

Construções perceptivas e ordem de palavras num texto medieval

Sandra Pereira

Centro de Linguística da Universidade de Lisboa (Portugal)

spereira@clul.ul.pt

Recibido o 12/01/15. Aceptado o 01/06/15

Perception constructions and word order in an Old Portuguese text

Resumo

A análise da ordem de palavras envolvidas nas estruturas de verbos perceptivos (*ouvir*, *sentir* e *ver*) com complementos infinitivos em português antigo constitui-se como o principal objetivo deste estudo. Partindo de trabalhos anteriores para o português europeu contemporâneo padrão (cf. Gonçalves 1999; Duarte & Gonçalves 2002) e dialetal (Pereira 2012), este estudo destaca a posição e a forma do chamado “sujeito da oração infinitiva”, pretendendo contribuir para o estudo das construções perceptivas (e, por extensão, das causativas, uma vez que apresentam semelhanças estruturais) na história do português.

Este estudo visa igualmente dar conta das vantagens da utilização de um sistema de anotação sintática para o estudo das propriedades gramaticais de construções muito específicas. Essa demonstração é feita a partir de um texto medieval, *O Livro de José de Arimateia*, na edição de Castro (1984). Este texto foi recentemente anotado (morfológica e sintaticamente) no âmbito do projeto WOChWEL, que segue os sistemas de anotação da equipa do *Tycho Brahe Corpus*, fortemente inspirados na anotação dos *Penn corpora*.

Palabras chave

Construções perceptivas, ordem de palavras, português antigo, anotação de corpus

Sumario

1. Introdução. 2. Projeto WOChWEL. 3. Construções Perceptivas em português europeu contemporâneo. 3.1. Construção de infinitivo flexionado. 3.2. Construção de Marcação Excepcional de Caso. 3.3. Construção de União de Orações. 3.4. Construção infinitiva gerundiva. 3.5. Construção com infinitivo flexionado e sujeito acusativo. 4. Anotação sintática das construções perceptivas. 5. Pesquisa das construções perceptivas. 6. Resultados. 7. Conclusões.

Abstract

The main objective of this paper is to analyse word order in structures with perception verbs (*to hear*, *to feel* and *to see*) with infinitive complements in Old Portuguese. Based on previous work for standard (cf. Gonçalves 1999; Duarte & Gonçalves 2002) and dialectal (Pereira 2012) contemporary European Portuguese, this study highlights the position and shape of the subject of the infinitive clause. It aims to contribute to the study of perception constructions (and, by extension, causative constructions, since they have structural similarities) in the history of Portuguese.

In this paper we also show the benefits of using a syntactic annotation system for the study of grammatical properties of highly specific constructions. The text under analysis is *O Livro de José de Arimateia*, in the edition of Castro (1984). This text was recently tagged and parsed within the framework of the WOChWEL project, which follows the annotation systems of the team of the *Tycho Brahe Corpus*, strongly influenced by the annotation of the *Penn corpora*.

Keywords

Perception constructions, word order, Old Portuguese, corpus annotation

Contents

1. Introduction. 2. The WOChWEL project. 3. Perception Constructions in contemporary European Portuguese. 3.1. Inflected infinitive construction. 3.2. Exceptional Case Marking Construction. 3.3. *Faire-Inf* construction. 3.4. The prepositional infinitive construction. 3.5. Inflected infinitive construction with accusative subject. 4. Syntactic annotation of perception constructions. 5. Searching perception constructions. 6. Results. 7. Conclusions.

1. INTRODUÇÃO

As construções causativas e perceptivas em português europeu contemporâneo (daqui em diante, PEC) foram já objeto de vários trabalhos (cf. para o PEC padrão, Raposo 1981, 1989; Gonçalves 1999; Gonçalves e Duarte 2001; Duarte e Gonçalves 2002; Silva 2005; Bossaglia 2013; e, para o PEC dialetal, Carrilho e Pereira 2010; Pereira 2012). Uma vez que os verbos causativos e perceptivos podem selecionar, de um modo geral, as mesmas construções não finitas, muitos autores tratam as construções causativas e as perceptivas em conjunto. Neste trabalho, optou-se por descrevê-las separadamente¹, partindo da assunção de que, em PEC: (i) uma construção como a infinitiva gerundiva, também chamada de PIC (do inglês, *Prepositional Infinitival Construction*) só é possível com verbos perceptivos e (ii) a denominada construção *Faire-par* (Guasti 2006), praticamente inexistente em PEC, só se atesta com verbos causativos. Adicionalmente, dando-se ênfase apenas às construções perceptivas em português antigo (daqui em diante, PA), a descrição sobre a ordem de palavras nas construções que ocorrem neste período torna-se mais exaustiva e refinada.

O recurso a *corpora* anotados sintaticamente traz vantagens evidentes. Uma vantagem incontornável é o facto de a pesquisa por construção (neste caso específico, as construções perceptivas) ser facilitada:

- i) os resultados são mais limitados e, por isso, a tarefa de seleção é mais fácil;
- ii) é possível refinar buscas, restringindo, por exemplo, ordem de palavras, tipo de preposição, etc;
- iii) permite buscas por categorias vazias.

Uma outra vantagem é a possibilidade de aparecerem resultados surpreendentes (como por exemplo, verbos inesperados que permitem determinada construção) que numa pesquisa por palavra não seriam encontrados. Além disso, no caso específico dos textos do WOChWEL, usar a anotação sintática para pesquisas diminui os impactos da falta de lematização e da não modernização gráfica dos textos.

2. PROJETO WOChWEL

Este trabalho insere-se no âmbito do projeto WOChWEL (*Word Order and Word Order Change in Western European Languages*)². É um projeto em curso no Centro de Linguística da Universidade de Lisboa que tem como objetivo fundamental o estudo da ordem de palavras numa perspectiva comparada, inspecionando vários tópicos como a ordem VS, extraposição, colocação dos pronomes clíticos, negação, entre outros. Nesse sentido, a anotação de textos do período mais antigo da língua constitui-se como objetivo paralelo. De facto, a disponibilização de recursos linguísticos relativos ao PA visa suprir uma lacuna existente no panorama português, colocando o português a par de outras línguas europeias.

Os dados usados neste estudo foram extraídos de *O Livro de José de Arimateia*, na edição de Castro (1984). Trata-se de um texto da literatura medieval (do ciclo arturiano Post-Vulgata) que chegou aos nossos dias através de uma cópia do século XVI a partir de um original do século XIII. Este é o primeiro de dois textos literários a ser anotados no âmbito do projeto WOChWEL³. Cada um dos textos é submetido a dois tipos de anotação: a anotação morfológica e a anotação sintática. A anotação morfológica é feita através de etiquetas POS (*Part of Speech*) e foi estabelecida pela equipa do projeto *Tycho Brahe Parsed Corpus of Historical Portuguese (Tycho Brahe Corpus)*⁴. Este tipo de anotação permite pesquisas mais simples, centradas na palavra e/ou na etiqueta POS e serve também de *input* à anotação sintática. A anotação sintática segue os princípios do

¹ Sobre um estudo das construções causativas no mesmo texto medieval, veja-se Pereira (2014).

² O WOChWEL é um projeto financiado pela Fundação para a Ciência e Tecnologia (FCT: PTDC/CLE-LIN/121707/2010). Os recursos produzidos no seu âmbito estão disponibilizados livremente em: <http://alfclul.clul.ul.pt/wochwel/oldtexts.html>. Informações relevantes sobre o projeto podem também ser encontradas no mesmo endereço eletrónico.

³ Está neste momento em curso a anotação de um outro texto, *A Demanda do Santo Graal* (cf. Toledo Neto, 2012-2015; Martins e Pereira 2014-2015; Martins et al. 2014-2015). Pretende-se ainda fazer a anotação de alguns textos notariais do mesmo período.

⁴ Cf. <http://www.tycho.iel.unicamp.br/~tycho/corpus/en/>

sistema de anotação dos *Penn Corpora of Historical English (Penn Corpora)*⁵ e, para o português, do *Tycho Brahe Corpus*, em que a informação é codificada de maneira a que construções sintáticas relevantes possam ser facilmente pesquisáveis. As pesquisas são feitas recorrendo ao *CorpusSearch*⁶, um motor de busca que, através de uma linguagem de *query* básica, oferece várias possibilidades de procura de estruturas sintáticas (cf. Carrilho e Magro 2010).

3. CONSTRUÇÕES PERCEPTIVAS EM PORTUGUÊS EUROPEU CONTEMPORÂNEO

Duarte e Gonçalves (2002) referem quatro construções não finitas selecionadas por verbos perceptivos: i) a construção com infinitivo flexionado; ii) a construção de Marcação Excepcional de Caso (ECM, do inglês *Exceptional Case Marking*); iii) a construção com predicado complexo (também chamada *União de Orações*, designação adotada neste trabalho, ou *Faire-inf*) e iv) a construção infinitiva gerundiva (PIC, cf. Raposo 1989). Além destas, Martins (2006: n4) e Bossaglia (2013) identificaram uma outra construção com infinitivo flexionado e sujeito acusativo.

3.1. Construção de infinitivo flexionado

Na *construção de infinitivo flexionado*, o verbo infinitivo exibe marcas de flexão e o sujeito⁷ da oração infinitiva é um sintagma nominal ou um pronome nominativo que ocorre entre os dois verbos:

(1a) A Maria viu as meninas brincarem.

(1b) A Maria viu elas brincarem.

3.2. Construção de Marcação Excepcional de Caso

Na *construção de ECM*, o infinitivo não está flexionado e o sujeito da oração não finita é um pronome clítico acusativo (cuja subida é obrigatória) ou um sintagma nominal, ocorrendo este entre os dois verbos:

(2a) A Maria viu as meninas brincar.

(2b) A Maria viu-as brincar.

(2c) A Maria viu as meninas **não** brincar.

(2d) A Maria viu as meninas penteá-**las** (**las** = as bonecas).

O domínio infinitivo tem alguma autonomia sintática permitindo a ocorrência de negação frásica (como em (2c)) e de clíticos complementos (como em (2d)).

Repare-se que no caso em que o sujeito da oração encaixada é um sintagma nominal no singular, há ambiguidade estrutural entre a construção de infinitivo flexionado e a construção de ECM:

(3) A Maria viu a menina brincar.

Como o verbo infinitivo não exibe flexão na terceira pessoa do singular e não é possível dizer se o sintagma nominal recebe caso nominativo (construção de infinitivo flexionado) ou acusativo (construção de ECM), frases como (3) são ambíguas em PEC.

⁵ Cf. <http://www.ling.upenn.edu/histcorpora/>

⁶ Cf. <http://corpussearch.sourceforge.net/>

⁷ Ao longo de todo o texto, a referência ao “sujeito da oração infinitiva” ou ao “sujeito encaixado” é feita de forma generalizada, mesmo quando diz respeito à construção de União de Orações, caso em que esse constituinte não é, na realidade, o sujeito (cf. Duarte e Gonçalves 2002: 164, n2).

3.3. Construção de União de Orações

Na construção de *União de Orações*, o sujeito encaixado é um pronome acusativo ou um sintagma nominal, que ocorre após a sequência verbal:

- (4a) A Maria viu brincar *as meninas*.
 (4b) A Maria viu-as brincar.
 (4c) *A Maria viu não brincar *as meninas*.
 (4d) *A Maria viu penteá-las *as meninas* (**las** = as bonecas).

O facto de o domínio infinitivo desta construção não legitimar negação frásica (como em (4c)) nem clíticos complementos (como em (4d)) torna-o mais defetivo do que o da construção anterior.

Ainda assim, note-se que as frases (2b) e (4b) são estruturalmente iguais mostrando a ambiguidade que existe entre a construção de ECM e a construção de União de Orações nos casos em que o verbo infinitivo é monoargumental (*brincar*) e o sujeito é um pronome acusativo (*as*).

Um outro aspeto da construção de União de Orações evidenciado por Duarte e Gonçalves (2002: 163) é a restrição encontrada na construção de União de Orações: o verbo infinitivo não pode ser transitivo, optando-se, nesse caso, pela construção ECM⁸, como fica demonstrado pelo contraste entre (5) e (6):

- (5) *Os espectadores viram desperdiçar o golo aos sadinos.
 (6) Os espectadores viram os sadinos desperdiçar o golo. [Duarte e Gonçalves (2002: 163)]

No entanto, Pereira (2012: 364) observa que essa restrição não se atesta nos dados dialetais (extraídos do *corpus* CORDIAL-SIN⁹), sendo possível encontrar frases em que o verbo encaixado é transitivo e o sujeito da oração infinitiva é um sintagma preposicional (cf. (7) e (8)) ou um pronome dativo (cf. (9))¹⁰:

- (7) Eu ainda vi fazer carvão *ao meu pai* também. C-UNS42
 (8) E tenho ouvido falar até *a mais pessoas*. S-AAL14
 (9) Mas ouvia *lhe* falar à mãe do meu marido, que ela usava muito essa barrela. N-STA34

3.4. Construção infinitiva gerundiva

De acordo com Raposo (1989) e Duarte e Gonçalves (2002), e.o., nestas construções o infinitivo não flexionado e o infinitivo flexionado encontram-se em variação livre e os clíticos associados à posição de sujeito do domínio encaixado ocorrem na forma acusativa:

- (10a) A Maria viu *as meninas* a brincar(**em**).
 (10b) A Maria viu-as a brincar(**em**).

⁸ É de salientar que estas restrições não se aplicam aos verbos infinitivos nas construções causativas:

i) O João mandou comer o bolo ao João.

⁹ *Corpus Dialetoal para o Estudo da Sintaxe*, disponível em: <http://www.clul.ul.pt/pt/investigacao/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects>

¹⁰ Pereira (2012: 366-368) encontrou outras duas construções não padrão:

(i) a construção de infinitivo pessoal não flexionado (o sujeito da oração infinitiva é um pronome nominativo e o infinitivo não exibe flexão):

i) Já vimos *elas* ir daqui para ali – assim cinco seguidas! A-TRC40

(ii) a construção de infinitivo preposicionado com sujeito nominativo (o sujeito da oração infinitiva é um pronome nominativo e o infinitivo é preposicionado):

ii) Mas o velho estava em casa a ver *ela* a fazer aquele guisado. A-CLH16

3.5. Construção com infinitivo flexionado e sujeito acusativo

Esta construção, referida em Martins (2006: n4) e Bossaglia (2013), caracteriza-se por exibir flexão no infinitivo e o sujeito da oração infinitiva, quando pronominaliza, é um pronome acusativo adjacente ao verbo finito:

- (11) Nunca os ouvi reclamarem o ordenamento do território.
 (12) Vi-os assinarem a paz com a maioria dos clãs. [Bossaglia (2013: 228)]

Gonçalves e Duarte (2001) e Duarte e Gonçalves (2002), ao estudarem as construções causativas e perceptivas, consideram haver, em PEC, uma defetividade crescente associada a estas construções (exceto a PIC)¹¹:

- (13) Infinitivo flexionado < ECM < Predicado Complexo [Gonçalves e Duarte (2001: 669)]

O facto de na formação do Predicado Complexo não ser possível a cliticização de pronomes complementos nem a presença do operador de negação frásica no domínio infinitivo torna a construção de União de Orações mais defetiva do que a construção de ECM, que permite ambos os fenómenos.

Martins (2006), num estudo sobre construções infinitivas na história do português, defende que, em PA, a construção ECM era mais defetiva do que atualmente, aproximando-se da construção de União de Orações quanto à cliticização de pronomes complementos e ao operador de negação frásica.

A forma e a posição do sujeito da oração infinitiva manifestam esta defetividade, uma vez que são diferentes em cada uma das construções. É neste aspeto fundamental que reside a particularidade da anotação sintática destas construções, como se demonstra na secção seguinte.

4. ANOTAÇÃO SINTÁTICA DAS CONSTRUÇÕES PERCEPTIVAS

O processo de anotação sintática inicia-se com a anotação morfossintática. Nesta secção são explicitadas todas as fases referentes ao processo de anotação.

Numa primeira fase, as frases do texto são submetidas ao *tagger* (cf. Kepler 2005), o qual é parte integrante da ferramenta *Edictor*, concebida pela equipa do projeto *Tycho Brahe Corpus*¹². O *tagger* atribui uma etiqueta morfossintática a cada palavra e um código de identificação a cada frase. Assim, em (14) e (15) apresenta-se uma frase antes e depois de ser submetida ao *tagger*, respetivamente:

- (14) E assi o fez e assentou-se nela.
 (15) E/CONJ assi/ADV o/CL fez/VB-D e/CONJ assentou-se/VB-D+SE nela/P+PRO ./ JAR41,3.46/ID

Depois de corrigidas todas as frases, a versão do texto com anotação morfossintática encontra-se completa e é disponibilizada para buscas mais simples e baseadas essencialmente na informação morfológica¹³. Paralelamente, esta versão constitui o *input* do *parser*, depois de feitas algumas operações no sentido de descontrair contrações e separar clíticos em posição enclítica. Veja-se as diferenças entre a frase (15) e a frase (16), após a preparação para o *parser*:

- (16) E/CONJ assi/ADV o/CL fez/VB-D e/CONJ assentou@/VB-D @se/CL em@/P @ela/PRO.

¹¹ Num outro quadro teórico, Silva (2012: 528)) propõe uma escala de gramaticalização das diferentes construções: VSV > VOV > VV, em que VV, isto é, a construção de União de Orações, "represents the highest degree of constructional grammaticalization while VSV [construção de infinitivo flexionado] the lowest".

¹² Esta ferramenta pode ser descarregada a partir da página do projeto *Tycho Brahe*: <http://www.tycho.iel.unicamp.br/corpus/index.html>

¹³ No caso de *O Livro de José de Arimateia*, a versão anotada morfossintaticamente (cf. Martins e Pereira 2012-2015) encontra-se disponível em: http://alfclul.clul.ul.pt/wochwel/documents/Jose_Arimateia.txt

Ou seja, “@” marca a existência prévia de uma contração entre dois constituintes (entre um verbo e um clítico e entre uma preposição e um pronome, em (16)).

O parser usado é o mesmo do *Tycho Brahe Corpus* (cf. Dan Bikel 2004). O resultado é um ficheiro de texto (ASCII) em que as frases apresentam a forma de parentetização indentada, como se mostra seguidamente:

```
( (IP-MAT (IP-MAT (CONJ E)
      (NP-SBJ *pro*)
      (ADVP (ADV assi))
      (NP-ACC (CL o))
      (VB-D fez))
  (CONJP (CONJ e)
    (IP-MAT (NP-SBJ *pro*)
      (VB-D assentou@)
      (NP-SE (CL @se))
      (PP (P em@)
        (NP (PRO @ela))))))
  (. .))
(ID JAR41, .42))
```

Além de um qualquer editor de texto, outros programas que permitem a edição deste *output* são o *CorpusDraw* (parte integrante do programa *CorpusSearch*) e o *Annotald*¹⁴. Enquanto neste a visualização é mais próxima da visualização do editor de texto (com indentação colorida), no primeiro as frases são visualizadas sob a forma de árvore. A correção semiautomática, recorrendo a *queries*, só se pode efetuar com o *CorpusDraw*.

O sistema de anotação sintática adotado pelo projeto WOCHWEL guia-se pelo manual de anotação sintática do *Tycho Brahe Corpus*, adaptado dos *Penn Corpora*. Trata-se de um sistema de anotação de constituintes que pretende ser teoricamente descomprometida, codificando fronteiras de constituinte e dependências estruturais, ao mesmo tempo que fornece informação categorial (se é um NP, PP, etc.¹⁵), gramatical (SBJ, ACC, DAT, VOC) e oracional (por exemplo, se é um oração comparativa – COMP –, adverbial – ADV –, exclamativa – EXL, etc.). São também codificados constituintes nulos (como os sujeitos) e alguns movimentos (como extraposição, deslocação à esquerda, subida do clítico, etc.).

Apresentam-se na Tabela 1 as etiquetas sintagmáticas, as etiquetas oracionais ou frásicas e as categorias vazias relevantes para a anotação das construções percetivas¹⁶.

¹⁴ Cf. <http://annotald.github.io/user.html> Este programa foi desenvolvido no âmbito da anotação de um *corpus* do islandês, o *Icelandic Parsed Historical Corpus*, que também adota o formato dos *Penn corpora*: http://linguist.is/icelandic_treebank/Icelandic_Parsed_Historical_Corpus_%28lcePaHC%29

¹⁵ Ao contrário das outras categorias, os verbos não projetam, não havendo nenhum nível VP.

¹⁶ Sobre a aplicação do conjunto das etiquetas e subetiquetas originais a um *corpus* do português, veja-se Carrilho e Magro (2011).

Etiquetas Sintagmáticas	
NP-SBJ	Sintagma Nominal (sujeito)
NP-ACC	Sintagma Nominal (acusativo)
NP-DAT	Pronome dativo
PP-SBJ	Sintagma Preposicional (sujeito)
PP	Sintagma Preposicional
VB	Verbo infinitivo
VB-F	Verbo infinitivo flexionado
VB-P	Verbo finito (no Presente)
Etiquetas Oracionais / Frásicas	
IP-MAT	IP independente ou coordenado
IP-INF	Oração infinitiva
Categorias vazias	
pro	sujeito nulo
arb	sujeito arbitrário
*	vestígio deixado pela subida do clítico

Tabela 1. Etiquetas relevantes na anotação das construções perceptivas

Como já se referiu anteriormente, o que há de particular na anotação das construções perceptivas é a forma como o sujeito é codificado¹⁷. Assim, o sujeito pode ser anotado:

i) como **arb** (de arbitrário), dentro de um NP-SBJ, nas construções em que o sujeito da oração infinitiva não está expresso, como em (17):

(17) E um cavaleiro lhe disse que ouvira $_{IP-INF} [\text{ }_{NP-SBJ} \text{ } *arb* \text{ } \text{dezer que forão cento e cincoenta. (JAR96,,28)}$

ii) como *NP-SBJ* (não nulo), nas construções em que o sujeito da oração infinitiva é um sintagma nominal, quer em posição pré-verbal, como em (18), quer em posição pós-verbal, como em (19):

(18) “Filha, disse ele, nom ouviste tu $_{IP-INF} [\text{ }_{NP-SBJ} \text{ } \text{os anjos cantar com sua alma, (JAR52,,41)}$

(19) De como el-rey Arfasão jouve em @o paço onde ouviu $_{IP-INF} [\text{ }_{NP-SBJ} \text{ } \text{cantar os anjos, (JAR116,,3)}$

iii) como *PP-SBJ*, nas construções em que o sujeito da infinitiva é um sintagma preposicional, como em (20):

(20) Mas Abel, que pouco dormia, acordou asinha e vio $_{IP-INF} [\text{ }_{PP-SBJ} \text{ } \text{a Caim. (JAR74,,9)}$

iv) como vestígio de um clítico coindexado com um pronome (acusativo em (21), ou dativo em (22)), que se encontra realizado no domínio superior:

(21) – Certo não, disse ele, mas eu $_{NP-ACC} \text{ } a \text{ } _{IP-INF} [\text{ }_{NP-SBJ} \text{ } \text{*} \text{ } \text{falar com o mestre de@ @os cristãos (JAR48,,15)}$

(22) Quando el-rey, que bem vio os golpes que dava, $_{NP-DAT} \text{ } \text{lhe}_1 \text{ } _{IP-INF} [\text{ }_{NP-SBJ} \text{ } \text{*} \text{ } \text{alçar a espada, nom teve atrevimento de esperar o golpe, (JAR98,,22)}$

Na secção seguinte, demonstra-se como se pode efetuar a pesquisa destas construções num *corpus* com este tipo de anotação sintática, como é o caso dos textos do projeto WOChWEL.

¹⁷ O mesmo tipo de anotação é aplicado às construções causativas, embora aqui se trate apenas as construções perceptivas.

5. PESQUISA DAS CONSTRUÇÕES PERCETIVAS

Nesta secção pretende-se, de um modo geral, apresentar a forma como a pesquisa destas construções se processa. As pesquisas são feitas no texto medieval¹⁸ recorrendo ao programa *CorpusSearch*, um programa para pesquisa e construção de *corpora* linguísticos no formato do *PennTreebank*.

Na Tabela 2 apresentam-se as funções de busca necessárias para procurar as construções percetivas que seleccionam uma oração não finita¹⁹.

função de busca	descrição
iDoms	x domina imediatamente y
precedes	x precede y
HasSister	x e y são imediatamente dominados pelo mesmo nó

Tabela 2. Funções de busca do *CorpusSearch* relevantes na pesquisa das construções percetivas

Tendo em conta a forma do sujeito da oração infinitiva, um ficheiro de *query* como o que se apresenta em (23) encontra grande parte das construções percetivas do texto.

(23)

```
query: (IP* idoms IP-INF)
      AND (IP-INF idoms {1}NP-SBJ|PP-SBJ)
      AND ({1}NP-SBJ|PP-SBJ idoms !\*pro\*)
      AND (IP-INF idoms VB)
```

Ou seja, este ficheiro vai procurar:

- um domínio infinitivo abaixo de IP (onde se incluem todos os casos da construção de infinitivo flexionado, ECM e União de Orações), condição essa que é expressa na primeira linha ((IP* idoms IP-INF)). Ficam excluídos todos os casos de infinitivo preposicionado, os quais são dominados por PP (e não por IP);
- um domínio infinitivo que domine um sujeito que é um sintagma nominal ou preposicional (AND (IP-INF idoms {1}NP-SBJ|PP-SBJ))²⁰ e não pode ser nulo (AND ({1}NP-SBJ|PP-SBJ idoms !*pro*))²¹;
- um domínio infinitivo que domine um verbo infinitivo sem flexão (AND (IP-INF idoms VB)).

No entanto, se se quiser filtrar ou separar os resultados tendo em conta a posição do sujeito da infinitiva, pode refinar-se o ficheiro de *query* e transformar-se (23) em (23a) e (23b):

(23a)

```
query: (IP* idoms IP-INF)
      AND (IP-INF idoms {1}NP-SBJ|PP-SBJ)
      AND ({1}NP-SBJ|PP-SBJ idoms !\*pro\*)
      AND (IP-INF idoms {2}VB)
      AND ({2}VB precedes {1}NP-SBJ|PP-SBJ)
```

¹⁸ Cf. Martins *et al.* (2013-2015).

¹⁹ Para mais informações sobre as funções de busca, consultar a página do programa *CorpusSearch* (<http://corpussearch.sourceforge.net>) ou Carrilho e Magro (2010) para a sua aplicação a um *corpus* do português.

²⁰ Dito de outra forma, esta condição vai procurar todos os casos de sujeitos, nominais ou preposicionais (NP-SBJ|PP-SBJ), que não dominem sujeitos nulos *pro* mas dominem qualquer tipo de categoria, como é o caso de: i) sujeito nulo que é *arb*, como em (17); ii) sintagma nominal, como em (18)-(19); iii) sintagma preposicional, como em (20) e iv) vestígio de um clítico que subiu (*), como em (21)-(22).

²¹ O sinal "!" é um operador lógico que exclui o constituinte ao qual se refere.

(23b)

```

query: (IP* idoms IP-INF)
      AND (IP-INF idoms {1}NP-SBJ|PP-SBJ)
      AND ({1}NP-SBJ|PP-SBJ idoms !\*pro\*)
      AND (IP-INF idoms {2}VB)
      AND ({1}NP-SBJ|PP-SBJ precedes {2}VB)

```

Em (23a) e (23b) foi introduzida uma condição sobre a ordem dos constituintes. Assim, como resultado de (23a), haverá apenas frases em que o verbo infinitivo precede o sujeito da oração infinitiva e o resultado de (23b) inclui exclusivamente frases em que o sujeito da oração infinitiva aparece antes do verbo infinitivo.

No caso das construções em que o infinitivo pode apresentar flexão, a pesquisa tem de ser ligeiramente diferente. Assim, para pesquisar casos de PIC, em que, além de o infinitivo ser flexionado, é introduzido pela preposição *a*, um ficheiro de *query* possível é o que se apresenta em (24):

(24)

```

query: (PP idoms IP-INF)
      AND (IP-INF idoms {1}NP-SBJ)
      AND ({1}NP-SBJ idoms !\*pro\*)
      AND (IP-INF idoms VB|VB-F)

```

Este ficheiro pesquisa orações infinitivas introduzidas por preposição²² ((PP idoms IP-INF)) e que contenham um sujeito que não pode ser nulo, prevendo a possibilidade de o infinitivo poder ou não ser flexionado (AND (IP-INF idoms VB|VB-F))²³.

Os casos de construção de infinitivo flexionado, por não serem introduzidos por preposição, exigem uma busca diferente. Em (25) apresenta-se uma pesquisa possível para estas construções:

(25)

```

query: ({1}IP-INF idoms NP-SBJ)
      AND ({1}IP-INF idoms VB-F)
      AND ({1}IP-INF HasSister VB-*)

```

Tendo em conta os ficheiros de *query* apresentados, na secção seguinte mostram-se os resultados obtidos.

6. RESULTADOS

O Livro de José de Arimateia apresenta, de acordo com a anotação sintática explicitada nas secções anteriores, 183 construções não finitas introduzidas por verbos perceptivos: 36 com *ouvir*, 4 com *sentir* e 143 com *ver*.

Quanto à forma do sujeito da oração infinitiva, a Tabela 3 resume as diferentes formas que ele pode assumir neste texto.

²² Pode criar-se uma nova linha especificando qual a preposição que introduz o infinitivo para que os resultados sejam mais específicos.

²³ A barra vertical "|" é um operador lógico que significa alternativa (= ou).

forma do sujeito da oração infinitiva	nº de ocorrências
sintagma nominal	90
pronome acusativo	48
nulo	24
pronome dativo	7
sintagma preposicional	6
oração relativa	5
pronome vos ¹	2
pronome se	1

¹ As frases com o pronome *vos* foram contabilizadas separadamente dos pronomes acusativos e dativos, uma vez que este pronome apresenta forma única para o caso acusativo e dativo.

Tabela 3. Forma do sujeito da oração infinitiva e nº de ocorrências

Como a Tabela 3 demonstra, a forma mais comum do sujeito encaixado é um sintagma nominal, que pode ocorrer antes (como em (26)) ou depois do verbo infinitivo (como em (27)), ou um sintagma preposicional (como em (28)), que ocorre sempre depois do complexo verbal. Quando assume a forma de um pronome, aparece sempre no domínio do verbo finito. É mais frequentemente um pronome acusativo (cf. (29)) mas também pode ser, por vezes, um pronome dativo (cf. (30) e (31)). O sujeito da oração infinitiva pode não estar expresso e ser nulo, como em (32).

(26) Ali poderiades ver **os homens** matar e derribar outros. (JAR101.,22)

(27) quando viram sair de ùa montanha **Nasciam e o duque Gaanor**, com muyta e fermosa companhia. (JAR99.,136)

(28) Mas, quando Tolomer viu chegar **a el-rei Evalac**, bem cuidou que nom teria vagar de lhe cortar a cabeça (JAR49.,72)

(29) Quando **o** ela ouviu falar tam atrevidamente, foy mays leda que de@ @antes (JAR111.,18)

(30) E, quando os judeus **lhe** vírão fazer taes maravilhas, houvérão@ @lhe grande enveja (JAR32.,24)

(31) – Pois eu **lhe@ @o** vi fazer, disse o cavaleyro, assi como vos eu digo. (JAR83.,103)

(32) – Muitas vezes ouvimos nós falar de Hipocras. (JAR81.,15)

Relativamente à forma e posição do sujeito da oração infinitiva, os dados mostram que, tal como em PEC dialetal (cf. Pereira (2012) e exemplos (7)-(9) acima), ele pode estar expresso sob a forma de um pronome dativo. Em (30) e (31) temos um verbo infinitivo que é transitivo e o sujeito é um pronome dativo que sobe para o domínio superior. Note-se que em (31), juntamente com o sujeito, sobe também o complemento direto do verbo infinitivo. Em (29), o pronome sobe para uma posição antes do sujeito, constituindo um caso de interpolação (do sujeito).

Não sendo sempre possível identificar a construção envolvida, a Tabela 4 mostra a distribuição das várias construções presentes no texto medieval, tendo em conta os diferentes verbos.

Tipo de construção	ouvir	sentir	ver	nº total de ocorrências
construção de União de Orações	7		71	78
construção de ECM	4	3	17	24
casos ambíguos entre ECM e União de Orações	6	1	42	49
outros casos	19		13	32
	36	4	143	183

Tabela 4. Distribuição do verbo por tipo de construção percetiva

A construção com verbos percetivos mais frequente neste texto é a construção de União de Orações, com 78 ocorrências não ambíguas, sendo a construção mais frequente com o verbo *ouvir* e com o verbo *ver*. Há 49 casos de ambiguidade entre a construção de ECM e a construção de União de Orações. Há ainda 24 frases com a construção de ECM, construção esta que é a mais

comum com o verbo *sentir*. Não há exemplos inequívocos da construção de infinitivo flexionado. Também não foram encontrados exemplos da construção infinitiva gerundiva nem da construção com infinitivo flexionado e sujeito acusativo²⁴. Relativamente às frases cujas construções não são identificáveis, há 32 casos em todo o texto: 24 por o sujeito da oração infinitiva não ocorrer (como em (32) acima); em cinco casos o referente desse sujeito é um sintagma nominal que é o antecedente de uma oração relativa (cf. (33)); duas construções estão envolvidas em casos de coordenação (cf. (34)); e há um caso em o que sujeito da oração infinitiva é o pronome *se* (cf. (35)).

(33) e logo me pareceo que aquela era **a fegura** que vira sair de@ @a capela e logo disse que aquilo creeria eu como quem o vira mui bem. (JAR51,,56)

(34) Pois, quando Ebalac vio **seu cunhado** tão esforçado e tão ardido e **cometer** tão esquivamente por seu amor de@ @ele e por livrar sua terra, vendo sua gram bondade, começou a chorar (JAR49,,23)

(35) Quando Celidonis **se** vio meter com aquela besta tam grande e tam dessemelhada, fez sobre si o sinal de@ @a cruz e encomendou@ @se a Nosso Senhor. (JAR78,,80)

Considerem-se agora os seguintes exemplos que ilustram as diferentes construções perceptivas identificadas no texto.

(36) Depois vio Joseph vir **a Nosso Senhor** em outra tal semelhança como lhe ele aparecera em@ @o cacere, (JAR41,,16)

(37) Depois bebeo ûu pouco de sangue e, depois que isto houve feito, vio estar **o anjo**, que tomou a patena e cobrio o caliz e ergueo@ @a em alto. (JAR43,,15)

(38) E, antes que as azes fossem partidas como haviam de estar, viram **os cristãos** sair todos armados de ùa montanha, os elmos todos enlazados. (JAR101,,16)

(39) quando se assentaram ambos em ùa camara, sentiram **o paço** só si tremer. (JAR59,,5)

(40) E, quando Josefes **o** vio vir, disse@ @lhe: – Moïs, nom te assentes aqui se nom es tal como deves, que tu te acharás mal. (JAR104,,7)

(41) e hoje nom podedes vós vir de longe, se nom andais mais de pressa de@ @o que **vos** eu vi vir. (JAR75,,38)

Em (36) e (37) apresentam-se exemplos da construção de União de Orações com um sintagma preposicional e um sintagma nominal, respetivamente, no domínio infinitivo (para além dos casos em (30) e (31) acima com pronomes dativos). Ambas as frases em (38) e (39) são aqui consideradas casos de ECM: o exemplo (38), estando o sintagma nominal no plural e não exibindo o infinitivo flexão, não levanta nenhuma dúvida; o exemplo (39), em PEC, seria um caso ambíguo entre construção de ECM e construção de infinitivo flexionado, uma vez que o sintagma nominal está no singular e a terceira pessoa do singular não exhibe flexão no infinitivo, ainda que possa tratar-se de infinitivo flexionado. No entanto, não existindo, neste período, evidência para a existência desta construção, os casos em que o sintagma nominal ocorre entre os dois verbos e é singular só podem ser casos de ECM²⁵. A frase (40) representa um caso ambíguo entre a construção de ECM e a construção de União de Orações, uma vez que o sujeito da oração infinitiva é um pronome acusativo, o que acontece em casos de ECM e em casos de União de Orações em que o verbo infinitivo seja monoargumental (como *vir*). O mesmo acontece com o exemplo (41), que apresenta ainda um outro tipo de ambiguidade: o facto de o pronome de segunda pessoa do plural, *vos*, apresentar a mesma forma para o caso acusativo e para o caso dativo.

²⁴ Neste texto medieval também não foram atestadas as construções dialetais descritas em Pereira (2012) para o PEC (cf. nota 10 acima).

²⁵ São 11 os casos em que o sintagma nominal é singular, ou seja, os casos em que, em PEC, haveria ambiguidade entre a construção de infinitivo flexionado e a construção de ECM.

Repare-se que, em todos os casos da construção de ECM, não há exemplos de cliticização de pronomes complementos nem da presença do operador de negação frásica no domínio infinitivo, havendo evidência, na linha de Martins (2006), para considerar o domínio infinitivo na construção de ECM mais defetivo em PA do que em PEC.

É de salientar que frases como a que se apresenta em i), em PEC, seriam consideradas casos de União de Orações, na medida em que o domínio infinitivo não legitima clíticos complemento, contrariamente ao que se passa na construção de ECM, e estes têm de subir para o domínio superior. No entanto, não havendo evidência, em PA, para afirmar que a construção ECM legitima clíticos complemento, este exemplo é ambíguo. Por se tratar de um caso em que o sujeito da oração infinitiva não ocorre, esta frase foi categorizada como “outros casos” (cf. Tabela 4).

i) E eu te farei certo de tal cousa de que nhum homem mortal nunca foi certo, e per ti será aberta e conhecida a todos aqueles que **a** ouvirem contar. (JAR03,,27)

Note-se que frases como as que se apresentam em (42) e (43) que, aparentemente, podem parecer ambíguas, foram consideradas instâncias da construção de União de Orações. Em PA como em PEC, estes verbos podem ter uma aceção em que são reflexivos (*mover-se* e *doer-se*), sendo essa a aceção que parece estar envolvida nestes exemplos. Segundo Gonçalves (1999: 321), a construção de ECM e a construção de União de Orações apresentam diferenças relativamente aos pronomes reflexos e anticausativos. Assim, eles podem ocorrer no domínio infinitivo da construção ECM mas não no da construção de União de Orações. Veja-se o contraste entre (42)-(43), por um lado, e (44)-(45): estes são exemplos da construção de ECM, cujo domínio infinitivo legitima a ocorrência do pronome *se* junto ao verbo infinitivo. O facto de este pronome não se encontrar atestado em exemplos como (42)-(43), permite identificá-los como construções de União de Orações. A frase (46), sendo exemplo único no texto, constitui um caso intrigante uma vez que o sujeito da oração infinitiva ocorre depois do complexo verbal, como numa construção típica de União de Orações, mas o verbo infinitivo exhibe o pronome *se*, como na construção de ECM. Na linha de Martins (no prelo), o exemplo foi considerado aqui como construção de ECM, na medida em que, em PA, a ordem dos constituintes era mais livre e não há evidência para a legitimação de clíticos reflexos e anticausativos no domínio infinitivo da construção de União de Orações.

(42) E, depois que lhes isto disse, aprouve@ @lhe com eles e, em esforço de@ @estes, fez estar toda sua companha, aquela que escapara, e disse a@ @os seus cavaleiros que nenhúu nom movesse ata que o vissem a ele *mover*. (JAR48,63)

(43) Entam começaram a chorar todos quantos i estavam, porque o viram *doer* de seu pecado por a vingança que Nosso Senhor tomara. (JAR102,21)

(44) Mas ali assima, sô ùa oliveira, vi **um homem** nom sey se é cavaleiro se pião *combater*@ @se com um fero jayam de terra estranha. (JAR88,33)

(45) E o mensageiro andou tanto que vio **Tolomer** *partir*@ @se. (JAR48,5)

(46) E, depois que assí esteve um pouco, vio em@ @o mar *alevantar*@ @se **úa gram tempestade**, que parecia que a pena se queria derribar. (JAR64,14)

Ainda sobre a construção de União de Orações, importa aqui destacar que, contrariamente ao que acontece em PEC padrão, o verbo finito e o não finito nem sempre ocorrem em posição adjacente. Vejam-se os seguintes exemplos:

(47) Quando Narpos, que ainda nada sabia de@ @esto, ouviu **assi** falar a seu padre, disse: – Sobre esto haveremos logo conselho. (JAR118,34)

(48) Eles aportaram em@ @a insola e ouvirão **dentro** soar ùa bozina. (JAR79,37)

(49) Tu viste **de teu sobrinho** sair um lago e de@ @ele saíam nove rios. (JAR66,87)

(50) Por a manhã, quando se levantou a gente de Roma e saíram de suas cassas, olharam contra a torre e viram **em@ @o çesto** estar ùu homem. (JAR83,45)

(51) E vio **entam sobre a penna** estar ùu pão mui negro de cevada e, quando o vio, foi mui ledado e correndo foi tomar o pam. (JAR65,48)

(52) E ele sobio em@ @o eyrado de@ @a torre e vio **perto de@ @o peytoril, sô as ameyas**, jazer ùa corda grande e forte e longa, e o cabo de@ @ela atado em@ @as ameyas. (JAR83,7)

Os exemplos (47) e (48) mostram a possibilidade de um advérbio ocorrer entre o verbo finito e o não finito. Em (49) e (50) estão casos em que complementos do verbo infinitivo ocorrem imediatamente antes dele. Finalmente, em (51) e (52) há mais do que um constituinte (um advérbio / sintagma adverbial e um sintagma preposicional) a interromper a adjacência entre o verbo finito e o verbo infinitivo.

7. CONCLUSÕES

Os dados mostram que, em PA, apenas duas construções são possíveis com verbos percetivos a seleccionar complementos infinitivos: a construção de ECM e a construção de União de Orações, não se atestando neste texto, como é expectável, a construção de infinitivo flexionado. A cons-

trução infinitiva gerundiva e a construção com infinitivo flexionado e sujeito acusativo também não ocorrem. Apesar de haver bastantes frases ambíguas entre a construção de ECM e a construção de União de Orações, há uma preferência clara por esta última, tal como refere Martins (2006). Tal tendência é também seguida pela generalidade dos dialetos do PEC, com exceção da Madeira onde a construção não se encontra atestada, como se mostra em Pereira (2012), mas não pelo PEC padrão.

Uma outra ligação que se pode estabelecer entre os dados dialetais do PEC e estes dados do PA é o facto de não existirem restrições sobre o verbo infinitivo da construção de União de Orações. Os exemplos (30) e (31) acima mostram que a construção é possível com verbos transitivos e o mesmo tinha já sido demonstrado para os dialetos do PEC por Pereira (2012) (cf. exemplos (7)-(10) acima).

Dado que neste texto medieval não se atestam exemplos que mostrem que, em PA, na construção de ECM era possível a ocorrência do operador de negação frásica ou da cliticização de clíticos complementos, assume-se, com Martins (2006), que o domínio infinitivo nestas construções é mais defetivo em PA do que em PEC.

Com este trabalho, mostrou-se igualmente a utilidade de *corpora* sintaticamente anotados para o estudo de construções específicas. Dando continuidade a este trabalho, uma tarefa futura será olhar para estas construções em outros períodos da língua em *corpora* com o mesmo tipo de anotação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bikel, Dan (2004): *On the Parameter Space of Generative Lexicalized Statistical Parsing Models*. Diss. PhD. University of Pennsylvania.
- Bossaglia, Giulia (2013): "Inflected /Non-inflected Infinitive Alternation in Causative and Perception Constructions of Contemporary European Portuguese: A Corpus-based Study", *Procedia – Social and Behavioral Sciences* 95, 220-230. <http://dx.doi.org/10.1016/j.sbspro.2013.10.642>
- Carrilho, Ernestina / Catarina Magro (2010): "A anotação Sintática do CORDIAL-SIN", in Ana Maria Brito et al. (orgs.), *XXV Encontro Nacional da Associação Portuguesa de Linguística – Textos Seleccionados*. Porto: Associação Portuguesa de Linguística, 225-241.
- Carrilho, Ernestina / Sandra Pereira (2010): "Causees in European Portuguese dialects: some observations on the properties and the position of the causee in causative constructions in CORDIAL-SIN", comunicação apresentada no *Wedsyn's First Workshop on Syntactic Variation*, IKER, Bayonne, Março 2010.
- Castro, Ivo (1984): *Livro de José de Arimateia* (Estudo e Edição do COD. ANTT 643). Dissertação de Doutoramento. Universidade Lisboa.
- Duarte, Inês / Anabela Gonçalves (2002): "Construções de subordinação funcionalmente defectivas: o caso das construções perceptivas em Português europeu e em Português brasileiro", in Anabela Gonçalves / Clara Nunes Correia (eds), *Actas do XVII Encontro da Associação Portuguesa de Linguística*. Lisboa: APL, 161-173.
- Gonçalves, Anabela (1999): *Predicados Complexos Verbais em contextos de infinitivo preposicionado do português europeu*. Dissertação de doutoramento. Universidade de Lisboa.
- Gonçalves, Anabela / Inês Duarte (2001): "Predicados complexos com verbos causativos e perceptivos do Português europeu", in Clara Nunes Correia / Anabela Gonçalves (eds), *Actas do XVI Encontro da Associação Portuguesa de Linguística*. Lisboa: APL, 227-239.
- Guasti, Maria Teresa (2006): "Analytic Causatives", em Martin Everaert / Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, vol. 1. Oxford: Blackwell, 142-172. <http://dx.doi.org/10.1002/9780470996591.ch6>
- Kepler, Fábio (2005): *Um etiquetador morfo-sintático baseado em cadeias de Markov de tamanho variável*. Dissertação de Mestrado. Instituto de Matemática e Estatística da Universidade de São Paulo.
- Martins, Ana Maria (2006): "Aspects of infinitival constructions in the history of Portuguese", in Randall S. Gess / Deborah Arteaga (eds.), *Historical Romance Linguistics: Retrospective and Perspectives*. Amsterdam & Philadelphia: John Benjamins, 327-355. <http://dx.doi.org/10.1075/cilt.274.20mar>
- Martins, Ana Maria (no prelo): "Infinitival complements of causative/perception verbs in a dia-

- chronic prespective”, in Anabela Gonçalves / Ana Lúcia Santos (eds.) *Complement Clauses in Portuguese: Syntax and Acquisition*, Amsterdam/Philadelphia: John Benjamins.
- Martins, Ana Maria / Sandra Pereira (2012–2015): POS-tagged José de Arimateia. CC licensed: WOChWEL by Centro de Linguística da Universidade de Lisboa [last update 14-11-2014].
- Martins, Ana Maria / Sandra Pereira / Adriana Cardoso (2013-2015): Parsed José de Arimateia. CC licensed: WOChWEL by Centro de Linguística da Universidade de Lisboa [last update 23-07-2014].
- Martins, Ana Maria / Sandra Pereira (2014-2015): POS-tagged Demanda do Santo Graal (chapters 01-300). CC licensed: WOChWEL by Centro de Linguística da Universidade de Lisboa [last update 19-12-2014].
- Martins, Ana Maria / Sandra Pereira / Adriana Cardoso (2014-2015): Parsed Demanda do Santo Graal (chapters 01-200). CC licensed: WOChWEL by Centro de Linguística da Universidade de Lisboa [last update 22-12-2014].
- Pereira, Sandra (2012): *Protótipo de um glossário dos dialetos portugueses com informação sintática*. Dissertação de doutoramento. Universidade de Lisboa.
- Pereira, Sandra (2014): “A anotação das construções causativas em textos medievais”, comunicação apresentada no CODIL13 -Terceiro Colóquio Internacional sobre corpora diacrônicos de línguas ibero-românicas, Universidade de Zurique, 23-25 de junho.
- Raposo, Eduardo Paiva (1981): *A Construção União de Orações na gramática do português*. Dissertação de doutoramento. Universidade de Lisboa.
- Raposo, Eduardo (1989): “Prepositional Infinitival Constructions in European Portuguese”, in O. Jaeggli / K. Safir (eds.), *The Null Subject Parameter*. Kluwer Academic Publishers, 277-305. http://dx.doi.org/10.1007/978-94-009-2540-3_10
- Silva, Augusto Soares da (2005): “Revisitando as construções causativas e perceptivas do Português: significado e uso”, in Inês Duarte / Isabel Leiria (orgs.), *Actas do XX Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: Associação Portuguesa de Linguística, 855-874.
- Silva, Augusto Soares da (2012): “Stages of grammaticalization of causative verbs and constructions in Portuguese, Spanish, French and Italian”, *Folia Linguistica* 46: 2, 513-552. <http://dx.doi.org/10.1515/flin.2012.018>
- Toledo Neto, Sílvio (2012-2015): *Edição da Demanda do Santo Graal*.