da regra de movimento. Esta hipótese, originalmente defendida por Kayne (1994), tem sido aplicada às línguas que evidenciam o fenómeno designado de *scrambling* por linguistas como Zwart (1997) (para o holandês) e Roberts (1997) (para estádios antigos do Inglês).

Se línguas como o holandês e o alemão forem consideradas línguas VO, tal implicará uma aplicação mais extensa do fenómeno de *scrambling*, de forma a derivar correctamente as ordens observadas em superfície. De igual forma, defendemos que a língua retratada quer pelos nossos textos do Latim Medieval (séculos XI

e XII), quer pelos do Português Antigo (séculos XIII e XIV) apresenta a ordem subjacente SVO, ainda que nos primeiros as ordens atestadas revelem que *scrambling* se aplica em larga escala, enquanto nos últimos a sua aplicação vai enfraquecendo diacronicamente.

2. Scrambling: do Latim Medieval ao Português Antigo

Ross (1967), propõe o termo *scrambling* justamente para dar conta da alegada ordem “livre” em línguas como o Latim, remetendo este fenómeno para a componente estilística e excluindo-o assim da componente transformacional. Este termo recobre a aceção de fronteamento ou elevação de constituintes vários, nomeadamente objectos directos, indirectos, ou ainda sintagmas preposicionais, dependendo de condições específicas de cada língua. No quadro de Princípios e Parâmetros e do Programa Minimalista, *scrambling* é reduzido à regra de Mover- α , sendo esta aplicada após “Spell-Out”, devido a propriedades discursivas das línguas em que se verifica este fenómeno.

Alguns linguistas (nomeadamente Roberts (1997) e Weerman (1997)) assinalaram já que, nomeadamente na história do Inglês e do Holandês, *scrambling* manifesta uma tendência para reduzir a sua frequência (desaparecendo eventualmente ou restringindo o seu uso). Dado que o sistema casual destas línguas sofre uma simplificação sensivelmente no mesmo período (com a perda das distinções Casuais), considera-se que estes dois fenómenos estão relacionados entre si.

A proposta de que *scrambling* seja um fenómeno relevante também em estádios históricos do Português (e não apenas nas línguas germânicas, onde *scrambling* se encontra largamente atestado) não é inovadora, tendo já sido defendida para outras línguas românicas como o italiano (cf. Belletti & Shlonsky (1995)), bem como para o Português Europeu Contemporâneo (cf. Costa (1998)) e para o Português Antigo (cf. Martins (2000a) e (2000b)).

2.1. Scrambling no Português Antigo

Martins (2000b) define o Português Antigo como uma língua (S)VO (como, aliás, o Português Europeu Contemporâneo), na qual a possibilidade atestada da ordem de constituintes OV é derivada, entre outros mecanismos, pelo fenómeno designado de *scrambling*. Sendo uma língua de sujeito nulo, o Português Antigo apresenta frequentemente frases sem sujeito realizado, a par de sujeitos realizados em posição pós-verbal (que podem eventualmente envolver *scrambling*), pelo que se conclui que a posição do sujeito por si só não deverá constituir um diagnóstico de estruturas de *scrambling*.

Já os contextos de interpolação se qualificam como testes diagnósticos fiáveis, uma vez que, de acordo com Martins (2000b), o clítico parece assinalar a fronteira entre o constituinte deslocado à esquerda ou constituintes focalizados, por um lado, e o constituinte movido por *scrambling*, por outro.

Como Martins (2000b) assinala, a ordem OV ocorre majoritariamente em frases encaixadas, sendo a ordem (S)OV menos frequente em orações principais (possivelmente porque nestas o verbo se move para uma posição funcional mais alta do que nas subordinadas, ainda que Martins (2000b) considere não haver evidência empírica para assumir a existência de V-2 no Português Antigo). A hierarquia dos diferentes constituintes em frases subordinadas seria então a apresentada em (4), e ilustrada em (3)) como *scrambling* de objecto directo ((3a.) e (3c.)), de um SP adjunto (3b.), de um complemento verbal infinitivo (3d.), de uma oração participial (3e.), de uma oração pequena (3f.) e de um SP complemento (3g.).

- (4) "COMP > Tópico > Foco > Clítico > XPs movidos por *scrambling*/ interpolados > V"
(Martins 2000b)

Assumindo, com base na evidência fornecida pela posição do sujeito (concretamente a possibilidade de sujeito nulo e sujeitos pós-verbais) bem como pela colocação de advérbios temporais (em particular, a ocorrência do verbo à esquerda de advérbios que ocupam posições mais baixas na estrutura) a existência de movimento do verbo do interior de VP para IP no Português Antigo, Martins (2000b) identifica este *scrambling* como adjunção a IP (ou *scrambling* médio) e não *scrambling* de VP (adjunção a VP ou *scrambling* curto, como proposto por Costa (1998) para o Português Europeu Contemporâneo). Esta análise assume assim que a posição relevante para a qual os constituintes se movem mediante *scrambling* seja a de um núcleo funcional de IP, uma vez que, no Português Antigo, estes objectos se movem para posições mais altas do que no Português Europeu Contemporâneo².

Para dar conta da perda, em termos diacrónicos, de *scrambling* de IP e da ordem (S)OV que lhe está associada, Martins (2000a) e (2000b) defende a existência de uma simplificação na arquitectura da frase, com a mudança de um IP desdobrado para um IP não desdobrado, a par da alteração das propriedades de AgrS. Isto é, este núcleo funcional deixou de permitir múltiplos Specs (as posições até aí disponíveis para os constituintes movidos por *scrambling*), razão pela qual *scrambling* de IP desaparece.

Como Martins (2000b.) assinala, do ponto de vista de uma teoria da mudança linguística que faça apelo à necessidade de explicar de que forma se processou a aquisição dessa mudança (cf. Lightfoot (1991); (1999)), é de prever que a evidência positiva que até aí desencadeara a aquisição do traço relevante deixe de estar disponível. A partir do século XVII, as mudanças na colocação dos pronomes clíticos originaram o desaparecimento gradual das estruturas de interpolação (com os clíticos a surgirem normalmente adjacentes ao verbo) e concomitantemente a perda de *scrambling* e da ordem (S)OV.

² Note-se que, em Martins (2000a) e (2000b), o termo Objecto é utilizado no sentido Larsoniano do termo, sub-entendendo-se que, numa concepção de tipo "VP-shell", os constituintes complementos e adjuntos se posicionam indistintamente no interior do VP.

2.2. *Scrambling* nos textos médio-latinos

A análise de *scrambling* aqui adoptada deverá ainda permitir-nos estabelecer um marco de comparação relativamente aos textos médio-latinos que nos propomos estudar. Neste sentido, parece-nos interessante assinalar que a produtividade deste mecanismo evolui diacronicamente no sentido de uma maior restrição, como assinalado por Thráinsson (2001: 186).

De facto, se assumirmos, com Martins (2000a) e (2000b) a existência de *scrambling* de IP na fase designada de Português Antigo, e havendo evidências para este fenómeno no Português Europeu Contemporâneo (ainda que com uma distribuição mais restrita), a questão que se coloca consiste naturalmente em saber se também a produção textual do período médio-latino que nos ocupa apresenta evidências desse mesmo fenómeno.

3. Dos pronomes proto-clíticos aos clíticos e a interpolação

Propomo-nos, nesta secção, estabelecer o percurso dos elementos considerados precursores dos pronomes clíticos, visando a observação do seu comportamento em contextos infinitivos nos dois *corpora* de textos (um constituído por textos médio-*latinos* e outro por textos medievais portugueses, respectivamente dos séculos XI/XII e XIII/XIV).

3.1. Dos proto-clíticos aos clíticos

Interessam-nos particularmente as formas do sistema pronominal latino (concretamente das formas de pronome pessoal e formas pronominais deícticas) nas quais reside a génese da categoria funcional núcleo de DP.³

Relativamente ao sistema de demonstrativos do Latim Clássico, este teria sido reorganizado, segundo Penny (1993: 145) (*apud* Dillet (1998)) da seguinte forma:

(5) Reorganização do sistema de pronomes demonstrativos do Latim:

Anafórico	Pessoal	Demonst.1	Demonst.2	Demonst.3	Ênfase	Identidade
IS	ILLE	HIC	ISTE	ILLE	IPSE	IDEM
ILLE	ILLE	ISTE	IPSE	*ACCU ILLE	*MEDIPSISSIMUS	
o	ele	este	esse	aquele	mesmo	

³ Recorde-se que esta categoria emerge na sintaxe românica com o aparecimento de dois novos elementos -os clíticos e os artigos- que partilham entre si uma evolução paralela.

De entre todas as formas em (5), a redução gradual da força deíctica (mais acentuada nos casos de IS, EA, ID e ILLUM, ILLAM, ILLUD) teria culminado na sua constituição como clíticos com formato segmental reduzido nas línguas românicas. O paradigma pronominal IS, EA, ID em particular, enquanto item puramente anafórico, tende a ser substituído pelo paradigma de ILLE, uma vez que resiste mais dificilmente às mudanças fonéticas que afectaram o Latim. ILLE herda assim as funções de IS, que eram já partilhadas entre os dois, dado que ambos permitem recuperar a referência relativamente a um elemento anteriormente mencionado.

A questão da emergência do clítico e do artigo tem sido considerada na literatura como um importante factor na transição do Latim para as línguas românicas, como assinala Vincent (1997: 149-150):

“I will suggest that in the transition from Latin to Romance we are able to observe the evolution of a pattern of configurationality, and that the differential developments of *ille* and *ipse* constitute one of the overt signals of a typological re-organization of Latin into Romance syntax”.

A hipótese relativamente consensual na literatura (cf. Aebischer (1948), Renzi (1976), Harris (1980)) de que os artigos e os clíticos derivam da mesma fonte parece encontrar, na maior parte das línguas românicas, evidência no facto de haver uma coincidência entre as formas de artigo e de clítico de objecto.

Assume-se, assim, no chamado Latim Clássico, a existência de duas classes de pronomes (na linha da distinção já tradicional e retomada em termos teóricos por Cardinaletti & Starke (1994)): fracos e fortes, não havendo, no entanto, atestações de clíticos pronominais no sentido de Kayne (1975). Só numa fase posterior (de transição entre o Latim e as línguas românicas) se verifica a emergência desta classe (etimologicamente derivada de ILLE e das formas de pronomes pessoais de 1ª e 2ª pessoas). Na terminologia de Cardinaletti & Starke (1994), tal corresponde a uma evolução de acordo com a seguinte escala: clíticos < pronomes fracos < pronomes fortes. A mudança de pronomes fracos para pronomes clíticos teria ocorrido tendo por base os pronomes pessoais de 1ª e 3ª pessoa, bem como a forma ILLE (enquanto marcador de informação dada), os quais ocorriam a par com IPSE (marcando informação topicalizada). Note-se que a reorganização do sistema deíctico latino culmina com a redução gradual deste a estas duas formas básicas (ILLE e IPSE), como se pode observar em (5).

3.1.1. Os (proto-)clíticos em contextos de orações infinitivas (século XI a XIV)

Tendo por base a reorganização do sistema pronominal latino acima descrita, procurámos observar em que fase deste percurso se encontram os (proto-)clíticos que ocorrem em contextos de orações infinitivas no nosso *corpus*.

Para esse efeito, foi dada particular relevância ao paradigma dos pronomes pessoais, bem como aos pronomes deícticos IS, EA, ID e ILLUM, ILLAM, ILLUD,

na qualidade de formas proto-clíticas. As restantes formas do paradigma deíctico assinaladas em (5) não serão aqui consideradas, dado que manifestam um percurso que se afasta consideravelmente do dos pronomes proto-clíticos.⁴

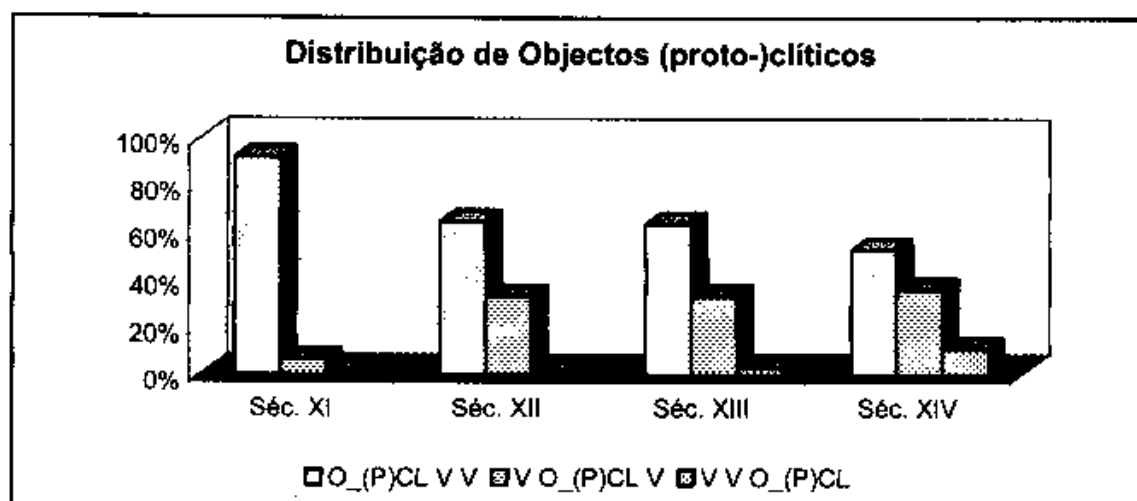
3.1.1.1. Distribuição de Objectos (proto-)clíticos

Num primeiro passo foram identificadas todas as ocorrências de pronomes proto-clíticos e pronomes clíticos em contextos infinitivos, no *corpus* dos séculos XI a XIV. De entre estas, foram consideradas, para efeito dos gráficos em (6) e (7), apenas as formas correspondentes a um único complemento do verbo (o complemento directo no caso de verbos transitivos e o complemento preposicionado, no caso de este ser o único complemento realizado). Deste modo, pretendeu-se delimitar um conjunto de formas, com uma frequência significativa, que pudesse integrar quer sintagmas plenos, quer pronomes (proto-)clíticos.

Partiu-se mais uma vez de descrições já existentes na literatura sobre a distribuição de pronomes clíticos no Português Antigo, relativamente aos quais se assume que a sua colocação diverge da dos DPs plenos. Procurámos assim esclarecer se os pronomes proto-clíticos dos textos médio-latinos e os pronomes clíticos do Português Antigo teriam diferente estatuto, o que se veio a confirmar.

Os dados relevantes são apresentados sob forma de gráfico, em (6):

(6) Gráfico I: Distribuição de Objectos (proto-)clíticos em contextos infinitivos nos séculos XI a XIV:



⁴ Relativamente aos deícticos HIC, HAEC, HOC e ISTE, ISTA, ISTUD, (assumindo que a distribuição destes elementos em contextos infinitivos terá paralelo nos restantes contextos finitos), a sua reduzida frequência parece apontar para um estágio de evolução em que estas formas são já preteridas por outras.

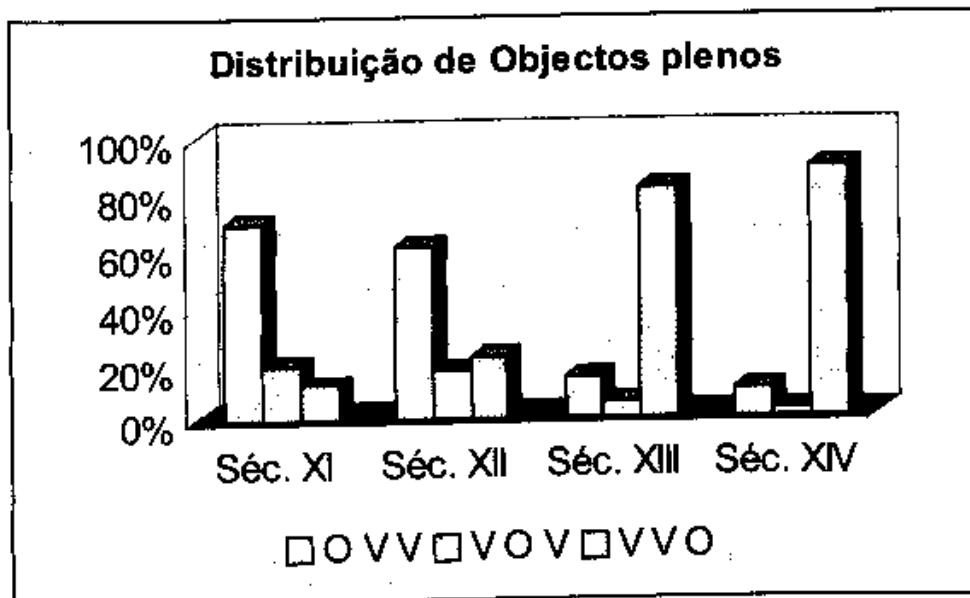
Da observação de (6) é possível concluir que a distribuição dos pronomes proto-clíticos e clíticos ao longo dos quatro séculos não apresenta diferenças consideráveis, ainda que os dados relativos ao século XI se diferenciem dos dos séculos seguintes. A próclise relativamente às duas formas verbais (finita e infinitiva) é claramente a opção mais frequente, ainda que a tendência diacrónica seja a de que esta opção perca terreno face às restantes alternativas, sem deixar, no entanto, de ser o padrão dominante. Note-se que esta evolução preconiza desde já uma tendência de mudança relativa à perda da próclise que se terá tornado mais pronunciada nos séculos XVI e XVII (veja-se Martins (1994)).⁵

Se, em contextos idênticos, a distribuição dos proto-clíticos não parece apresentar, comparativamente à dos pronomes clíticos, diferenças relevantes⁶ quanto a padrões de colocação daqueles, a distribuição dos DPs plenos (nos séculos XI e XII, bem como nos dois séculos subsequentes) poderá fornecer-nos indicações mais precisas quanto ao real estatuto dos proto-clíticos.

3.1.1.2. Distribuição de Objecto pleno

Os dados relevantes encontram-se sob forma de gráfico em (7):

(7) Gráfico II: Distribuição de Objectos plenos em contextos infinitivos nos séculos XI a XIV:



⁵ Tal tendência terá culminado, nos nossos dias e nomeadamente nas camadas mais jovens, pela crescente opção pelo uso de ênclise em detrimento da próclise (veja-se, nomeadamente Duarte & Matos (2000)).

⁶ As diferenças mais significativas parecem residir não numa mudança de padrões de colocação, mas antes numa maior ou menor expressividade em termos quantitativos.

Abstraindo da ordem relativa das duas formas verbais ($V_{FIN} V_{INF}$ ou $V_{INF} V_{FIN}$), observa-se que o gráfico (7) apresenta até ao século XII inclusivé, e em termos gerais, padrões de distribuição semelhantes aos verificados no gráfico (6) para os Objectos representados pelos pronomes proto-clíticos. A colocação do proto-clítico à esquerda dos dois verbos assemelha-se à distribuição dos Objectos plenos ocupando maioritariamente uma posição à esquerda dos dois verbos (em consequência do movimento do DP via *scrambling*). A única diferença detectada parece residir no aumento significativo da ordem $V V O_{pleno}$ já no século XII, com a percentagem de atestações de objecto final (22%) a ultrapassar as ocorrências em que o Objecto se encontra em posição medial (17%). Ainda assim, a opção pela periferia esquerda permanece claramente o padrão dominante, com 61% das atestações. Note-se que, nos nossos dados, tal distribuição não terá correspondente na evolução dos pronomes (proto-)clíticos pelo menos até ao século XIV.

Há ainda a assinalar que é justamente nos séculos XIII e XIV que se verifica uma distribuição radicalmente distinta entre Objectos (proto-)clíticos e dos DPs plenos, com os DPs plenos a ocorrerem, em larga maioria, na periferia direita dos dois verbos (81% no século XIII e 88,5% no século XIV) e com os clíticos maioritariamente proclíticos (64% no século XIII e 53% no século XIV).

3.1.1.3. Do estatuto XP a X^o

A distribuição paralela dos proto-clíticos e dos DPs plenos nos séculos XI e XII não deverá constituir surpresa, tendo-se verificado posteriormente a evolução dos primeiros no sentido de um progressivo enfraquecimento. Tal desenvolvimento não é aliás, único no domínio das línguas românicas: já Menéndez Pidal (1964: II, 30) havia notado a distribuição paralela entre pronominais plenos e (proto-)clíticos (dativos e acusativos) em Castelhana Antigo.

Defenderemos, pois (na linha de Rivero (1986) e Duarte & Matos (2000)) que, diacronicamente, os pronomes (proto-)clíticos terão passado de um estatuto sintagmático (i.e., são núcleos de projecções máximas ou XPs, com a mesma distribuição dos pronomes tónicos e dos DPs não-pronominais) para o de núcleos (X^os). Para Duarte & Matos (2000), as diferenças observadas entre estádios de língua como o Português Antigo e o Português Europeu Contemporâneo resultam de diferentes condições de verificação de traços, (concretamente das propriedades formais intrínsecas dos clíticos, os quais apresentam um estatuto híbrido entre X^o e XP). Com base em dados da diacronia e da aquisição da linguagem, Duarte & Matos (2000) consideram que a diferente distribuição dos pronomes clíticos pode ser explicada assumindo uma mudança na especificação formal dos clíticos quanto ao tipo de hospedeiro que estes requerem. Os (proto-)clíticos terão assim evoluído a partir de um estádio inicial, atestado nos nossos textos dos séculos XI a XIV, no qual a distribuição padrão é a que ocupa a posição mais à esquerda e os (proto-)clíticos apresentam um estatuto não-específico de tipo XP. No decurso da sua evolução, estes

elementos sofrem um processo de reanálise que lhes confere um estatuto afixal: a ênclise generaliza-se como ordem não-marcada, com o clítico e o seu hospedeiro a ocorrerem num mesmo núcleo funcional AgrO. Tal mudança traduz-se assim, de acordo com Duarte & Matos (2000), numa maior especificação do clítico, que passa a ser formalmente especificado como núcleo (X^o) hóspede de V (“V-host”), adjungindo-se apenas a núcleos verbais.

3.2. A Interpolação

Outro dos argumentos em favor do estatuto sintagmático dos proto-clíticos e dos clíticos (ainda que em diferentes graus) baseia-se na atestação, no nosso *corpus* textual, do fenómeno de interpolação.

3.2.1. Descrição do fenómeno:

O mecanismo de interpolação consiste na extração do (proto-)clítico e sua adjunção a uma determinada posição, que assumiremos ser, no período por nós considerado, de adjunção indiferentemente a categorias máximas e mínimas. O (proto-)clítico pode assim surgir separado do verbo temporalizado, quebrando-se a adjacência estrita entre estes dois elementos:

- (8) a. proinde calumniari temptauerit et nos **eam** uobis in iudicio auxilio uestro adiuti diuindicare et uestre parti auctorizare noluerimus aut non ualuerimus (1094, D&C351)
 “e nós a_PCL vos_PCL em juízo ajudados pelo vosso auxílio vingar (...) não quisermos”
- b. ada de alchaide nichil accipiat alchaide per uim, nisi quod **ei** milites amore suo dare uoluerint. (1179, DR296)
 “nada aceite o alcaide por força, nem o que lhe os soldados, por seu amor, quiserem dar”
- c. pero se **lha** el nō quis(er) tolh(er) nē se(us) h(er)deyros nō lha tolhã. (1280?, FR, 117r)
- d. u en seu au(er) qual q(ue)r corregalho pela soldada & se **o** pela soldada nō pode coreger correga-o p(er) q(ua)nto lhy acharẽ & se corpo pelo corpo (1340-1360, CS1)

Em (8a.), o proto-clítico *eam* ocorre à esquerda de V, tendo sido interpolados à sua direita a seguinte sequência de constituintes: objecto indirecto <sintagma preposicional adjunto <sintagma participial <verbo infinitivo. Já em (8b.), a adjacência entre o proto-clítico dativo *ei* e o verbo é quebrada pela sequência: sujeito <sintagma adjunto <verbo infinitivo. A interpolação é ainda atestada nos séculos XIII e XIV, como ilustrado em (8c.) e (8d.), que apresentam, respectivamente, as seguintes

seqüências de constituintes interpolados: sujeito <advérbio de negação e sintagma preposicional <advérbio de negação.

Enquanto evidência para uma abordagem dos (proto-)clíticos como morfemas que não são especificados relativamente ao tipo de hóspede ao qual se adjungem – tanto se podem adjungir a XP (categorias sintagmáticas) como a X^o (núcleos) – a interpolação apresenta alguma variação diacrónica quanto à sua frequência.

3.2.2. Frequência da interpolação

Para efeito das tabelas que se seguem, foram consideradas a totalidade de formas (proto-)clíticas atestadas em contextos infinitivos no nosso *corpus* (1959 ocorrências).

Note-se que a questão da real frequência dos contextos de interpolação (como assinalado nomeadamente por Martins) deverá ser avaliada tendo em conta os contextos que efectivamente apresentam interpolação comparativamente àqueles que, apresentando constituintes em posição passível de permitir a interpolação, não apresentam a distribuição de interpolação. Tal metodologia permite assim, de forma mais correcta, excluir, de entre os contextos que evidenciam próclise, aqueles que não manifestando interpolação, também não deverão ser considerados como contextos de não-interpolação, uma vez que não apresentam uma configuração potencial de interpolação. Concretamente, apenas foram consideradas como configurações em que potencialmente poderá ocorrer interpolação aquelas que reúnem os seguintes elementos: um constituinte desencadeador de próclise, um pronome (proto-)clítico e um ou mais constituintes situados à esquerda do verbo finito, em potencial posição de interpolação. No universo destas configurações, teremos assim, por um lado, aquelas em que realmente ocorre a interpolação (cf. nomeadamente (8)); por outro lado, os verdadeiros casos de não-interpolação (como os apresentados em (9)):

- (9) a. *damus a uobis ipsa ereditate per ubi uos illa potueritis inuenire cum omni tota sua prestatione.* (1100, D&C416)
 “damos a vós a mesma herdade por onde vós a puderdes encontrar”
- b. *Si autem eum noluerint adiuuare* quos uocauerit #V solidos pariant et sint contradicti concilio. (1183, DR313)
 “Se também o não quiserem auxiliar, os que chamar paguem 5 soldos”
- c. E se d(e)poys o poder au(er) ou de q(ue) o peytar, d(é) o s(er)uo ((ou))o p(re)ço a sseu dono e seya quite. (1280?, FR, 141ar)
- d. Este custume se guarda p(er) esta guisa se cõ a ferida assíjnada o q(ue)r faz(er) p(er) íuramêto. pode-lhe diz(er) a outra p(ar)te que o nõ pode ffazer p(er) seu íuramento (1350c, CS3, 6va)

À luz deste critério, a frequência de contextos de interpolação apresenta-se em (10):

(10) Tabela III: Frequência de (proto-)clíticos, próclise, configurações potenciais de interpolação, de verdadeira interpolação e de não-interpolação no total de contextos infinitivos:

	(Proto-)clíticos	Próclise (P)CL V	Configurações Potenciais de Interpolação	
			Interpolação	Não-Interpolação
Século XI	267	205 = 77%	57/60 = 95%	3/60 = 5%
Século XII	198	120 = 61%	88/93 = 95%	5/93 = 5%
Século XIII	743	425 = 57%	88/177 = 50%	89/177 = 50%
Século XIV	751	330 = 44%	91/139 = 65%	48/139 = 35%
TOTAL	1959	1080	324/469	145/469

De acordo com os dados apresentados em (10), a proporção de contextos em que ocorre interpolação é, relativamente ao total de potenciais contextos passíveis de apresentarem interpolação, a opção claramente mais frequente nos nossos dados dos séculos XI e XII (com 95% de contextos de interpolação). Tal frequência decresce nos dados do Português Antigo dos séculos XIII e XIV, sem que, no entanto, desça abaixo dos 50% (50% no século XIII e 65% no século XIV). Note-se que, desta forma, se torna evidente a elevada frequência da interpolação nos quatro séculos que aqui consideramos. Tal distribuição corrobora igualmente a existência de *scrambling* nos períodos assinalados, fenómeno relativamente ao qual a frequência da interpolação é usada como teste diagnóstico.

4. Conclusões:

Com o presente trabalho pretendeu-se analisar a variação OV/VO em contextos infinitivos, dum ponto de vista diacrónico (concretamente em textos não literários produzidos em território português entre os séculos XI e XIV).

Assumimos, como hipótese de trabalho, a proposta de Kayne (1994) segundo a qual as línguas são universalmente VO, sendo a ordem OV derivada em resultado da aplicação da regra de movimento. Esta hipótese tem sido avançada para as línguas que evidenciam o fenómeno designado de *scrambling* por Zwart (1997) para o holandês e Roberts (1997) para estádios antigos do Inglês. De igual forma, defendemos que os estádios de língua retratados quer pelos nossos textos latino-

-portugueses (séculos XI e XII), quer pelos relativos ao Português Antigo (séculos XIII e XIV), partilham uma mesma ordem subjacente VO, ainda que, nos primeiros, as ordens atestadas impliquem que *scrambling* se aplica em larga escala.

Tendo em consideração a proposta de Martins (2000b), segundo a qual os contextos de interpolação se qualificam como testes diagnósticos fiáveis para a existência de *scrambling* (com o clítico a assinalar a fronteira entre o constituinte deslocado à esquerda ou constituintes focalizados, por um lado, e o constituinte movido por *scrambling*, por outro), procedemos à análise dos contextos em que ocorrem pronomes proto-(clíticos). A observação da frequência de Objectos proto-(clíticos) e plenos nos nossos dados vem corroborar a hipótese de que, diacronicamente, os primeiros terão evoluído a partir de um estatuto sintagmático (i.e., são núcleos de projecções máximas ou XPs) para o de núcleos (X's).

Qualificando-se como contexto diagnóstico para a existência de *scrambling*, o facto de haver mais interpolação nos textos médio-latinos (com 95% nos dois séculos) do que nos do Português Antigo (50% e 60% nos séculos XIII e XIV, respectivamente) vem justamente confirmar a ideia de que a produtividade da interpolação terá constituído evidência positiva para a aquisição de *scrambling*, com a posição do (proto-)clítico a servir de diagnóstico.

Referências bibliográficas:

- AEBISCHER, Paul (1948) "Contribution à la proto-histoire des articles *ille* et *ipse* dans les langues romanes", *Cultura Neolatina*, 8, pp. 181-203.
- BELLETTI, A. & U. Shlonsky (1995) "The Order of Verbal Complements: A Comparative Study" *Natural Language and Linguistic Theory*, 13, pp. 489-526.
- CARDINALETTI, Anna, & Michal Starke (1994) "The Typology of Structural Deficiency: a Case Study of Three Classes of Pronouns" In Henk van Riemsdijk (ed.) (1999) *Clitics in the Languages of Europe*, Berlin/New York: Mouton de Gruyter, pp. 33-82.
- COSTA, João (1998) *Word Order Variation. A constraint-based approach*, The Hague: Holland Academic Graphics.
- DILLET, Montse Batllori (1998) "El sintagma nominal: génesis del artículo y distribución de los complementos nominales desde uma perspectiva diacrónica" in *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, vol. 1: *Grammatica storica delle lingue romanze*, Niemeyer: Tübingen, pp. 57-67.
- DUARTE, I. & G. MATOS (2000) "Romance Clitics and the Minimalist Program" in Costa (ed.) *New Comparative Studies in Portuguese Syntax*, Oxford: Oxford University Press, pp. 116-142.
- HARRIS, M. (1980) "The Marking of Definiteness in Romance" in Fisiak, J. (ed.) *Historical Morphology*, Berlin: Mouton de Gruyter, pp. 141-156.
- HUBER, Joseph (1933) *Gramática do Português Antigo*, Lisboa: Fundação Calouste Gulbenkian, 1986, Tradução do alemão por M. Manuela G. Delille.
- KAYNE, R. S. (1994) *The Antisymmetry of Syntax*, Cambridge, Massachusetts: MIT Press.

- KAYNE, R. S. (1975) *Syntaxe du français – Le cycle transformationnel* (trad. Pierre Attal), Éditions du Seuil, Travaux Linguistiques Coll., Paris.
- LIGHTFOOT, D. (1991) *How to Set Parameters: Arguments from Language Change*, Cambridge, Massachusetts: MIT Press.
- MAROUZEAU, J. (1922) *L'ordre des mots dans la phrase latine*, (vol. I: *Les groupes nominaux*), Paris: Champion.
- MARTINS, A. M^a (2000a) "Parameter setting in the history of Portuguese: The loss of OV/VO – Considerations on clause structure, word order variation and change", comunicação apresentada ao Colóquio "Lisbon 2000: Statistical Physics, Pattern Identification and Language Change", Universidade de Lisboa.
- MARTINS, A. M^a. (2000b) "On word order variation and change: the loss of IP-scrambling in Portuguese", comunicação apresentada ao Going Romance 2000, Universiteit Utrecht.
- MARTINS, A. M^a. (1994) *Clíticos na História do Português*, Dissertação de Doutoramento, Lisboa: FLUL.
- MATTOS E SILVA, Rosa V. (1989) *Estruturas Trecentistas. Elementos para uma Gramática do Português Arcaico*, Lisboa: INCM.
- MENÉNDEZ PIDAL, Ramón [1926] [1950³] (1980⁹) *Orígenes del Español: Estado lingüístico de la Península Ibérica hasta el siglo XI*, Madrid: Espasa-Calpe, 9^a ed. (segundo a 3^a ed., com correcções e adições).
- MENÉNDEZ PIDAL, Ramón 1964 (ed.) *Cantar de Mio Cid: Texto, gramática y vocabulario*, Madrid: Espasa-Calpe, 4^a edição.
- OGANDO, V. (1980) "A colocación do pronome átono en relación co verbo no galego-português medieval", *Verba* 7, pp. 251-282.
- PÁDUA, M. P. Canaes e M. de (1960) *A Ordem das Palavras no Português Arcaico (Frasas de Verbo Transitivo)*, Coimbra: Instituto de Estudos Românicos.
- PANHUIS, Dirk G. (1982) *The Communicative Perspective in the Sentence. A Study of Latin Word Order*. Amsterdam: John Benjamins.
- PENNY (1993) *Gramática histórica del español*, Barcelona: Ariel.
- PINKSTER, Harm (1995) *Sintaxis y Semántica del Latín*, Madrid: Ediciones Clásicas. (trad. esp. de M. E. Torrego e J. de la Villa).
- RENZI, L. (1976) "Grammatica e storia dell' articolo italiano", *Studi di Grammatica Italiana* 5, pp. 5-42.
- RIVERO, M. Luisa (1986) "Parameters in the Typology of Clitics in Romance and Old Spanish" *Language* 62, pp. 774-807.
- ROBERTS, Ian (1997) "Directionality and word order change in the history of English" in A. van Kemenade & N. Vincent (eds.), *Parameters of Morphosyntactic Change*, Cambridge: Cambridge University Press, pp. 397-426.
- ROSS, J. R. (1967) *Constraints on Variables in Syntax*, Dissertação de Doutoramento, MIT. [publicado em 1986 como *Infinite Syntax!*, Norwood NJ: Ablex.]
- THRÁINSSON, Höskuldur (2001) "Object Shift and Scrambling" in M. Baltin, & Chris Collins (eds.) *The Handbook of Contemporary Syntactic Theory*, Oxford: Blackwell, pp. 148-202.
- VINCENT, N. (1997) "The emergence of the D-system in Romance" in A. von Kemenade & N. Vincent (eds.) *Parameters of Morpho-syntactic Change*, Cambridge: Cambridge University Press, pp. 149-169.

- WANNER, Dieter (1987) *The Development of Romance Clitic Pronouns. From Latin to Old Romance*, Berlin: Mouton de Gruyter.
- WEERMAN, F. (1997) "On the relations between morphological and syntactic case", in A. van Kemenade & N. Vincent (eds.) *Parameters of Morphosyntactic Change*, Cambridge: Cambridge University Press, pp. 427-459.
- ZWART, C. J.-W. (1997) *The Morphosyntax of Verb Movement: A Minimalist Approach to Syntax of Dutch*, Dordrecht: Kluwer.