

# Du traitement automatique des textes espagnols du Moyen Age à l'analyse sémantique : Une voie plantée d'importants jalons

JEAN ROUDIL  
(Paris)

## 1. Introduction

1.0. Index, glossaires et vocabulaires des textes médiévaux espagnols sont généralement l'expression de deux conceptions divergentes selon les publics que l'on veut atteindre: l'une répond à un souci d'exhaustivité, l'autre à une exigence d'utilité à l'égard du lecteur moderne<sup>1</sup>.

### 1.1. L'exhaustivité peut se manifester:

au niveau des adresses: *tous* les mots sont inventoriés et retenus pour la constitution du lexique, et à celui des références: *tous* les lieux de présence des mots dans le texte sont indiqués. Ce type de réalisation est rare.

au niveau des adresses: *tous* les mots sont inventoriés, mais non à celui des références: les lieux de présence des mots dans le texte ne sont indiqués que de façon complète ou partielle, ou pas du tout, suivant que le mot est conçu comme étant banal et courant, ou comme rare et intéressant. Mots perdus, mots sans vie, mots fossiles, bref, le critère de traitement est déterminé par la langue actuelle.

au niveau des exemples: la recontre d'un mot «jamais vu» ou «rare», ou peu étudié, a bien souvent piqué la curiosité du philologue qui se faisait alors une joie d'extraire d'autres textes les passages où ledit mot se trouvait par bonheur. Et l'on amassait avec délectation des données diverses, date d'écriture et nature des textes important peu!

---

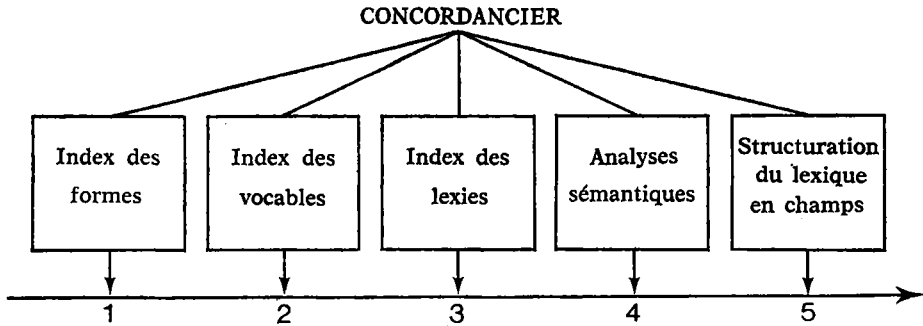
<sup>1</sup> Voir déjà, à ce sujet, Jean ROUDIL, «Contribution au dictionnaire onomasiologique de l'espagnol médiéval», in *Cahiers de Lexicologie*, t. 25, 1974, 2, pp. 3-4.

1.2. L'estimation des connaissances lexicales du lecteur du texte, à qui l'on veut éviter des difficultés de compréhension, détermine, elle, la retenue de tel ou tel mot. Seules certaines unités figurent dans la macrostructure, et la microstructure ne comportera que des définitions choisies, en étroite relation avec la complexité du signifié.

1.3. Index, glossaires et vocabulaires ainsi constitués sont une contribution, imparfaite, au *Dictionnaire de l'espagnol médiéval*.

Le traitement automatique des textes a ouvert une nouvelle voie. Des chercheurs doivent la parcourir ensemble et veiller à uniformiser les critères suivis ou à offrir, pour le moins, des données et des résultats utilisables par des équipes différentes<sup>2</sup>.

Je tiens pour essentielles cinq étapes, qui font l'objet du présent article et dont l'ordre chronologique s'inscrit sur l'axe ci-dessous:



## 2. *Index des formes*

### 2.1. DÉFINITION.

L'index des formes est un simple relevé alphabétique des formes d'un texte, c'est-à-dire des unités graphiques séparées par un blanc, avec mention de la fréquence absolue de celles-ci mais sans indication des références<sup>3</sup>.

Son établissement repose sur deux principes:

- a. les formes distinctes constituent chacune un item;
- b. les homographes ne sont pas discriminés et figurent donc ensemble.

<sup>2</sup> Sur un début de collaboration entre des groupes de recherche des Universités de Poitiers, de Paris VIII et de Paris XIII, voir René PELLEN, Véronique HUYNH-ARMANET, Jean ROUDIL, «Traitement automatique de textes espagnols du Moyen-Age et analyse des données», in *Cahiers de Linguistique Hispanique Médiévale*, t. 2, 1977, pp. 271-274 (diffusion: Librairie Klincksieck, 11, rue de Lille, Paris).

<sup>3</sup> Il n'est naturellement pas exclu de fournir des références dans un index des formes.

Deux remarques s'imposent:

a. le contour d'une forme dépend —bien malheureusement— des critères adoptés par l'éditeur; avant toute réalisation, il convient de mener une réflexion critique, approfondie et prudente, sur la qualité du texte qui fait l'objet de l'index et le degré atteint dans le respect ou la déformation du ou des manuscrits;

b. les formes du texte peuvent être respectées intégralement ou donner lieu à une pré-édition.

## 2.2. EXPÉRIENCES.

### 2.2.1. *El Fuero de Cuenca (Códice Valentino)*<sup>4</sup>.

Texte utilisé: *Fuero de Cuenca (Formas primitiva y sistemática: texto latino, texto castellano y adaptación del Fuero de Iznatoraf)*. Edición crítica, con introducción, notas y apéndice por Don Rafael de Ureña y Smenjaud, Madrid, 1935.

Préalablement à toute réalisation d'index, j'ai pris le parti de séparer de nombreuses formes groupées.

Par exemple:

#### Conjonctions et articles ou pronoms

elos → e los  
 olo → o lo  
 quela → que la  
 que lo → que lo  
 siel → si el.

#### Prépositions et articles

al, alos, ala, alas → a l, a los, a la, a las  
 del, delos, dela, delas → d el, de los, de la, de las  
 enel, enlos, enla, enlas → en el, en los, en la, en las  
 entrel → entre l  
 fastal → fasta l  
 sobrel → sobre el.

<sup>4</sup> Voir Jean ROUDIL, *La tradition manuscrite forale de Cuenca. Essais de linguistique textuelle*, thèse d'Etat, Paris, 1972. La partie intitulée «Lexicométrie» traite très longuement et en détail des problèmes théoriques posés par les index des formes, des vocables et des lexies; elle est suivie des différentes réalisations des index.

## Prépositions et pronoms ou adjectifs démonstratifs ou possessifs

asu	→ a su
auos	→ a uos
della, dellos	→ d ella, d ellos
desta, destas, desto	→ d esta, d estas, d esto
desu	→ de su
enella	→ en ella
sobrella	→ sobr ella.

## Prépositions et verbes

asaber	→ a saber
adexar	→ a dexar.

## Prépositions et substantifs

apuerta	→ a puerta
atuerto	→ a tuerto
defuero	→ de fuero
denoche	→ de noche.

## Verbe et préposition

ade	→ a de.
-----	---------

## Verbes et pronoms

rreçibala	→ rreçiba la
destruyanlas	→ destruyan las
prendanle	→ prendan le
denles	→ den les
despennenlo	→ despennen lo
defendiendose	→ defendiendo se
adiestro	→ a diestro
amedias	→ a medias
asabiendas	→ a sabiendas
asiniestro	→ a siniestro.

Ces séparations opérées, j'ai donné à l'index des formes deux présentations:

## a. l'une alphabétique:

agua	24
aguado	1
aguaducho	1, etc.
↓	↓
forme	fréquence absolue

b. l'autre par ordre de fréquence décroissante.  
Pour chaque fréquence l'ordre est alphabétique.

Au début de chaque fréquence importante, le nombre total de mots qui l'offrent est indiqué.

dia	219 (3)
querelloso	219
vn	219
sobre	209
conçejo	205
calonna	199 (2)
fasta	199, etc.

J'ai pu établir que:

a. le texte du *Fuero de Cuenca (Códice Valentino)* compte 64.035 formes;

N (étendue du texte): 64.035

b. le nombre de formes différentes y est de 3.926.

### 2.2.2. *Poridat de las poridades*<sup>5</sup>.

Texte utilisé: Seudo ARISTÓTELES, *Poridat de las poridades*. Edición de Lloyd A. Kasten, Madrid, 1957.

Dans un premier temps, j'ai préféré respecter les formes offertes par le texte; deux index ont été ainsi réalisés:

a. l'index alphabétique des formes;

b. le premier index des formes par ordre de fréquence décroissante, ou index primitif.

Mais cette conservation des unités graphiques m'a contraint à envisager un second index des formes ou index primitif corrigé, où les formes groupées ont été séparées.

Par exemple:

del d el  
al a l.

Les données fournies par ce second index sont comparables à celles du *Fuero de Cuenca* —(par similitude des critères suivis; voir sous 2.2.1.)—; elles me permettent d'établir que:

a. l'étendue du texte —(N)— est de 14.259 formes;

b. le nombre de formes différentes est de 2.367.

<sup>5</sup> Voir Jean Roudil, «Documents lexicométriques: «Poridat de las poridades» (Index des formes et index des vocables)», in *Cahiers de Linguistique Hispanique Médiévale*, t. 2, 1977, pp. 119-169.

### 2.2.3. *Primera Crónica General.*

Texte utilisé: *Primera Crónica General de España que mandó componer Alfonso el Sabio y se continuaba bajo Sancho IV en 1289*. Publicada por Ramón Menéndez Pidal, Editorial Gredos, Madrid, 1955.

L'index<sup>6</sup> reproduit alphabétiquement les formes telles qu'elles apparaissent dans l'édition mentionnée ci-dessus.

### 2.2.4. *El Libro de los gatos.*

Texte utilisé: l'excellente transcription, respectueuse des graphies, réalisée par une de mes étudiantes, Mademoiselle Annie-Noëlle Peïdro, et qui doit donner lieu à une édition collective comportant plusieurs travaux.

L'étude des graphies des textes espagnols du Moyen Age est une tâche urgente à réaliser, mais bien rares sont les éditions qui permettent de la mener à bien.

L'édition du *Libro de los gatos* rend compte très fidèlement et très rigoureusement du manuscrit. Elle permettra la confection, en premier lieu, d'un index nouveau, l'index alphabétique des *formes graphiques*, réellement graphiques et originales, qui sera suivi d'index de conception semblable à celle décrite précédemment.

## 3. *Index des vocables*

### 3.1. DÉFINITION.

Si l'index des formes est synonyme de dispersion —chaque forme représentant une adresse sans distinction aucune de la valeur, du sens et de la fonction— celui des vocables est synonyme de regroupement et de séparation. Il doit conserver un degré de généralité suffisamment élevé pour pouvoir être comparé aux index de vocables établis à partir d'autres textes; le moyen d'y parvenir est d'abstraire les mots de leurs réalisations de «norme», contextuelles, dont l'index des lexies aura le privilège de l'étude.

La répartition des formes par espèces, la distribution des catégories grammaticales qui donneront lieu ensuite à l'établissement de pourcentages divers est ici le travail essentiel.

---

<sup>6</sup> Il constitue le n.º 4 bis des *Cahiers de Linguistique Hispanique Médiévale*, Paris, 1979.

En effet:

- a. les homographes sont séparés;
- b. les hétérographes sont réunis:
  - groupement des formes différentes sous une seule adresse;
  - groupement des formes fléchies:
    - $\alpha$ / singulier et pluriel d'un substantif,
    - $\beta$ / masculin et féminin, singulier et pluriel d'un adjectif,
    - $\gamma$ / modes, temps, personnes d'un verbe;

c. l'appartenance d'un vocable à une catégorie grammaticale est indiquée; elle est déterminée par le fonctionnement dans le texte soit du vocable, soit des composants de celui-ci.

La lemmatisation incombe au chercheur; elle est réalisée sur la base des résultats fournis sur ordinateur (concordanciers).

L'index des vocables revêt deux présentations: l'une est alphabétique, l'autre est une liste des vocables par ordre de fréquence décroissante.

### 3.2. EXPÉRIENCES: *El Fuero de Cuenca (Códice Valentino)* et *Poridat de las poridades*<sup>7</sup>.

#### 3.2.1. *Présentation de l'index des vocables.*

Le principe est le suivant: un vocable à forme unique est suivi, sur la même ligne:

- a. du symbole chiffré indicateur de la catégorie grammaticale à laquelle il appartient;
- b. du chiffre indicateur du nombre total d'occurrences du vocable lemmatisé;
- c. de chiffres qui renvoient au texte étudié (pages, lignes, paragraphes).

Les symboles chiffrés que j'ai choisis sont:

- 0 Substantifs.
- 1 Verbes.
- 2 Adjectifs qualificatifs.
- 3 Adjectifs autres que qualificatifs.
- 4 Adverbes.
- 5 Pronoms.
- 6 Noms propres.
- 7 Articles.

---

<sup>7</sup> Cf. les notes 4 et 5.

- 8 Prépositions.
- 9 Conjonctions.
- 10 Chiffres.
- 11 Lettres et abréviations.

Il est souhaitable que d'autres chercheurs les retiennent.

En présence des différentes graphies et formes, formes fléchies et fonctions d'un vocable, des options de présentation sont prises; elles font toujours l'objet d'une description minutieuse.

### 3.2.2. *Catégories grammaticales et lemmatisation. Aspects généraux.*

La répartition des formes par catégories et leur intégration dans un index des vocables pose d'importants problèmes théoriques, dont l'ampleur et la complexité ont pour compensation l'état agréable et fécond d'a-connaissance dans lequel vous place l'index des formes. En vrac, sous *que* et sous *fuere*, sous *vezino* et sous *judio*, etc. les occurrences totales sont fournies, sans distinction aucune et sans arbitraire imposé. Or c'est bien souvent arbitrairement et trop systématiquement que les éléments du vocabulaire ont été répartis entre diverses classes «eu égard à leur nature, leur formation ou leur fonction»<sup>8</sup>. Le cloisonnement ne tarde pas à nous apparaître d'une rigidité implacable et bien vite nous interrogeons-nous sur les motifs du passage d'une catégorie à une autre.

Il est donné à une unité de changer de catégorie: tel adjectif fonctionne comme substantif dans une séquence déterminée, et tel verbe comme substantif. Comment intégrer ces fluctuations dans un index, rigide par définition? Deux solutions sont à envisager:

1. les fonctions d'adjectif et de substantif d'une unité sont réunies sous une seule adresse qui forme un vocable;
2. chaque fonction différente constitue un vocable

adjectif	vocable 1
adjectif → substantif	» 2
adjectif → adverbe	» 3
verbe	» 4
infinitif → substantif	» 5, etc.

Cette ultime classification a l'avantage de la clarté et de la netteté; sa maniabilité est certaine dans l'utilisation statistique de l'index des vocables. Son inconvénient majeur est qu'elle est peu flexible, car une unité linguistique en fonction n'appartient pas irrévocablement à une catégorie; elle a certes une vocation grammaticale majoritaire, mais, intégrée dans un énoncé, il lui est donné d'acquérir une autre valeur. En fin de compte elle se définit par l'acceptation ou le refus d'un certain nombre d'options.

<sup>8</sup> J. MAROUZEAU, *Lexique de la terminologie linguistique*, Paris, 1961, p. 44.



UNITÉS	OPTIONS	ACCEPTATION +	REFUS -
Unités I	Unique	Substantif	Les autres catégories
Unités II	Double	1. Substantif 2. Adjectif	»
Unités II'	»	1. Substantif 2. Adverbe	»
Unités III	Triple	1. Substantif 2. Adjectif 3. Adverbe	»
Unités IV	Unique	Adjectif	»
Unités V	Double	1. Adjectif 2. Substantif	»
Unités V'	»	1. Adjectif 2. Adverbe	»
Unités VI	Triple	1. Adjectif 2. Substantif 3. Adverbe	»
Unités VII	Unique	Verbe	»
Unités VIII	Double	1. Verbe 2. Substantif	»
Unités IX	Unique	Participe passé	»
Unités X	Double	1. Participe passé 2. Adjectif	»
Unités X'	»	1. Participe passé 2. Substantif	»
Unités XI	Triple	1. Participe passé 2. Adjectif 3. Substantif	»

Les unités à options forment un système virtuel dont la réalisation ou la non-réalisation relèvent de l'actualisation. La chance d'apparition des options, dans un type d'énoncé déterminé, doit faire l'objet d'études statistiques; l'on peut en effet prévoir le fonctionnement d'une unité comme adjectif ou comme substantif et lui attribuer un pourcentage sûr. L'établissement des fluctuations chiffrées d'une unité entre catégories grammaticales diverses contribuera à la définition et à la fixation des normes linguistiques.

Mais le problème de l'index des vocables présente deux aspects: celui de la codification des unités et celui de leur répartition sous un lemme. J'ai décidé que des unités différemment codifiées pouvaient figurer sous la même adresse-mère, c'est-à-dire sous le même vocable. La valeur linguistique de celui-ci est aléatoire parfois, mais il s'agit d'une notion nécessaire en statistique, qui s'oppose très bien à celle de forme et trouve ainsi sa pertinence.

Par exemple, pour ce qui est du *Fuero de Cuenca*:

LEMES	FONCTIONS
fiel	} adjectif et substantif
porfioso	
presente	
propinco	
propio	
quarto (quarta)	
querelloso	
siedmo (siedma)	} adjectif et adverbe
falso	
saluo	
sacado	adjectif et préposition
	verbe et préposition.

A noter que la réunion des formes fléchies du verbe sous un vocable implique logiquement celle des unités à fonctions variées.

A ce système, toujours susceptible de donner lieu à d'autres regroupements, des aménagements peuvent être apportés. C'est ainsi que j'ai considéré comme vocables à part entière:

a. l'infinitif substantivé (0)

comer

b. le substantif issu d'un adverbe (0)

demas

c. le substantif issu d'un participe passé (0).

Ainsi:

acusado, alquilado, asoldadado, cansados, casada, casados, caydo, cononbrados, criado, culpado, dado, demandado, desafiado, dicho, encomendado, enpennado, enplazado, escripto-escrito, fecho, ferido, fiado, finado, foydo, gannado, llagados, mandado, muerto, poblado, prendado, preso, rreptado, senbrado, sobre leuado, torçeado, tornada, venida, vista, yda.

d. l'adjectif issu d'un participe passé (2)

menguado.

### 3.2.3. *Catégories grammaticales et lemmatisation. Aspects particuliers.*

Les problèmes spécifiques, propres à chaque texte, seront toujours présentés en détail; c'est ainsi que dans ma thèse<sup>9</sup> j'aborde les points suivants:

- Le participe passé *preso*;
- Identité du lexème verbal, interférences et échanges au niveau des signifiés et des grammèmes: type *prender/prender*;
- Fuere*: unité verbale à deux têtes signifiées;
- Adjectif et substantif;
- Substantifs —(ou autres catégories)— intégrés dans un ensemble;
- Que*;
- Formes contractées de l'article défini;
- Se* + verbe;
- Unité + *mente*;
- Questions diverses plus spécifiques.

### 3.2.4. *Exploitation.*

Le texte du *Fuero de Cuenca (Códice Valentino)* a 2.017 vocables (V —nombre des vocables— = 2.017).

N devient, dans l'index des vocables: 62.345.

Le compte de ces vocables permet de dresser différents tableaux dont la nature est fonction du but recherché; pour le texte mentionné ci-dessus, j'ai dressé:

- a. un tableau de distribution des fréquences;
- b. un tableau de distribution des catégories grammaticales;
- c. un tableau de distribution des catégories grammaticales par ordre de fréquence décroissante;
- d. un tableau de distribution des formes verbales par ordre de fréquence décroissante;
- e. un tableau de fréquence décroissante des vingt-cinq premiers substantifs;
- f. un tableau de fréquence décroissante des vingt-cinq premiers verbes.

## 4. *Index des lexies*

### 4.1. DÉFINITION.

Le relevé des formes fournit une matière brute. L'index des vocables est une première reconstitution qui permet, d'une part, une utilisation

<sup>9</sup> Cf. note 4.

poussée à des fins statistiques et qui offre, d'autre part, une distribution des unités par catégories grammaticales. Parce qu'il est un outil devant servir à des études comparatives, son critère se doit de demeurer analytique, dans la mesure du possible; mais il s'ensuit des incohérences linguistiques que seul est capable de rectifier et de corriger l'index des lexies<sup>10</sup>.

Celui-ci est fonctionnel; il tient compte en effet de l'environnement d'une unité et de sa capacité à former un ensemble complexe avec les segments qui l'entourent.

#### 4.2. EXPÉRIENCE: *El Fuero de Cuenca (Códice Valentino)*.

L'index des lexies est à la fois un héritage et une amélioration de l'index des vocables. L'héritage repose sur la conservation de la plus grande partie des lexies simples et sur le maintien intégral des lexies complexes dont la place dans l'index des vocables s'explique par leur utilisation générale et, donc, leur appartenance au vocabulaire commun.

Par exemple:

açada	cab adelante
acalonnador	cada vno
camarera	de nuevo
desmemoriado	tan demientra.

Mais certaines lexies, réunies sous un vocable, devront être élevées au rang de vocable, de lemme. Ainsi, *por sienpre, para en sienpre, para a sienpre*, qui, dans l'index des vocables, figurent sous *sienpre*, devront constituer trois lemmes dans celui des lexies.

L'amélioration consiste à introduire de nouveaux vocables caractérisés par leur étendue: ce sont les lexies complexes que le caractère analytique de l'index des vocables m'avait interdit d'intégrer. Ce nouveau regroupement de certaines unités en une seule peut bouleverser les fréquences et les catégories grammaticales de l'index des vocables.

Par exemple:

*ancho* est codifié 0, les occurrences de cette forme étant au nombre de 6. Nous lisons *en lo ancho* deux fois; *en ancho* quatre fois. Dans l'index des lexies, *en ancho* élevé au rang de lemme, serait codifié comme adverbe, à cause de son incidence au verbe.

De la même façon, la prise en considération de *en alto* comme lexie modifiera les données de l'index des vocables:

<sup>10</sup> Pour l'index des vocables, les mots sont sortis et extraits de leur situation. Dans l'index des lexies, ils y sont remplacés: situation «fuero», par exemple. Je retrouve ici ce que dit Bernard POTTIER: «En français, *eau, glace, vapeur*, sont des lexies simples, mais *eau courante, eau chaude* (dans la situation «hôtel») sont des lexies complexes.» («Le domaine de l'ethnolinguistique», in *Langages*, t. 18, 1970, p. 4.)

I. DES VOCABLES

en alto } 2 vocables } en  
          }          } alto  
          } Codifiés } .préposition  
          } comme : } .substantif

I. DES LEXIES

en alto } 1 lexie  
          } Codifiée  
          } comme : adverbe.

Le choix du critère à suivre pour décider qu'un ensemble d'unités médiévales constitue une lexie complexe est de la plus haute importance. C'est la fréquence élevée de la réalisation d'une suite d'unités qui permet de trancher, fréquence que l'on oppose aux occurrences des unités prises individuellement.

Je parle de lexies complexes:

- 1.° lorsque deux ou trois unités ne sont employées que réunies, ensemble, sans connaître isolément d'emploi;
  - 2.° lorsque deux ou trois unités sont employées plus fréquemment ensemble qu'isolément, connaissant ainsi un pourcentage élevé d'emploi.
- Parmi les lexies établies, je distingue:

a. celles qui sont d'un emploi général.

Par exemple:

I. DES VOCABLES		I. DES LEXIES				
			<i>lexies</i>	<i>occurrences</i>	<i>vocables</i>	<i>occurrences</i>
amor	0 4	—	[por amor de ]	2		
			[por amor d ]	1		
					amor	1
cabo	0 5	—	en cabo de	3		
		—	al cabo de	1		
		—	al cabo	1		
[fijo, 74 ]	0 76	—	fi de fodido	2		
[fi, 2 ]					fijo	74

*Fi* ne se réalise que dans la lexie *fi de fodido*.

menester	0 16	—	[ es menester ]	2		
			[ menester fuere ]	2		
			[ fuere menester ]	9		
			[ fueren menester ]	1		
		—	oviere menester	1		
		—	fiziere menester	1		
ojo	0 7	—	[ a ojo de ]	1		
			[ a ogio de ]	1		
		—	estouiere a ojo	1		
					— [ojo ]	3
					[ojos ]	1
ventura	0 36	—	por ventura	36		

b. *celles qui sont propres à la norme forale et qui la caractérisent*<sup>11</sup>.

## I. DES VOCABLES

## I. DES LEXIES

canpo 0 43

## lexies occurrences

— [	.derriipten lo en canpo e	} 16
	saluden lo en conçejo (9)	
	.derreipten lo en canpo e	
	saluden lo en conçejo (1)	
	.derriipten lo en canpo e	
	salude lo en el conçejo (1)	
	.derriipten lo en canpo (1)	
	.sea creydo e luego	
derreptado en el canpo (1)	}	
.sea creydo e en canpo		
— [	derreptado (3)	]

Soit 16 occurrences où sont simultanément en présence le lexème verbal *derrept-* et *canpo*. *Derrept-* (*riepro*) modifie le signifié de *canpo*, qu'il particularise et restreint; par implication logique, il contient même en lui le concept *canpo*, pouvant ainsi se passer de l'actualisation concrète et de la présence de ce dernier. *Derrept-*, en première instance, ne se conçoit pas sans *canpo*; réciproquement, cette unité, dans le sens de 'champ de bataille judiciaire', enferme le concept 'défi, défier' qui l'informe toujours, même non réalisé. Ainsi:

el juez τ los alcades demuestrenles los mojonos del canpo τ partanles el sol., p. 529, 2, ll. 7253-7255

caya 1j 35

— [	.caya del pleito (26)	} 28
	.caya del pleyto (1)	
	.caya de pleito (1)	

creydo 1d 124

— sea creydo 120

doze 3 79

— [	.salue se con doze	} 53
	vezinos (46)	
	.se salue con doze	
	vezinos (6)	
	.saluen se con doze	
— [	vezinos (1)	]

entrega 0 2

— ayan entrega de las calonnas 2

fiador 0 15

[	.fiador de saluo (1)	} 11
	.fiadores de saluo (10)	

fiadores 0 26

fiadura 0 12

[	.fiadura de saluo (8)	} 9
	.fiaduras de saluo (1)	

fiaduras 0 2

prenda 0 126

[	.defendiere prenda (10)	} 13
	.prenda defendiere (1)	
	.defendieren prenda (2)	

<sup>11</sup> Par exemple: *como fuero es, como dicho es*.

Rapprochée des occurrences de *prenda*, la fréquence de *defender prenda* peut sembler réduite. Pour déterminer alors s'il y a lexie ou non, il faut sortir du cadre limité d'un fuero particulier et envisager l'ensemble des textes dits «fuero». Nous prenons ainsi conscience de la présence obligatoire, dans chaque texte, de la formule *defender prenda*.

rresponda	1j	91	
par	0	20	— rresponda a su par 18
rriempto	0	22	— [ .rrespond { —ER al rriempto (1) —ER a rriempto (1) — A al rriempto (2) — A a su par a rriempto (1) } 15
			[ .respond { — A a rriempto (8) — A a rriempto (1) —AN a rriempto (1) } 15

On peut constater, devant *rriempto*, la présence fréquente du lexème verbal *rrespond-* suivi de la préposition *a*, avec ou sans actualisateur.

salga	1j	112	
enemigo	0	85	— [ .salga enemigo (68)
enemjgo	0	21	salga enemjgo (18)
enemiga	0	3	.salga enemiga (1) } 90
			.salga por enemigo (2)
			.salga por sienpre enemjgo (1)

Cette lexie peut être agrandie puisque, après *salga enemigo/enemjgo* nous rencontrons:

por vn anno, 45 fois;  
por sienpre, 10 fois.

*salue* 1j 118.

Après *se salue/salue se* sont fréquemment placés les segments suivants:

con doze vezinos, 52 fois;  
con dos de quatro cononbrados, 14 fois.

Une lexie appartenant au vocabulaire commun peut s'intégrer au vocabulaire juridique par sa situation dans un texte de lisibilité bien précise, le fuero, et par la fréquence élevée qu'elle y atteint; elle devient spécifique de la terminologie juridique. Ainsi, par exemple:

escogencia	0	4	[ .e esto sea en escogencia del querelloso (3) e esto sea en escojençia del querelloso (1) .pero en escojençia sea del querelloso (2) .e esto sea en escogencia de los parientes del muerto (1) } 7
escojençia	0	3	

L'élément nominal placé après *de* pouvant changer —bien que *querellosa* soit l'unité normalement attendue—, la lexie à fixer est: *ser en escogencia de*.

$$\text{lioures } 0 \quad 9 \quad \left. \begin{array}{l} \text{—} \left[ \begin{array}{l} .fiziere \text{ lioures } (3) \\ .lioures \text{ fiziere } (3) \\ \text{lioures non fiziere } (1) \\ .lioures \text{ non faga } (1) \end{array} \right\} 8 \end{array} \right\}$$

De nombreux ensembles formés avec *fazer* pourraient être considérés.

$$\text{saber } 1a \quad 43 \quad \left. \begin{array}{l} \text{—} \left[ \begin{array}{l} .es \text{ a saber } (30) \\ .es \text{ de saber } (6) \end{array} \right\} 36 \end{array} \right\}$$

Pour terminer, je noterai que la prévisibilité d'une réalisation n'est pas à confondre avec la possibilité de formation d'une lexie; elle est une condition nécessaire de celle-ci, mais non suffisante, la lexie représentant une étape postérieure dans laquelle les éléments qui la composent parviennent à un état de liaison ou de soudure très stable.

Les occurrences de l'adjectif *fechizo* sont au nombre de cinq; elles se situent toutes, régulièrement, après *juez*. L'on ne saurait cependant parler de la lexie *juez fechizo*, les deux segments gardant une grande autonomie et pouvant, successivement, faire place à d'autres segments:

$$\begin{array}{l} \text{juez} \longrightarrow \text{alcalde} \\ \text{fechizo} \left. \begin{array}{l} \longrightarrow \text{jurado} \\ \longrightarrow \text{annal.} \end{array} \right\} \end{array}$$

### 5. *Analyses sémantiques*

Un concordancier permet une lecture verticale des unités qui, à l'horizontale, se meuvent dans leurs environnements immédiats et médiats, et de gauche et de droite. Situation privilégiée qu'il faut exploiter en décrivant mécaniquement les unités dont on indique avec minutie, exhaustivement et de façon structurée la situation et la distribution dans le texte.

Cette étape accomplie, indispensable par son apport à la découverte de la signification, il convient de dévoiler la substance sémique des unités.

Ainsi ai-je procédé, méthodologiquement, pour étudier le signifié de:

- Fuero/dicho.
- Calonna.
- Coto.
- Exir/salir.



#### 6. *Structuration du lexique en champs. Conclusion*

Le dernier travail à accomplir est la structuration du lexique de l'oeuvre que l'on étudie. Peu de réalisations ont été faites jusqu'à présent. Fort heureusement nous disposons d'excellentes et fécondes réflexions théoriques dues à l'éminent collègue auquel nous rendons hommage ici<sup>12</sup>.

Il entre dans mes projets d'étudier le lexique de la *Primera Crónica General* du point de vue de sa structuration. La classification des adjectifs est commencée; dès que possible, je livrerai les premiers résultats de cette recherche.

---

<sup>12</sup> Je renvoie le lecteur aux travaux de M. E. COSERIU et, plus particulièrement, à l'article «Vers une typologie des champs lexicaux», in *Cahiers de Lexicologie*, t. 27, 1975, 2, pp. 30-51.

