

La distribució de la informació contextual en els elements estructurals d'un article de diccionari: col·locacions, restriccions lèxiques i definició

Judit Feliu Cortès
Joan Soler Bou
Institut d'Estudis Catalans

The main goal of this paper is to discuss the need of improving the encoding dimension of general monolingual dictionaries by considering the treatment of non-phraseological lexical combinations—particularly collocations. This will be attained through an analysis of the notion of collocation from a lexicographic point of view taking into account both empirical and theoretical approaches to the general problem of lexical combinations. Moreover, lexical information extracted from corpus must be analysed and included in general monolingual dictionaries bearing in mind that it can be distributed in different elements of an entry-definition, example, etc.—depending on each dictionary structure. Thus, authors will shed some light on the lexicographic task in order to determine whether the co-occurrence of two lexical items must be retrieved and, if so, how and where this lexical and semantic information should be organised. In this sense, it will be demonstrated that the definition and the example fields are not enough to retain and reflect the real use of collocations extracted from corpus. Some guidelines will be provided in order to help the lexicographer in the dictionary-making task concerning the distribution of the extracted information in the collocation field and also in the semantic-restriction pattern and its corresponding definition.

Introducció

L'estudi i l'interès pels fenòmens relacionats amb l'expressió lexicogràfica del significat contextual (i molt particularment de les col·locacions) en els diccionaris generals han adquirit progressivament més importància arran de la generalització i de la sistematització de la lexicografia basada en corpus, i de la constatació de la importància de considerar aquest tipus d'estructures com a elements clau en els processos de traducció, aprenentatge de llengües, i altres activitats molt estretament relacionades amb la lexicografia. El concepte de col·locació està, sens dubte, directament vinculat a la lingüística de corpus, fins al punt que la majoria de treballs que aborden aquesta qüestió s'interessen pels mecanismes de detecció i d'extracció d'aquestes unitats sobre repertoris de dades textuais més o menys extensos. Aquest concepte, però, té límits imprecisos, i el seu tractament en el programa de descripció dels diccionaris generals monolingües no ha estat encara abordat amb la profunditat que mereix.

Des d'una perspectiva estrictament lexicogràfica, es pot dir que, a part del seu valor eminentment descriptiu, les col·locacions adquireixen una importància cabdal en el desenvolupament de les funcions *codificadores* del diccionari (és a dir, les relacionades amb els usos *actius*) (Cf. Kroman *et al.* 1984). En aquest sentit, l'usuari d'un diccionari ha de saber que l'activació del substantiu *confiança* es fa mitjançant el verb *inspirar*, i que la intensificació de sentit de *lector* es fa mitjançant l'adjectiu *voraç*, per exemple. En el programa de descripció dels diccionaris, però, també és molt important el tractament de les col·locacions des de la perspectiva *descodificadora*, que dona compte dels tipus de substantius que *s'inspiren* o de les diverses entitats que poden resultar *voraces*. Aquestes dues funcions condicionen, en darrer terme, l'orientació dels diversos diccionaris específics sobre col·locacions: així com el BBI s'ocupa fonamentalment de la dimensió codificadora, el REDES està orientat a la funció descodificadora.

L'objectiu fonamental d'aquesta contribució —que parteix de la base que les col·locacions han de ser tractades també en els diccionaris generals monolingües i que han d'ocupar un espai rellevant en un article de diccionari— és l'anàlisi de la fenomenologia lingüística que presenten les col·locacions,

amb la finalitat de determinar si els elements estructurals i els constituents en què sol organitzar-se un diccionari general són adequats i suficients per al tractament d'aquest tipus d'estructures. Aquesta determinació es basa, sobretot, en els condicionaments que aquestes estructures plantegen sobre el mètode de descripció lexicogràfica, en especial sobre l'element estructural dels diccionaris que té per funció l'explicitació del significat del definit: la definició lexicogràfica.

Les col·locacions: aproximació empírica

Les col·locacions s'han definit des de diversos punts de vista, però l'aparició del terme es relaciona amb aproximacions de base empírica¹ que no parteixen de la inducció teòrica ni incorporen nivells abstractes a l'explicació del fenomen.

Aquesta aproximació es pot relacionar, a grans trets, amb la interacció formulada per Sinclair (1991) entre el *principi de lliure combinació* i el *principi de coocurrència*. D'acord amb el primer principi, un text o un enunciat qualsevol es pot entendre com el resultat d'una sèrie d'eleccions lliures només controlades per les regles gramaticals, que es pot presentar a través d'arbres on els elements terminals (els elements lèxics) són de lliure elecció. El principi de coocurrència estableix, en canvi, que els mots no apareixen a l'atzar en un text, i que les restriccions imposades per la gramàtica no són suficients per explicar les tendències a la coocurrència que presenten alguns elements lèxics. Segons aquest principi, un mot pot veure limitades les seves possibilitats de combinar lliurement amb un altre i, malgrat això, formar frases de sentit analitzable componencialment.

Les col·locacions, des d'aquest enfocament, se solen caracteritzar per estar formades per elements que apareixen *freqüentment* de manera conjunta en un corpus. Tot i que no es tracta d'una definició formal i no exclusiva d'altres possibilitats de combinació, aquesta caracterització —que determina les estratègies de detecció de col·locacions en corpus (Church & Hanks 1990; Kilgarriff & Tugwell 2001) mitjançant la relació entre diferents magnituds freqüencials— sembla que és a la base de molts treballs sobre col·locacions i fonamenta alguns dels criteris que segueixen determinats diccionaris.

Les col·locacions: aproximació semàntica i definició formal

La semàntica lèxica ha abordat el fenomen de les col·locacions (o, més ben dit, de les estructures que solen ésser considerades col·locatives) a partir de diferents punts de vista. Aquests punts de vista es poden relacionar amb tres conceptes generals, que permeten establir tres caracteritzacions no coincidents, però superposables parcialment les unes amb les altres i susceptibles d'originar definicions diverses:

- les *solidaritats lèxiques* plantejades per Coseriu (1977);
- les *restriccions lèxiques* que són, de fet, no tan sols una ampliació conceptual dels mateixos principis en què es basen les *restriccions seleccionals* de Katz i Fodor (1963), sinó un intent de sistematitzar les restriccions lingüísticament pertinents, només parcialment coincidents amb el fenomen de les col·locacions,
- la definició de col·locació de Mel'čuk (1998) —lligada amb la noció de *funció lèxica* (FL)—, dins del marc teòric de la Teoria sentit-text.

¹ Aquesta aproximació implica una anàlisi del fet col·locatiu lligat fonamentalment a l'observació i a l'anàlisi empírica del llenguatge, més que no pas un concepte induït des de la lingüística teòrica. De fet, quan Firth (1951) introdueix el terme *col·locació* (en el sentit de les possibilitats de combinació lèxica d'un mot), l'usa per parlar de les expectatives d'aparició mútua (*mutual expectancy*) com un procés de significat per coocurrència, diferent del significat conceptual i situacional. Sinclair (1991) formula de manera similar la seva definició de col·locació ("Collocation is the occurrence of two or more words within a short space of each other in a text", p. 170), i, sobretot, pel que fa a la metodologia d'avaluació de la importància d'una coocurrència en relació amb la freqüència dels elements que la componen.

Aquestes tres aproximacions diferents delimiten i descriuen diferents tipus de realitats lingüístiques. D'acord amb Coseriu (1977) una combinació (en especial allò que ell anomena *solidaritats lèxiques multilaterals*) pot estar determinada per un d'aquests tipus d'elements:

- a) un lexema: *clau*, en el cas de *reblar* ("doblegar la punta de") *un clau*.
- b) un arxilexema: «cavall» + *renillar*, on [cavall] es pot considerar l'arxilexema dels lexemes *corser*, *poni*, etc.
- c) una classe semàntica: *cometre* + [acció punible], on l'*acció punible* és la classe que agrupa el camp lèxic format per *error*, *crim*, *falta*, *pecat*, *delicte*, *imprudència*, *injustícia*, etc.

L'aproximació basada en les restriccions lèxiques té una de les seves concrecions més notables en el diccionari REDES. Aquest diccionari recull sistemàticament les restriccions significatives que els predicats imposen als seus arguments. Una unitat predicativa pot seleccionar nocions de naturalesa molt diversa: objectes físics, llocs, esdeveniments, propietats, etc.

Des del nostre punt de vista, l'enfocament d'aquests tipus de restriccions és una altra manera d'interpretar el que Coseriu anomena solidaritats lèxiques. Bosque distingeix entre selecció categorial (*selecció-c*) i selecció semàntica (*selecció-s*). Així, la restricció de subjecte que imposa una unitat com *emprendre* és una selecció-c (*algú*, *un ocell*, *un vehicle*, etc.), però la de complement d'objecte és semàntica (selecció-s): *s'emprenen* (és a dir, es *comencen*) coses equiparables a desplaçaments o viatges, que solen tenir inici i final; *s'emprenen camins* però no *carreteres* o *viaranyes*, *s'emprèn el vol*, però no altres accions com un *salt*, per exemple.

Per la seva banda, la definició *lingüística* de Mel'čuk (1998: 30) es basa en el caràcter *adaptable* per idiomaticitat d'un dels dos constituents de la col·locació:

A collocation of **AB** of language **L** is a semantic phraseme of **L** such that its signified 'X' is constructed out of the signified of one of its two constituent lexemes –say, of **A**– and a signified 'C' [$'X' = 'A \oplus C'$] such that a lexeme **B** expresses 'C' only contingent on **A**

Mel'čuk distingeix dos casos possibles d'aplicació d'aquesta definició en funció del fet de si B té dins el diccionari el sentit de C o si, per contra, no el té. Cadascun dels dos casos es divideix, al seu torn, en dues possibilitats més. Aquest autor exemplifica la no inclusió de B en el diccionari amb el sentit de C amb estructures del tipus *fer un favor* i *succe morè* i, els d'inclusió de B amb el sentit de C, amb exemples del tipus *el cavall renilla* o *mantega rànica*.

Aquesta formulació recull, doncs, el caràcter idiomàtic, i alhora transparent, d'aquest tipus de combinacions lèxiques. Aquest procés s'expressa a través del manteniment del significat d'un dels constituents, que selecciona un sentit adaptat per al lexema amb el qual combina (*malaltia fulminant*, *afirmar categòricament*). Tradicionalment, s'ha utilitzat un nombre tancat i reduït d'aquestes funcions (*magn*, *oper*) per sistematitzar alguns tipus de col·locacions, però no s'ha aprofundit en l'etiquetatge de les col·locacions utilitzant totes les possibilitats de FL sintagmàtiques i paradigmàtiques que ofereix aquesta teoria.

Les col·locacions en la perspectiva de la lexicografia general

Les dues aproximacions (freqüencial i semàntica respectivament) sobre què s'entén per col·locació des de la lexicografia general no són totalment coincidents pel que fa a la determinació de l'objecte descrit. Aquesta dicotomia es palesa quan Bosque renuncia a la utilització del terme *col·locació* en el seu diccionari. Aquest fet té una implicació especial en la lexicografia general. Considerades des d'un punt de vista freqüencial, les col·locacions i les anomenades restriccions de selecció no es diferencien en un gran nombre de casos; ara bé, considerades des del punt de vista semàntic, les col·locacions només afecten les bases que es mantenen semànticament inalterades.

Com ja ha estat dit a bastament, les col·locacions juguen un paper de primer ordre en el desenvolupament de les funcions tant actives (codificadores) com passives (descodificadores) del diccionari, tant per als aprenents de la llengua com per als usuaris en general. Tot i així, la manera com la informació col·locativa (o, més generalment, sintagmàtica) es distribueix en els

diferents elements estructurals del diccionari varia substancialment d'un diccionari a un altre. No existeixen, de moment, anàlisis serioses d'aquesta distribució, ni tampoc propostes metodològiques clares formulades des de la metalexicografia en aquest sentit.

L'element estructural del diccionari que ha hagut d'adaptar més les seves característiques al tractament de fenòmens de tipus col·locatiu (i en general totes les qüestions relacionades amb l'expressió del significat contextual en els diccionaris) és la definició lexicogràfica, que ha estat concebuda (Rafel 2006) com la unió de dos tipus de constituents: els elements *intrínsecs* i els elements *extrínsecs*.²

La qüestió que esdevé fonamental en aquesta contribució, però, és la següent: pot la definició lexicogràfica distingir eficaçment entre els diferents tipus d'elements i estructures que determinen l'existència d'una col·locació concreta a partir de les dades extretes d'un corpus textual? Pot diferenciar adequadament aquest tipus d'estructures d'altres modalitats sintagmàtiques, com ara les restriccions lèxiques o les solidaritats lèxiques?

Un element extrínsec és un constituent de la definició que pot presentar diversos graus d'abstracció lingüística: adaptant la divisió de Coseriu que hem comentat més amunt, podem establir que un element extrínsec que representa una solidaritat lèxica multilateral pot representar un lexema, un arxilexema, o una classe semàntica. Per altra banda, la definició lexicogràfica no ha desenvolupat, almenys en la lexicografia convencional, cap sistema per distingir entre aquests tres tipus de casos. En general, els casos que requereixen una explicació més específica són precisament les *col·locacions*, que per aquest motiu requereixen la presència d'elements estructurals específics, ja que la definició no pot delimitar-les adequadament a partir dels seus elements extrínsecs. Per exemple, en el cas posat anteriorment de *reblar* hem vist que el CTILC mostra l'existència d'una selecció lexemàtica respecte de *clau*. Es dona també aquesta selecció en altres casos com *clavar*? Almenys la definició que en dona el DIEC2, a partir de la manera com expressa el seu element extrínsec, ens ho pot fer pensar. El CTILC sembla que confirma que és així.

reblar v. tr. Doblegar la punta (d'un clau tan llarg que, un cop clavat, surt per l'altre costat) per tal que el clau estigui més fortament fixat i la punta no pugui fer mal.

clavar v. tr. Fer entrar a cops de martell (un clau) dins un cos sòlid. *Clavar un clau a la paret.*

En altres casos, però, la presència d'un element extrínsec no es correspon amb una selecció lexemàtica de les combinacions del definit amb els seus arguments, sinó que representen opcions més complexes lligades a un arxilexema o bé a una classe semàntica. Vegem-ho en el cas següent:

interposar v. tr. (...) Formalitzar (un recurs legal). *Interposar un recurs de nul·litat, d'apel·lació, administratiu.*

En aquest cas, la informació sobre la restricció lèxica apareix com a element extrínsec en la definició i també en els exemples d'ús. Aquesta informació també es recull en el substantiu corresponent:

recurs m. Acció de demanar l'anul·lació o la modificació d'un acte administratiu o d'una resolució judicial. *Interposar un recurs. Recurs ordinari. Recurs d'apel·lació, de cassació, de revisió.*

El CTILC, per la seva banda, corrobora l'especialització d'aquesta restricció lèxica, però alhora en dona d'altres que, amb una freqüència menor, també actualitzen aquest sentit d'*interposar*. Això es pot veure en la Taula 1:

² La denominació d'aquests dos constituents com a *elements extrínsecs* i *elements intrínsecs* respectivament correspon a Rafel (1989). Especialment els elements extrínsecs, o certes entitats més o menys equivalents en altres perspectives d'anàlisi, han estat denominats de manera diversa. Per citar-ne dues de rellevants: *entourage* (Rey-Debove 1971), *contorno* (Seco 1979).

<i>Lemes</i>	<i>Total</i>	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5
recurs	29						10	15	1	1	2
plet	3						1	2			
reclamació	3							3			
demanda	2							2			

Taula 1

En el cas d'*interposar*, doncs, la definició recull com a element extrínsec una restricció lèxica que en termes de Coseriu podríem determinar com un arxilexema, sense recollir, però, la resta de combinacions possibles que dona el corpus per a aquest ús. Alhora, aquesta restricció és expressada com a combinació lèxica en l'article d'*interposar* i de *recurs* respectivament; aquestes combinacions lèxiques són incloses en el mateix espai estructural de l'article (el corresponent a l'exemple lexicogràfic) i amb la mateixa tipografia, però obeeixen a motivacions i a mecanismes de delimitació ben diferenciats.

No sempre es dona, però, aquesta correspondència entre les informacions que s'incorporen als diccionaris. Examinem un altre exemple del DIEC2:

emprendre *v. tr.* (...) Posar-se a fer, començar (una cosa, especialment que suposa un esforç considerable, que enclou dificultats, riscos, etc.). *Emprendre una feina, una tasca, una obra, un negoci, un viatge. Emprendre una campanya contra la implantació dels nous impostos. Emprendre una carrera. No gosar emprendre res.*

En aquest cas, el diccionari explicita un element extrínsec de naturalesa genèrica, i es fa especialment necessari d'incloure les estructures combinades en què intervé.

Per a aquest mateix verb, el CTILC dona les dades de coocurrència que es poden veure en la Taula 2 sobre *emprendre* seguit o precedit de substantiu.

<i>Lemes</i>	<i>Total</i>	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5
camí	452	11	12	24	23	6	33	248	54	30	11
viatge	218	4	6	4	9	3	14	126	39	7	6
tasca	159	5	10	10	4	4	3	93	21	7	2
marxa	123	1	1	5	98	6	9	3			
obra	102	3	6	10	7	5	5	49	9	6	2
vol	93	1	3	1	2	77	6	3			

Taula 2

camí *m.* (...) Tira de terreny més o menys ampla que va d'un indret a un altre disposada expressament per a transitar-hi. *Un camí que va de can Boquet a Sant Mateu. Per a anar a can Pibernat hi ha dos camins. És un camí molt dret: no fa gaires tortes. Hi ha un camí que fa dreuera. Un camí planer. Un camí rost, tirós. Un camí abandonat, debolit. Prendre un camí. Girar de camí. Prendre un camí per altre, errar-se de camí. Anar fora de camí, sense camí. Fer, obrir, un camí. Camí de desemboscar, bosquetà, bosquerol. Camí d'abeurada. Camí de sirga.*

viatge *m.* (...) Anada que hom fa per terra, per mar o per aire per transportar-se a un lloc notablement distant, a una altra ciutat, a un altre país. *Un viatge a Amèrica. Un viatge de plaer. Un viatge al voltant del món. Ha caigut malalt durant el viatge. Emprendre un viatge. El viatge de nuvis.*

tasca *f.* (...) Treball que hom té l'obligació de fer, que li han assenyalat o s'ha assenyalat. *Fer algú la seva tasca. Anar a la tasca. Posar-se a la tasca. Deixar la tasca. Les tasques d'una acadèmia. Acabar la tasca del dia.*

marxa *f.* (...) Moviment amb què la tropa, un seguici, una processó, etc., desfila en un cert ordre. (...) Moviment d'un cos de tropes que va d'un punt a un altre. *Un batalló en marxa.*

obra *f.* (...) Aplicació de l'activitat humana a un fi. *Posar-se a l'obra. Posar les seves virtuts en obra.*

Una primera constatació a partir de l'observació d'aquestes dades (deixant de banda el fet que el DIEC2 no és un diccionari basat en corpus) és que el diccionari té una certa voluntat de recollir aquest tipus de combinacions, dins d'un element estructural de l'article que no es diferencia dels exemples lexicogràfics. En el cas de les combinacions recollides a l'entrada verbal, la motivació és reflectir les restriccions lèxiques; en el cas dels substantius, les estructures del tipus VERB + SUBSTANTIU incloses són pròpiament col·locacions que activen determinades funcions lèxiques (*Prendre un camí, Girar de camí, Fer, obrir, un camí / Emprendre un viatge / Anar a la tasca. Posar-se a la tasca. Deixar la tasca, posar-se a l'obra, etc.*). Cal fer notar, però, que només una de les combinacions analitzades repeteix des de tots dos punts de vista.

Evidentment, en la consideració lexicogràfica d'aquestes qüestions cal tenir en compte diversos aspectes relatius al mètode de descripció lexicogràfica. Aquests aspectes són:

- a) La divisió en accepcions diferents del conjunt d'usos d'un mot, que determina estructuralment el tractament de les diferents combinacions: per exemple, en l'article de *camí* no es recull cap sentit "no físic" d'aquesta unitat, que és el que segurament s'activa mitjançant el verb *emprendre*.
- b) La *direccionalitat* oposada de les col·locacions i de les restriccions lèxiques: en l'expressió *emprendre un viatge*, *viatge* és una restricció lèxica del verb *emprendre*, mentre que des del punt de vista de *viatge*, *emprendre* és l'expressió de la seva activació verbal d'inici. Aquesta diferència, segons que es consideri un o altre constituent, és bàsica en un diccionari general, ja que cal tractar separatament cadascun dels dos elements. Des del punt de vista d'*emprendre* l'estructura és tractable com un element extrínsec de la definició. Des del punt de vista de *viatge*, es tracta de la determinació el seu verb d'activació.
- c) El tractament adequat de tota aquesta casuística no només requereix el concurs d'una anàlisi pertinent, sinó també de dades empíriques que assenyalin adequadament la importància que adquireixen les combinacions lèxiques de cada article de diccionari.

Com a conseqüència de tot el que s'ha dit, cal afegir un últim exemple que permeti mostrar que la «transcendència» de la col·locació *ignorància crassa* és diferent si la considerem des del punt de vista de la descripció del constituent *ignorància* (que no experimenta cap adaptació de significat) o del constituent *crassa* (que intensifica el significat d'*ignorància*).

ignorància *f.* (...) Estat de qui és ignorant. Una ignorància crassa. *La ignorància d'un fet.*

cras crassa *adj.* (...) Gros, s'aplica a *error, engany, ignorància.* *Viuen en una crassa ignorància: ignoren les coses més elementals.*

Com es pot veure en les dues entrades corresponents del DIEC2, ambdues recullen en l'apartat d'exemplificació la col·locació *ignorància crassa*, però el tractament definicional de la col·locació és sensiblement diferent, fins al punt que no es pot parlar de definició lexicogràfica pròpiament dita en el cas de l'adjectiu, sinó més aviat d'una fórmula que permet reflectir la seva funció merament intensificadora.

La distribució de la informació contextual: esbós de proposta d'aplicació

La revisió presentada fins al moment ha de permetre de dur a terme una reflexió que comporti una aplicació eficaç d'aquests criteris semàntics per tal de recollir de la manera més adequada possible la informació col·locativa que el lexicògraf pot extreure de les dades d'un corpus textual informatitzat. La informació que segueix vol ser un esbós de proposta metodològica que es pugui aplicar en una de les etapes clau en el procés de compleció d'un diccionari descriptiu.

El mot objecte de descripció lexicogràfica que guiarà aquesta proposta és *derogar*. Com en els casos anteriors, la informació que proporciona el DIEC2 dona peu a formular-nos les preguntes següents: sempre és *algú* qui *deroga* una *disposició* o una *lei* (informació sobre selecció-c)? La informació entre parèntesis de la definició (*disposició, lei*) fa referència a unitats lèxiques concretes o a unitats semàntiques que inclouen una classe semàntica? Com s'han

d'interpretar si, seguidament, l'exemple reprèn la unitat *lleï*? És necessari que sempre aparegui en context un nou «instrument jurídic» (*nova disposició* o *lleï*) per tal que algú pugui *derogar*?

derogar v. tr. (...) Revocar (una disposició o una llei) amb una nova disposició o llei.
Derogar una llei.

lleï f. (...) Norma jurídica de caràcter general i obligatori, dictada pels òrgans que tenen atribuït el poder legislatiu. *Lleis polítiques, civils, criminals, penals. Establir una llei. Derogar una llei.*

Aquestes qüestions prèvies no semblen tenir una resposta clara a la llum de la informació que conté cada un dels elements estructurals del diccionari. Efectivament, la funció descodificadora, a partir de la definició i de l'exemple, queda acomplerta, però no es pot dir el mateix de la funció codificadora perquè sembla que l'usuari del diccionari no serà capaç de deduir si també és pot *derogar*, per exemple, un *article*. Tanmateix, destaca positivament, en aquest cas, el compliment de la bidireccionalitat de la informació recollida en el diccionari, que contribuirà en certa mesura a la funció codificadora.

Des de l'experiència lexicogràfica del treball basat en corpus proposem, per tal de donar compte de manera exhaustiva però sense repetir informació en els diferents elements estructurals del diccionari i, sobretot, amb la voluntat de no deixar informació encoberta per a l'usuari, distribuir la informació de la manera següent. A partir de la cerca de coocurrències, s'observa que la selecció argumental d'una unitat com *derogar* està formada, pel que fa a l'objecte directe, per: *lleï, decret, constitució, article, carta, disposició*. I pel que fa a les unitats que poden fer de subjecte del verb apareixen (posicions -1 i -2): *lleï, decret, constitució* però també *república, reialesa, faraó, santedat* (Taula 3).

Lemes	Total	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5
lleï	20				5	1	7	3	4		
decret	12				2	1	7	1	1		
constitució	8					1	5	2			
article	5						4	1			
any	4						2				2
carta	3							1	1	1	
disposició	3							3			
alcorà	1				1						
ordre	1				1						
judici	1				1						
república					1						
reialesa					1						
potestat					1						
faraó					1						
santedat						1					

Taula 3

Com a verb transitiu que és, *derogar* presenta un patró d'inserció sintàctica representable mitjançant la fórmula [N1 V N2]. Fent un procés d'abstracció i generalització a partir de l'anàlisi de corpus, el lexicògraf pot arribar a establir que la classe semàntica del constituent N2 (és a dir, de les entitats susceptibles d'ésser *derogades*) és, per exemple, i entesa com a etiqueta (selecció-s), la de *text legal* (que inclou *disposició*, *lleï*, *constitució*, *article*, etc.). En canvi, pel que fa al constituent N1, el lexicògraf detecta que aquest argument pot estar ocupat per un humà (subspecificat com a [+humà amb *autoritat*]) o bé per un element idèntic a N2, és a dir, un *text legal* que en *deroga* un altre (selecció-c). Aquesta informació es confirma en la mostra contextual següent:

1054:0:23:26:8 segueix: © Per a obviar tots aquests inconvenients fou **derogat** l'any 1830 el sistema tradicional d'arqueig basat en la
 1815:0:112:10:2 Fra Menors, que de cap manera havia volgut Sa Santetat **derogar** l'article 245 de les Constitucions Generals de l'Ordre:
 1758:0:16:34:5 resulta, per les grans arteries fluvials, l'obligació de **derogar** sovint les condicions propies a quiscuna de les regions
 369:0:9:20:3 establia l'article 290 de l'Estatut que la República va **derogar** per decret de 16 de juny de 1931. VI De la competència
 1431:0:24:14:9 i d'un cost molt menor. Un Decret-Llei del dia 16 de juny **deroga** el Decret del 1930 que igualava els sous dels militars
 2017:1:65:33:4 Respecte a les reformes militars deuria tal Govern **derogar** el Decret, convertint-lo en projecte de llei per a
 327:0:50:16:4 públic, i nega a l'actuació privada no sols la virtut de **derogar** la llei escrita, sinó àdhuc la de fer-la evolucionar; és a
 2844:0:36:16:2 millor negoci de Paradela. Ja ho crec! En Ciriac Curriola **derogà** les lleis constitucionals. Implantà la tirania. I Paradela
 1171:0:66:30:6 vigent des d'abril de 1976. La Llei de Reformes Laborals **deroga** les normes que prescrivien la necessitat d'autorització
 1167:0:155:34:3 manada fer pel veguer i el batlle. Aquesta redacció **derogà** totes les ordinacions anteriors sobre assegurances
 2013:0:8:30:3 á las Senyoras Infantas de la successió á la Corona está **derogát** per la Pragmática Sanció de 1789, que es la última llei que
 2013:0:30:1:6 de horrores y de catástrofes espantosas, un Real decret que **derogás** la Pragmática Sanció: si aixó es tan cert com se ha
 1183:2:13:19:2 que incideix directament en algunes matèries sense **derogar** els preceptes d'aquestes que resulten afectades, com és el
 2013:0:26:2:2 ho exijís; ab tanta major rahó si la llei promulgada **deroga** un Reglament informal, perjudicial, y nullo. Pero encára en
 2083:0:29:37:8 "No 's podia pas dubtar, diu Titus Livi, que la llei fos **derogada** per unanimitat y aixís ho fou al cap de vint anys de sa

Finalment, la proposta de distribució de la informació obtinguda a través del corpus per al patró sintàctic més senzill, a partir de l'anàlisi de les dades, és:

derogar v. [N1 V N2] (N1=autoritat; N2=text legal) [Un càrrec, una autoritat, un període polític]₁ revocar [un text legal]₂ amb un altre de nou. *Derogar una llei, derogar un decret, derogar la constitució, derogar un article.*

Conclusió

L'acompliment de les funcions codificadores i l'adequació descriptiva d'un diccionari general exigeix un tractament adequat dels diferents tipus de fenòmens lligats al significat contextual i a les restriccions lèxiques. Ara com ara, i malgrat que la metalexicografia ha elaborat mecanismes de definició que poden tractar satisfactòriament alguns d'aquests fenòmens, resta pendent la qüestió de com han de ser incorporades les col·locacions, sobretot quan se les considera des del punt de vista de la seva base lèxica (el mot determinat). En aquest sentit, la consideració de les col·locacions com un element estructural propi del diccionari esdevé una necessitat objectiva per a la millora descriptiva de les obres lexicogràfiques.

Per últim, una bona descripció i caracterització dels diferents elements estructurals d'un article de diccionari ha de permetre seleccionar adequadament la informació que proporcionen els corpus textuais. En aquest sentit, les dades proporcionen informació semàntica i, també, informació col·locativa, essencials per a la descripció d'un mot. En aquesta contribució hem intentat establir les línies mestres d'una possible distribució de la informació contextual entre el camp de la definició, el camp assignat a col·locacions i l'apartat dels exemples en un article de diccionari. Creiem que una visió integradora del contingut que ha de recollir cada un d'aquests elements ajudarà a obtenir una major sistematització de les dades que, al seu torn, esdevindrà essencial per als processos de codificació i de descodificació duts a terme per l'usuari del diccionari.

Referències bibliogràfiques

- Alonso, M. (2002). "Colocaciones y contorno en la definición lexicográfica". *Lingüística española actual* XXIV (1). 63-96.
- Benson, M.; Benson, E.; Ilson, R. (1986). *Lexicographic description of English*. Països Baixos: John Benjamins Publishing Company.
- Bosque, I. (2004). "La direccionalidad en los diccionarios combinatorios y el problema de la selección léxica". Dins Cabré i Monné, T. (ed.). *Lingüística teórica: anàlisi i perspectives* 1. Bellaterra: Universitat Autònoma de Barcelona. 13-58.
- Church, K. W.; Hanks, P. (1990). "Word Association Norms, Mutual Information and Lexicography". *Computational Linguistics* 16. 22-29.
- Coseriu, E. (1977). *Principios de semántica estructural*. Madrid: Gredos.
- Firth, J. R. (1951). "Modes of meaning". *Essays and Studies* 4. 118-149.
- Fontenelle, T. (1997). *Turning a bilingual dictionary into a lexical-semantic database*. Tübingen: Max Niemeyer.
- Katz, J. J.; Fodor, J. A. (1963). "The Structure of a Semantic Theory". *Language* 39 (2). 170-210.
- Kilgarriff, A.; Tugwell, D. (2001). "WORD SKETCH: Extraction and Display of Significant Collocations for Lexicography". *Proc. ACL workshop on COLLOCATION: Computational Extraction, Analysis and Exploitation*. Toulouse. 32-38.
- Koike, K. (2001). *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Alcalá de Henares: Universidad de Alcalá de Henares.
- Kroman, H.; Riiber, T.; Rosbach, P. (1984). "Active and passive bilingual dictionaries". *LEXter'83 Proceedings. Papers from the International Conference on Lexicography*. Tübingen: Max Niemeyer. 207-215.
- Mel'čuk, I. (1998). "Collocations and Lexical Functions". Dins Cowie, A. *Phraseology. Theory, analysis and Applications*. Oxford: Oxford University Press. 23-53.
- Muñiz, E. M. (2005). "Notas para unha análise comparativa das nocións de *restricción léxica e colocación*". *Cadernos de Fraseoloxía Galega* 7. 177-190.
- Rafel, J. (1989). "Sobre la normalització dels diccionaris catalans: elements intrínsecs i extrínsecs en les definicions lexicogràfiques". Dins *La Corona d'Aragó i les llengües romàniques. Miscel·lània d'homenatge per a Germà Colon*. Tubinga: Gunter Narr.
- Rafel, J. (2006). "Elements extrínsecs en les definicions lexicogràfiques: teoria i aplicació". Bernal, E.; DeCesaris, J. (ed.). *Palabra por palabra. Estudios ofrecidos a Paz Battaner*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra. 201-217.
- Rey-Debove, J. (1971). *Étude linguistique et sémiotique des dictionnaires français contemporains*. The Hague: Mouton.
- Seco, M. (1979). "El 'contorno' en la definición lexicográfica". Dins *Homenaje a Samuel Gili Gaya (in memoriam)*. Barcelona. 183-191.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Soler, J. (2006). *Definició lexicogràfica i estructura del diccionari*. Barcelona: Institut d'Estudis Catalans.

Diccionaris

- [BBI]. Benson, M.; Benson, E.; Ilson, R. (1986). *The BBI Combinatory Edictionary of English. A Guide to Word Combinations*. Amsterdam: John Benjamins Publishing Company.
- [DEC]. Mel'čuk, I. et al. (1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*. Montréal: Les Presses de l'Université de Montréal. Vol. I. (1984); vol. II (1988); vol. III (1992); vol. IV.
- [DICE]. *Diccionario de colocaciones del espanyol*. Projecte dirigit per Margarita Alonso i en curs de desenvolupament a la Universitat de A Coruña.

[DIEC2]. *Diccionari de la llengua catalana*. Barcelona: Institut d'Estudis Catalans, 2007.

[REDES]. Bosque, I. (dir.) (2004). *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: Ediciones SM.