

A DISTRIBUIÇÃO DE *CONHECER* NUM *CORPUS* DO PORTUGUÊS MEDIEVAL

CARLOS ROCHA

(Praxis XXI/Universidade de Oxford)

Introdução

A categorização da actividade mental em Português encontra-se repartida por um conjunto de itens e expressões, do qual faz parte o lexema *conhecer*. Numa perspectiva diacrónica e diatópica, verifica-se que este verbo reforça a sua afinidade com outros verbos de actividade mental, não só a nível da significação, mas também no respeitante às suas estruturas de complementação. Com efeito, nota-se no português medieval e em certos usos individuais e regionais uma maior variedade estrutural quanto aos complementos de *conhecer*, sobretudo no que se refere à subcategorização de um complemento oracional.¹ Neste tipo de ocorrências, estão marcados valores semânticos que se parecem ou se identificam com as significações de *perceber*, *compreender*, *aperceber-se* e *reconhecer*. Inclua-se ainda neste grupo o verbo *saber*, que se mantém em "sinonímia lexicológica" com *conhecer* (Vilela 1994: 101).

Nesta comunicação, pretendo, por um lado, dar conta da distribuição de *conhecer* num *corpus* de textos medievais portugueses e, por outro, detectar alguns factores de variação e mudança semânticas nessa distribuição. A perspectiva de análise baseia-se nas metodologias da Linguística de *corpus*, assumindo, com Sinclair (1991: 53-65), que sentido e estrutura estão de tal maneira associados que entre eles há constantes e contínuos ajustamentos. Estas metodologias permitem dar base empírica a gramáticas de carácter indutivo, especialmente úteis quando se trata de estudar sincronias do passado². Na interpretação dos dados, também recorro a propostas da teoria dos protótipos e da teoria dos campos semânticos.

1. Um *corpus* do Português medieval

O *corpus* utilizado para este trabalho é constituído por 1,5 milhões de palavras (formas ou *tokens*) repartidas por 26 edições de textos e colecções de textos medievais, de diferentes géneros e tipos textuais. Para a criação do *corpus* foram utilizadas três bases de textos, o *Archive of Old Portuguese Texts* (AOPT), o *Corpus Informatizado do Português Medieval* (CIPM), e a *Biblioteca Virtual de Autores Portugueses* (BVAP) (ver Fontes). Os textos que não se encontravam nestes arquivos foram digitalizados propositadamente para esta investigação.

O *corpus* estende-se por cerca de trezentos anos, incluindo textos que vão de 1248 (reinado de D.Afonso III) a cerca de 1520. As amostras de cada tipo ou género textual podem não ser representativas de todo este período, visto haver descontinuidades de produção e documentação. Noutras situações, dado que a digitalização de textos exige certa atenção, não houve tempo para acrescentar ao *corpus* amostras mais extensas, como foi o caso com os textos legais, que não puderam incluir as cartas posteriores a 1433.

O quadro 1 indica as dimensões e organização do *corpus*. Como se pode verificar, a criação do *corpus* teve em conta a variação tipológica textual. Para isso, defini duas classes fundamentais: prosa não narrativa (classe I) e prosa narrativa (classe II). Dentro da classe I, houve que separar os textos legais como uma subclasse (Ia), dado apresentarem características bem individualizadas.

Quadro 1- Distribuição das palavras (*tokens*) do *corpus* (ver Fontes)

	periodo	total de palavras	% sobre subtotal	% sobre total
Prosa não narrativa- classe I	s. XIV	23287	10%	
	s. XV	202256	90%	
subtotal		222543		12%
Textos legais- classe Ia	1248-1325	116526	20%	
	1325-1433	452182	80%	
subtotal		568708		30%
Prosa narrativa- classe II	s. XIV	523682	48%	
	s. XV	421350	38%	
	s. XVI	148269	14%	
subtotal		1093301		58%
TOTAL DO CORPUS		1884552		100%

A distinção entre prosa narrativa e não narrativa funda-se num certo número de expectativas quanto a características funcionais e formais dos textos. Espera-se que um texto não narrativo tenha mais marcada a presença do sujeito ou dos interlocutores do discurso, como se pode constatar em muita da prosa de edificação religiosa (cf. Cintra, 1957, Cepeda, 1962,) ou nas cartas notariais (cf.

Duarte 1986, Marques, 1984 e 1991, Marques et al. 1982, Marques e Dias, 1990 e Marques e Dias 1986). Supõe-se que a prosa narrativa é mais "objectiva", isto é, menos comprometida com a subjectividade no discurso. Esta distinção é uma questão de grau, pois que se verifica uma certa interpenetração das marcas gramaticais e discursivas de uma e outra classe num mesmo texto. É assim que os diálogos em textos narrativos configuram subunidades estruturais com uma tipologia própria, por inscreverem a presença das relações de intersubjectividade nos textos narrativos. De modo inverso, o discurso abstracto, de carácter teológico ou moral, embora seja aqui classificado como prosa não narrativa, pode apresentar uma redução das marcas de primeira e segunda pessoas. O critério de narratividade é, pois, aqui assumido provisoriamente, necessitando de uma base teórica e empírica mais forte.³

Quanto à classe Ia, ela resulta de se observarem características muito específicas em textos notariais e jurídicos. Tais características prendem-se, por um lado, com o grau de idiomatização de cada texto, aspecto que se torna mais perceptível quando se contrastam as fórmulas notariais, que são fixas, com sequências não formulaicas.⁴ Por outro lado, trata-se de textos com uma terminologia própria, a qual pode reflectir-se na emergência de estruturas específicas, como é o caso do uso de *conhecer de* (ver 2.1.1.).

Alguns tipos de texto foram excluídos do *corpus*. São eles os textos versificados, que abrangem a poesia trovadoresca, a poesia do Cancioneiro de Resende, até aos autos de Gil Vicente. Esta exclusão deve-se principalmente ao facto de, na produção de texto em verso, haver um longo hiato entre o fim da produção trovadoresca galaico-portuguesa e o começo da poesia palaciana coligida no Cancioneiro de Garcia de Resende. Além disso, os contextos versificados requerem uma análise mais minuciosa do que aqui é possível. Como se sabe, os subsistemas métrico e rimático podem interferir no comportamento sintáctico e semântico dos itens lexicais.⁵

Refira-se ainda que os textos electrónicos possibilitaram a aplicação de programas de concordâncias, as quais, por sua vez, facilitaram a criação de uma base de dados. Obtiveram-se assim 1437 registos contendo formas de *conhecer*, os quais foram classificados segundo as seguintes variáveis: localização na edição, localização no documento, contexto, classe textual, estrutura morfológica, estrutura sintáctica e grau de idiomatização.

2. Usos de *conhecer* dos séculos XIII a XVI

Para este trabalho, limitei-me à identificação e inventariação da subcategorização dos complementos directos de *conhecer*, como uma das etapas metodológicas para aceder ao conjunto das suas acepções. Por complemento directo, entendi um complemento seleccionado pelo verbo, seja ele nominal, oracional ou construído por regência preposicional. As acepções que se prendem

com a descrição semântica do sujeito de *conhecer* não foram por enquanto consideradas.

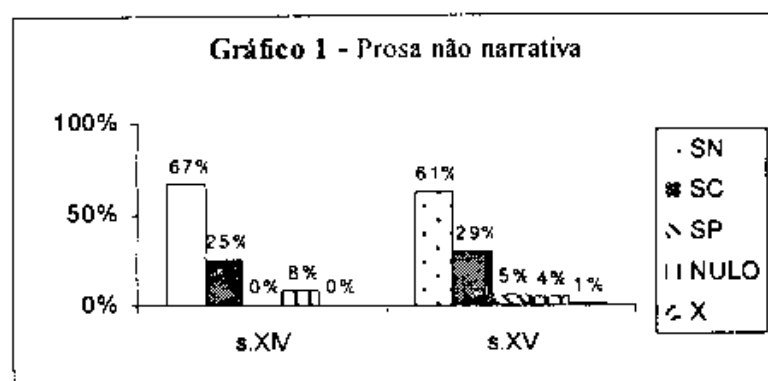
Identifiquei cinco classes de complementação: SN, para o complemento nominal; SC, para o complemento oracional; SP, para o complemento preposicionado; 0 para o constituinte zero ou para o emprego intransitivo do verbo; e X para a oração pequena.⁶ Quanto à distribuição diacrónica, as classes temporais são definidas por século, excepto na secção 2.1.1., em que, por uma questão de obter amostras quantitativamente significativas, optei por dois intervalos temporais, o primeiro, de 1248 a 1325, e o segundo, de 1325 a 1433.

2.1. Prosa não narrativa (classe I)

A distribuição das frequências absolutas de *conhecer* nesta classe encontra-se representada no quadro 2. O gráfico 1, baseado em percentagens, visa facilitar a leitura desta distribuição. Os dados foram extraídos das seguintes edições: Carstens-Grockenberger 1961, Cepeda 1962, Cintra 1957, Dias 1982, Mota et al. 1965, Nascimento 1989, Neto 1959-1960 e Piel 1989. Optei por uma certa flexibilidade na definição dos intervalos temporais, de modo a compatibilizarem-se com a datação de cada texto, a qual é, muitas vezes, feita por estimativa, sem um ano ou período de tempo claramente identificados.

Quadro 2 - frequências absolutas de *conhecer* na não prosa narrativa (classe I)

datação	SN	SC	SP	0	X	total
s.XIV	8	3	0	1	0	12
s.XV	177	83	15	12	2	289



A prosa não narrativa parece mostrar que, do séc. XIV para o séc. XV, houve um alargamento da complexidade estrutural de *conhecer*. Com efeito, passa-se de uma situação em que *conhecer* é empregue com SN, SC e 0, para outra situação em que todas as classes estruturais estão presentes. É preciso, no entanto, ter presente que o facto de nem todas as classes estarem representadas

no primeiro intervalo temporal pode ser consequência do tamanho da amostra. Esta condicionante tem, pois, de ser considerada na avaliação do contraste entre os totais para cada século (quadro 2): 289 do séc. XV contra apenas 12 do séc. XIV.

De todas as formas, assinala-se que *conhecer*+SC é, ao contrário do que parece verificar-se em Português contemporâneo, a segunda estrutura mais frequente nesta classe. Vê-se mesmo um ligeiro aumento percentual de frequência de SC no segundo intervalo temporal (de 25%, no séc. XIV, para 29%, no séc. XV, cf. gráfico 1). Esta observação, porém, pode apenas decorrer da distorção causada pelo já referido contraste entre os efectivos das amostras para cada período.

2.1.1. Textos legais (classe Ia)

A distribuição de *conhecer* em textos notariais e jurídicos é apresentada pelos quadros 3 e 4. As edições utilizadas foram, a saber: Duarte 1986, Coelho 1989, Ferreira 1987, Garvão 1990, Marques 1984, Marques 1991, Marques et al. 1982, Marques e Dias 1986, Marques e Dias 1990 e Rodrigues 1992.

Conhecer tem um comportamento particular nas fórmulas mais ou menos fixas dos textos legais. Estas fórmulas têm um funcionamento e temporalidade próprias, dada a sua reduzida composicionalidade, permitindo a conservação de arcaísmos. Pude identificar 55 ocorrências de *conhecer* em contextos fixos de dois tipos. Um corresponde a *notificationes* (parte introdutória de cartas notariais) do tipo apresentado em (1) e (2):

- (1) conhoscam todos aqueles que esta carta virem e ouvirem que nos...
(Duarte 1980: 31)
- (2) Conhoscam quantos esta carta viren que nos (...) mandamos ...
(Coelho 1989: 21)

Outra sequência é a do mesmo tipo em (3):

- (3) conhosco & cõffesso Que deuo dar uos (...) Trijnta libras de Portugal
(AOPT, cf. Parkinson 1983)

Há ainda a registar 2 ocorrências da expressão *conhecendo por verdade* (Marques 1984). É preciso notar que estas expressões podem não ser as únicas presentes no texto. O carácter formulaico do texto notarial leva a que, frequentemente, muitas sequências se assemelhem de texto para texto.

As *notificationes* incluem instâncias de *conhecer* + SC num período compreendido entre 1248 e 1325. No período seguinte, há apenas 2 ocorrências de *conhecer* + SC, registadas em documentos provenientes da região do Barroso (Marques 1984: 104 e 243). Já em Rocha 1994, sugeri que, nessas fórmulas,

conbecer se encontrava em variação livre com *saber*, em documentos dos reinados de D. Afonso III e D. Dinis, ou seja, entre 1248 e 1325. Após este período, tal distribuição cessa ou é muito reduzida. Quanto à distribuição das fórmulas do tipo em (3), o número de ocorrências, embora muito escasso, distribui-se pelos dois intervalos temporais, 1248-1325 e 1325-1433.

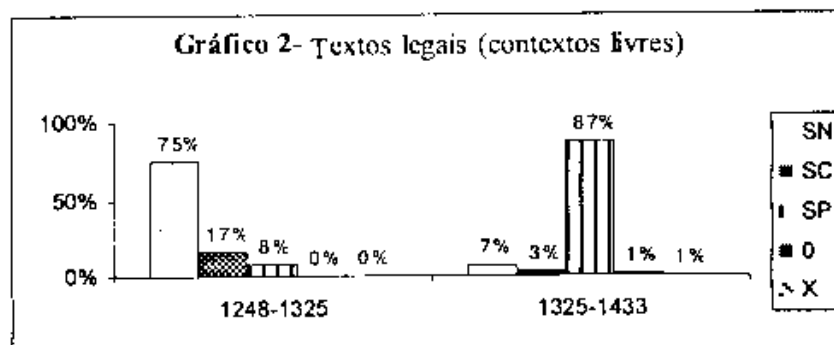
Quadro 3- *Conbecer* + SC em contextos fixos (os tipos indicados referem-se aos exemplos 1-3)

	1248-1325	1325-1433	Total
tipo (1)/(2)	48	2	50
tipo (3)	3	2	5

Outra particularidade da subclasse Ia, é a distribuição de *conbecer*+SP, sendo *de* o núcleo de SP. A distribuição de *conbecer de* é mais perceptível em contextos livres, como se observa no quadro 4 e no gráfico 2.

Quadro 4- Frequências absolutas de *conbecer* em contextos livres

	SN	SC	SP	0	X	total
1248-1325	18	4	2	0	0	24
1325-1433	11	5	131	2	1	150



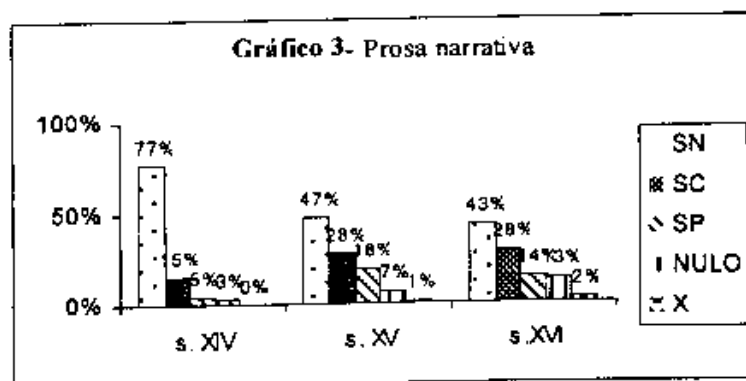
O aumento da frequência de *conbecer* + SP salienta-se no segundo intervalo temporal. Esta distribuição parece ser própria dos contextos livres dos textos notariais e jurídicos, conforme se poderá concluir pelo confronto do gráfico 2 com os outros gráficos presentes neste trabalho. Assim, *conbecer de* surge como um item característico das práticas discursivas do direito. Designa a intervenção dos representantes do poder judicial e deve ser interpretado, confirmando a aceção proposta por Machado 1987, como 'instruir um processo, uma causa'.

2.2. Prosa narrativa (classe II)

Para esta classe, foi possível recolher não só um maior número de textos, mas também dar-lhe representatividade num período de tempo mais extenso, de princípios do séc. XIV ao primeiro quartel do século XVI. Assim, nos nove textos narrativos seleccionados (Brocardo 1994, Castro 1985, Calado 1991, Cepeda 1982-1989, King 1978, Lucas 1988, Macchi 1975, Mattoso 1980, Piel e Nunes 1988)⁷, identificaram-se 907 ocorrências de *conbecer*, com a distribuição temporal que se pode observar no quadro 5. O gráfico 3 permite constatar que há um contraste entre o séc. XIV e os sécs. XV-XVI, quanto às percentagens obtidas por cada tipo de complemento.

Quadro 5 - Frequências absolutas de *conbecer* na prosa narrativa

	SN	SC	SP	0	X	total
s. XIV	371	73	22	15	0	481
s. XV	149	88	58	21	3	319
s. XVI	46	30	15	14	2	107



Nota-se, do primeiro intervalo temporal para os seguintes, uma descida de *conbecer*+SN e uma subida dos outros complementos, sobretudo de SC, que sobe de 15% no séc. XIV para 28% nos sécs. XV-XVI. De qualquer modo, os resultados obtidos confirmam a situação observada na prosa não narrativa (exceptuando os textos legais). Vê-se que a subcategorização predominante é SN, seguida de SC. SP e 0 ficam a alguma distância deste complementos e X, ou seja, a oração pequena, tem uma frequência muito baixa.

3. Direcionalidade e categorização semânticas

Os dados obtidos tendem para uma grande semelhança estrutural de *conbecer* com verbos que, actualmente, lhe são semanticamente mais próximos, entre eles, *saber*.⁸ Contudo, é preciso assinalar a diferença fundamental que separa *conbecer* de *saber*, quer na Idade Média, quer modernamente: no *corpus* em estudo e através da intuição dos usos contemporâneos, não se regista ou não

se admite *conhecer* com um infinitivo não flexionado, isto é, preenchendo a função de semi-auxiliar, que é atribuída a *saber* + infinitivo (p.ex. *saber ler*).

Além disso, a afinidade entre *conhecer* e *saber* não tem sido suficiente para sustentar tendências evolutivas divergentes. No quadro 6, aplica-se a proposta de Traugott 1989, acerca de três tendências (suficientes e não necessárias) no desenvolvimento da polissemia de certos itens verbais, o que permite evidenciar a especificidade de *conhecer*.

A leitura do quadro 6 deixa a impressão de que *conhecer* tem dificuldades de enquadramento nestas tendências. A tendência I não é seguida pela sua forma latina, pela simples razão de que esta já correspondia à conceptualização de uma situação interna.⁹ A tendência II encontra-se reflectida na fórmula injuntiva de documentos notariais, mas não parece ter tido continuidade em termos de uma passagem à tendência III, ou seja, ao alargamento a valores e usos modais. Com efeito, *conheço lá* não tem a mesma função de marcador discursivo de *sei lá* (com valor de hesitação, dúvida). Tampouco se regista a mesma lexicalização de sequências que, no caso de *saber*, conduziram ao arcaico *quicá* ou à forma galega *seica*.¹⁰ A evolução de *conhecer* será, pois, um exemplo de como as relações de sinonímia nem sempre permitem a generalização de determinadas tendências evolutivas.

Quadro 6- Direccionalidade da diacronia de *saber* e *conhecer*
(baseado em Traugott 1989: 34-35)

	SABER	CONHECER
Tendência I: situação externa > situação interna (avaliativa, perceptiva, cognitiva)	SAPERE ('saborear') > <i>saber</i> (port., cast.), <i>savoir</i> (fr.), etc.	COGNOSCERE ('ter conhecimento, conhecer') > <i>conhecer</i> (port.), <i>conocer</i> (cast.), <i>connaître</i> (fr.)
Tendência II: situação externa/ interna > situação textual e metalinguística	(i) S. XIII: "Sabiam todos aqueles que esta carta virem que eu Dom Afonso ..." Duarte 1984 (ii) português contemp.: <i>fica sabendo que...</i>	(i) s. XIII: "conhoscam quantos esta carta virem e leer ouvirem como nos ..." Duarte 1984 (ii) português contemp.: ?? <i>fica conhecendo...</i>
Tendência III: situação textual e metalinguística > atitude do falante em relação à proposição	(i) <i>quicá</i> (arc.), <i>seica</i> (gal.) (ii) <i>sei lá</i>	? <i>conheço lá</i>

Apesar de a sinonímia com *saber* não parecer condicionar a diacronia de *conhecer*, não se poderá dizer que a evolução do lexema foge a qualquer tipo de solidariedade estrutural ou sistêmica. É certo que *conhecer* tem mantido, ao longo dos séculos, uma parte da sua significação numa certa estabilidade. Trata-se dos valores referentes à situação de conhecer alguém ou alguma coisa. Uma das representações mais típicas ou centrais de 'conhecer' será a do relacionamento com outros seres humanos, donde decorrem sentidos como 'travar conhecimento', 'aprofundar o conhecimento de alguém', 'aprofundar a convivência com alguém' (até ao sentido bíblico de *conhecer*) e mesmo 'reconhecer'. Há, no entanto outras acepções e estruturas mais marginais que, como se verá, podem ser encaradas através de processos de indução motivados pelas relações de afinidade estrutural ou semântica de *conhecer* com outros itens lexicais (como sejam *saber*, *reconhecer* ou, contemporaneamente, *perceber*). Tais acepções e estruturas podem eventualmente atingir frequências apreciáveis em determinados períodos.

Estas tendências poderão inscrever-se naquilo que D. Geeraerts (1997: 64-65) classifica como poligênese semântica (*semantic polygenesis*). Geeraerts define este conceito como relativo à existência de sentidos transitórios no desenvolvimento diacrónico das categorias lexicais (1997: 64). O fenómeno consistirá na descontinuidade histórica dos sentidos periféricos, em contraste com a continuidade das interpretações mais salientes de um item lexical (1997: 65).¹¹ Este comportamento inscreve-se no quadro da categorização prototípica, na medida em que convoca um modelo de organização da significação segundo categorias possuidoras de um núcleo estável ("core") e de uma periferia difusa ("fuzzy"). O contraste entre centro e periferia é definido pelo conceito de saliência.¹²

Esta perspectiva poderia explicar melhor, por exemplo, a motivação léxico-semântica da construção *conhecer* + SC. Com efeito, leituras menos salientes ou extensões semânticas ocasionais da significação de *conhecer* poderiam levar à emergência de determinadas estruturas, também elas menos características ou salientes. Esta hipótese exige maior base empírica, com uma maior representatividade do *corpus*, no sentido de incluir uma amostra de textos notariais mais ampla e de assegurar intervalos temporais mais reduzidos (por exemplo de cinquenta anos, em vez dos de cerca de cem, conforme se encontram nos gráficos 1, 2 e 3 nas secções anteriores). Além disso, seria desejável que também pudessem ser abrangidos diferentes tipos de textos até aos dias de hoje, para se apurar se *conhecer* + SC corresponde a um sentido que entretanto se perdeu ou se é a manifestação de uma significação latente.¹³

Esta hipótese exigiria ainda discutir o estatuto de *conhecer* com uma relativa livre (p.ex, em *conhece o que se passa no Ministério*). Lembre-se que,

com *saber* e outros verbos de atitude mental, as relativas podem ser confundidas com interrogativas indirectas (p.ex., *vou saber o que se passa*). Esta ambiguidade também poderia ajudar a perceber melhor a flutuação histórica dos usos de *conhecer* + SC.

O aumento da frequência de *conhecer* + SC, principalmente no século XV, pode ainda ser visto como um efeito da organização histórica da classe ou do campo semântico a que *conhecer* pertence. Segundo a lista proposta por Silva (1989:606), que identifica a classe semântico-sintáctica, dos “verbos de cognição”, *conocer* selecciona quer um SN, quer “F *que*”. Este traço leva-a a incluí-lo na mesma lista que *aparecer* (= ‘parecer’), *atender* (= ‘esperar’), *constituir*, *criar*, *cuidar*, *duvidar*, *entender*, *leer*, *osmar* (= ‘pensar’, ‘imaginar’), *pensar*, *saber*, *reconhecer*, *nembrar-se*, *conselhar* (Silva 1988: 606-609). Do ponto de vista lógico-semântico, esta lista poderá ser reduzida a *saber*, *conhecer*, *reconhecer* e *nembrar-se*, dado estes serem verbos factivos, isto é, verbos que pressupõem a verdade da proposição subjacente ao complemento oracional (ver Vilela 1994:88).

Deste grupo de verbos, salienta-se *reconhecer*, por ser um sinónimo (quase) exacto de *conhecer*, quando selecciona SC. Buscas automáticas no *corpus* aqui em estudo mostraram que, embora *reconhecer* esteja atestado, desde os documentos do reinado de D.Dinis, a sua frequência é sempre baixa em comparação com outros “verbos de cognição”, como *saber* ou *entender*, que também seleccionam uma oração. Nos séculos XIII e XIV, *reconhecer* tem 6 ocorrências no AOPT (documentos do reinado de D.Dinis) e 2 ocorrências na *Vida e Paixões dos Apóstolos*. Como item disponível entre 1248 e 1325, é também atestado por Machado (1987), nas *Cantigas de Santa Maria*, e, como já se disse, por Silva (1989: 610 e 740), que regista ocorrências de *reconhecer*, quer com SN, quer com SC nos *Diálogos de S. Gregório* (manuscrito do século XIV, segundo Silva 1989: 56-60).

No *corpus*, verifiquei ainda que, para os séculos XV e XVI, na classe I (prosa não narrativa), apenas se encontram 4 ocorrências de *reconhecer* no *Livro da Cartuxa* (Dias 1982), do reinado de D.Duarte. Dentro da prosa narrativa (classe II), apenas o *Flos Sanctorum* (Lucas 1988) e a *Crónica de D. Pedro de Meneses* (Brocardo 1994) atestam o lexema com, respectivamente, 2 e 6 ocorrências.

Os dados parecem então indicar que, apesar de *conhecer* e *reconhecer* constituírem opções igualmente válidas para os utentes da língua, era *conhecer* + SC que tinha uso preferencial, sobretudo do séc. XV ao primeiro quartel do séc. XVI. É preciso sublinhar, contudo, que a sinonímia com *reconhecer* apenas recobre parte das acepções (‘reconhecer’, ‘admitir’) de *conhecer* + SC. Há que somar os sentidos de ‘perceber’, ‘compreender’, que, no *corpus*, não parecem estar associados a *reconhecer* mas sim a *conhecer* + SC, sobretudo no séc. XV. De

qualquer modo, se por um lado, o uso de *conhecer* + SC já foi mais alargado, porque a configuração do seu campo semântico ainda não incluía itens como *perceber* ou *compreender*, por outro lado, isso também é concomitante do escasso uso de um sinónimo, ou seja, *reconhecer*. Assim, o aumento do uso de *conhecer* + SC não parece apenas devido a uma lógica de estrutura. Isto é, o uso do lexema não se alargou só por falta de itens para 'perceber', 'compreender'; também beneficiou da baixa frequência de um item alternativo para 'reconhecer', 'admitir'.

Estas observações têm a vantagem de, ao reconstituírem o "traçado" de um campo semântico, considerarem sobreposições de sentido, como as de *conhecer* e *reconhecer*.¹⁴ Assim se captam os processos de construção da significação (ou da categorização semântica), os quais não são discretos nem claramente opositivos, antes deixando margens ou fronteiras indistintas ou pouco nítidas (*fuzzy*), favoráveis à partilha total ou parcial dos mesmos traços ou atributos semânticos.¹⁵ Resta saber como, quando e em que registos passou *reconhecer* + SC a suplantar *conhecer* + SC, sem que o tenha podido erradicar.

Conclusão

A utilização de um *corpus* informatizado visou facilitar a observação da correlação entre, por um lado, a variação sintáctica e semântica de *conhecer* e, por outro, a variação diacrónica e diafásica (em termos de uma tipologia textual). Constatou-se que os contextos (fixos e livres) dos textos legais (sobretudo dos notariais) mostravam uma distribuição característica, com um tipo de complementação que só aí ocorre. É disso exemplo o uso típico de *conhecer de* em contextos livres e da distribuição de *conhecer* + SC nos contextos fixos. A lenta subida de SC na prosa narrativa e na prosa não narrativa contraria a tendência que seria de esperar, no sentido da sua diminuição gradual até à distribuição actual.

Observe-se que a falta de textos legais do século XV no *corpus* faz sentir a necessidade de alargar o seu quantitativo. Só assim haverá forma de determinar se, durante os séculos XV-XVI, a distribuição em contexto livre confirma ou não as mesmas tendências dos textos em prosa narrativa e não-narrativa.

A explicação do aumento de *conhecer* + SC nos séculos XV-XVI requererá certamente várias abordagens, intra e extralinguísticas. A hipótese de poligénese semântica de Geeraerts 1998 e uma teoria dos campos semânticos mais sensível a comportamento prototípicos (semânticos e estruturais) apontam o caminho para o futuro desenvolvimento desta pesquisa.

Fontes

1. Edições

Alguns dos textos a seguir referenciados foram retirados de outros *corpora*. Sempre que tal se verifique, segue-se à referência uma remissão entre parênteses rectos.

- Brocardo, M.T., *Crónica do Conde D.Pedro de Meneses de Gomes Eanes de Zurara*, tese de doutoramento, Lisboa: Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa
- Calado, A.A. 1991, *Estoria de Dom Nuno Alvrez Pereyra: edição crítica da Cronica do Condestabre*, Coimbra: Universidade
- Carstens-Grockenberger, D. 1961. *Buch von den trei Tugenden* [Livro das três virtudes], Münster Westfalen: Aschendorffsche Verlagsbuchhandlung
- Castro, I. (coord.) 1985. *Vidas de Santos de um Manuscrito Alcobacense*, Lisboa: Imprensa Nacional - Casa da Moeda
- Cepeda, I.V. 1962. *A Linguagem da 'Imitação de Cristo' [versão portuguesa de Frei João Álvares]*, Lisboa: Centro de Estudos Filológicos
- 1982-1989. *Vida e Paixões dos Apóstolos: ms. alcobacense 280 da B.N. de Lisboa confrontado com a edição de Lisboa, 1505/ Bernardo de Brihuega*, Lisboa: Instituto de Investigação Científica - Centro de Linguística da Universidade de Lisboa
- Cintra, M.A.V. 1957, *Livro de Solilóquio de Santo Agostinho*, Lisboa: Centro de Estudos Filológicos
- Coelho, M.H.C. 1989, *O Baixo Mondego nos Finais da Idade Média (Apêndice Documental)*, vol. II, Lisboa: Imprensa Nacional - Casa da Moeda
- Dias, J.J.A. 1982. *Livro dos Conselhos de D.Duarte (Livro da Cartuxa)*, Lisboa: Estampa
- Duarte, L.F. 1986. *Documentos em Português da Chancelaria de D.Afonso III*, tese de mestrado, Lisboa: Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa [ver também CIPM]
- Ferreira, J.A. 1987. *Afonso X. Foro Real*, Lisboa: Instituto Nacional de Investigação Científica [ver também CIPM]
- Garvão, M.H. 1990. *Foros de Garvão*, tese de mestrado, Lisboa: Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa [ver também CIPM]
- King, L. 1978. *Crónica do Conde D.Duarte de Meneses / Gomes Eanes de Zurara*, Lisboa: Universidade Nova
- Lucas, C.A. 1988, *Ho Flos Sanctorum em lingoagê: os santos extravagantes*, Lisboa: Instituto Nacional de Investigação Científica
- Macchi, G. 1975. *Crónica de D.Fernando / Fernão Lopes*, Lisboa: Imprensa Nacional - Casa da Moeda [ver também BVAP]

- Marques, A.H.O. 1984. *Chancelarias Portuguesas - Reinado de D.Pedro I*, Lisboa: Instituto Nacional de Investigação Científica
- Marques, A.H.O. 1991, *Chancelarias Portuguesas - D. Afonso IV*, Lisboa, Instituto Nacional de Investigação Científica
- et alii 1982. *Cortes Portuguesas - Reinado de D.Afonso IV (1325-1357)*, Lisboa: Imprensa Nacional - Casa da Moeda
- e Dias, N.J.J.P. 1986. *Cortes Portuguesas - Reinado de D.Pedro I*, Instituto Nacional de Investigação Científica
- e Dias, J.J.A. 1990. *Cortes Portuguesas - Reinado de D. Fernando I (1367-1383)*, Lisboa: Imprensa Nacional - Casa da Moeda
- Mattoso, J. 1980. "Livro de Linhagens do Conde D. Pedro" in *Portugaliae Monumenta Historica. Nova Série*, 2^o vol, Lisboa: Academia das Ciências de Lisboa
- Mota, J.A. et alii 1965. *Livro das Aves*, Rio de Janeiro: Instituto Nacional do Livro
- Nascimento, A.A. 1989. "Regras para enformarmos os menynos en latim", *Euphrosyne*, vol. 17, pp. 209-232
- Neto, S.S. 1959-1960. Regra de S. Bento, *Revista Brasileira de Filologia*, vol. 5, t. I-II, pp. 21-46
- Piel, J.M. 1989. *Livro da Enseñança de Bem Cavalgar Toda Sela*, Lisboa: Imprensa Nacional - Casa da Moeda
- e Nunes, I.F. 1988, *A Demanda do Santo Graal*, Lisboa: Imprensa Nacional - Casa da Moeda
- Rodrigues, M.C.M. 1992. *Costumes de Santarém*, tese de mestrado, Lisboa: Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa [ver também CIPM]

2. Corpora electrónicos

- Archve of Old Portuguese Texts (AOPT)*, Parkinson, S.R., Modern Languages Faculty, University of Oxford (ver nas Referências Parkinson 1983)
- Base de Datos da Lírica Profana Galego-Portuguesa*, Brea., M (coordenadora), Santiago de Compostela: Instituto Ramón Piñeiro (<http://www.cirp.es/CIRPDB/BASIS/meddb>)
- Biblioteca Virtual de Autores Portugueses (BVAP)*, Castro, I., Amado, T., Ribeiro, C.A. e Morão, P. (coordenadores), Lisboa: Biblioteca Nacional
- Corpus Informatizado do Português Medieval (CIPM)*, Departamento de Estudos Linguísticos, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa

Notas

- 1 Em Nascimento et al. (1987: 222), regista-se a seguinte sequência: *conbece o que é um tremço?* O exemplo é problemático na actualidade. Mário Vilela considera que esta

estrutura é “normal” em Português, com o sentido de ‘reconhecer/perceber’ (1994: 97). Em contrapartida, o *Dicionário de Português Básico*, coordenado pelo mesmo investigador (Vilela 1991), não inclui tal estrutura na entrada de *conhecer*. Não fica, pois, claro o estatuto do complemento oracional de *conhecer* no português contemporâneo.

2 A importância de uma gramática indutiva tem sido referida nos estudos do Português medieval: “Para uma sincronia passada da língua, uma gramática mais adequada em uma primeira etapa de conhecimento sistemático do objecto em estudo será (...) uma gramática descritiva, indutiva que opere sobre inventários que se definam como representativos.” Silva 1989: 44.

3 Cabe, a este propósito, referir a análise multi-dimensional de Douglas Biber (1988 e Biber et al. 1998) que procura captar a variação funcional dos contextos de ocorrência (1998: 145-146): “In this approach [multi-dimensional analysis], co-occurrence patterns are identified quantitatively, based on the actual distributions of linguistic features in a large corpus of texts [...] Each set of co-occurring features is called a “dimension” of variation.[...] After the linguistic features defining a dimension are identified through factor analysis, the dimension is interpreted functionally, in terms of the situational, social, and cognitive functions most widely shared by the linguistic features. This interpretation is based on the assumption that co-occurrence reflects shared function.”

4 A hipótese da presença de rotinas de idiomatização em textos não literários pode ser problemática. Veja-se Sinclair (1991: 109-110) sobre o papel do princípio opcional e do princípio idiomático (*open-choice principle* e *idiom principle*) na actividade discursiva.

5 Para a poesia trovadoresca, consulte-se “on-line” a *Base de Dados da Lírica Profana Galego-Portuguesa* (ver Fontes).

6 Considerei como oração pequena as estruturas em que um SN, com a função de sujeito, tem como predicado um SP ou SA (ver Raposo 1992: 217), p.ex.: s. XIV: *veem per ella nom a conbecendo por coutada* (Marques e Dias 1990); s.XVI: *E elles, conbecendo a cousa feita e veendo as dictas fontes mui suaves* (Lucas 1988).

7 Incluí na primeira classe temporal Piel e Nunes 1988 (*A Demanda do Graal*) e Cepeda 1982-1989 (*Vida e Paixões dos Apóstolos*), apesar de o primeiro texto se reportar a um manuscrito datado do século XV (ver prefácio de Ivo Castro em Piel e Nunes 1988) e o segundo ser uma cópia do de 1442-1443 (Cepeda 1982-1989 e 1995) Nestas obras há a presença de algumas formas que hoje são tidas como arcaicas ou galegas, como a amálgama de pronomes *cho* (< *ti* + *o*; ver Huber 1986: 113, 174/175), o que permite recuar a época da sua produção.

8 Na Idade Média, a redundância enfática de certas expressões e fórmulas em textos notariais, permite deduzir relações de sinonímia: “conhocemos & confessamos Que teemos e tragemos & possoymos de uossa mão en uosso nome” (AOPT, cf. Parkinson 1983). A redundância nesta expressão encontra-se não só no uso de *conhocer* e *confessar*, mas também em *teemos* e *tragemos* e *possoymos*, que partilham a noção de posse.

9 Ver Buck 1988 para a etimologia de NOSCERE e COGNOSCERE.

10 Sobre *quiçá*, ver Machado 1987. Sobre *seica*, ver Alvarez et al 1986: 466.

- 11 Geeraerts (1997: 68) sintetiza o problema do seguinte modo: "Typical of this synchronous dynamism [the flexibility of lexical categories] is the emergence of incidental readings in the range of application of a word together with the fact that one and the same reading may diachronically emerge at different moments and on different grounds."
- 12 Sobre prototipicalidade, categorias e saliência, ver Geeraerts 1997: 10 e 20.
- 13 Para questões metodológicas relativas à identificação da poligénese semântica, ver Geeraerts 1997: 65.
- 14 Tal conclusão poderá chocar com os princípios fundamentais da teoria dos campos semânticos, conforme foram propostos por Trier (1933): a) não há sobreposições de sentido num mesmo campo; b) não há lacunas no sistema (Lehrer 1985: 284). No entanto, esta perspectiva tem sido continuamente revista (ver Lehrer 1985, Vassilyev 1974).
- 15 Sobre o conceito de *fuzziness*, ver Ungerer e Schmid 1996: 15.

Referências

- Álvarez, R. et alii 1986. *Gramática Galega*, Vigo: Editorial Galaxia
- Biber, D. 1988. *Variation across Speech and Writing*, Cambridge: Cambridge University Press
- 1998. *Corpus Linguistics. Investigating Language, Structure and Use*, Cambridge: Cambridge University Press
- Buck, C.D. 1988. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*, Chicago e Londres: The University of Chicago Press
- Campos, M.H.C. e Xavier, M.F. 1991. *Sintaxe e Semântica do Português*, Lisboa: Universidade Aberta
- Cepeda, I. V 1995. *Bibliografia da Prosa Medieval em Língua Portuguesa. Subsídios*, Lisboa: Instituto da Biblioteca Nacional e do Livro
- Geeraerts, D. 1997. *Diachronic Prototype Semantics. A Contribution to Historical Lexicology*, Oxford: Clarendon Press
- Lehrer, A. 1985. "The influence of semantic fields on semantic changes" in Fisiak, J. (ed.) *Historical Semantics. Historical Word-Formation*, Berlim, Nova Iorque, Amsterdão: Mouton Publishers
- Machado, J.P. 1987. *Dicionário Etimológico da Língua Portuguesa*, Lisboa: Livros Horizonte
- Nascimento, M.F.B., Marques, M.L.G. e Cruz, M.L.S. 1987. *Português Fundamental. Vol II. Métodos e Documentos. Tomo I. Inquérito de Frequência*, Lisboa: Instituto de Investigação Científica/ Centro de Linguística da Universidade de Lisboa
- Parkinson, S.R. 1983. "Um arquivo computadorizado de textos medievais portugueses", *Boletim de Filologia*, 28, Lisboa: Centro de Estudos Linguísticos, pp. 241-252
- Raposo, E.P. 1992. *Teoria da Gramática. A Faculdade da Linguagem*, Lisboa: Editorial Caminho

- Rocha, C 1994. "‘Saber’ e ‘conhecer’ em fórmulas e expressões fixas num ‘corpus’ de textos notariais dos sécs. XIII a XV", *Actas do X Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa: Colibri, pp. 481-491
- Silva, R.V.M 1989. *Estruturas Trecentistas. Elementos para uma Gramática do Português Arcaico*, Lisboa: Imprensa Nacional - Casa da Moeda
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*, Oxford: Oxford University Press
- Traugott, E. 1989. "On the rise epistemic meanings in English: an example of subjectification in semantic change", *Language*, vol. 65, n° 1, pp.32-55
- Taylor, J. 1995. *Linguistic Categorization. Prototypes in Linguistic Theory*, Oxford: Clarendon Press
- Trier, J. 1931. *Der deutsche Wortschatz im Sinnbezirk des Verstandes. Die Geschichte eines sprachlichen Feldes*, Heidelberg: Winter
- Ungerer, F. e Schmid, H.-J. Schmid. *An Introduction to Cognitive Linguistics*, Londres e Nova Iorque: Longman
- Vassilyev, L.M. 1974. "The theory of Semantic Fields: a Survey", *Linguistics*, 137, pp. 79-93
- Vilela, M. 1994. *Tradução e Análise Contrastiva: Teoria e Aplicação*, Lisboa: Caminho
- 1991, *Dicionário do Português Básico*, Porto: Edições Asa