

DEFINICIÓ LEXICOGRÀFICA
I ESTRUCTURA
DEL DICIONARI

Tesi doctoral

Universitat de Barcelona

Departament de Filologia Catalana

(Secció de Lingüística)

Programa de doctorat: *La normativa: teoria i pràctica*

Bienni 1989-1991

Joan Soler i Bou

DEFINICIÓ LEXICOGRÀFICA
I ESTRUCTURA
DEL DICCIONARI

Tesi doctoral

per optar al títol de doctor en Filologia Catalana

Director

Joaquim Rafel i Fontanals

Barcelona, 2003

TAULA GENERAL

AGRAÏMENTS.....	1
1 INTRODUCCIÓ.....	3
1.1 La lexicografia: canvis recents.....	3
1.2 Enfocament del treball	10
1.3 Objectius i hipòtesis del treball.....	15
1.4 Estructura del treball	18
2 ELS DICCIONARIS I EL MODEL DEL LÈXIC.....	23
2.1 Del lexema a l'entrada de diccionari.....	23
2.2 Delimitació formal de l'entrada de diccionari	31
2.3 Forma gràfica de l'entrada	35
2.4 Categorització gramatical	40
2.5 Paradigma de flexió	48
2.6 Entrades hipolexemàtiques	54
2.6.1 Clítics	55
2.6.2 Afixos lèxics.....	57
2.6.3 Irregularitats flexionals.....	59
2.6.4 Marques de flexió quasilèxiques	61
2.6.5 Categories de semantemes.....	62
2.7 Relacions entre macroestructura i microestructura	62
3 ELS MODELS DEL LÈXIC I EL DICCIONARI.....	73
3.1 Creativitat lèxica i diccionaris.....	73
3.2 El diccionari i els lèxics electrònics	80
3.3 Un model de diccionari teòric: les funcions lèxiques	88
3.4 Un model de lèxic basat en les relacions semàntiques: WordNet.....	98
3.5 Representació de sentits relacionats: el lèxic generatiu	103
4 DEFINICIÓ LEXICOGRÀFICA I EQUIVALÈNCIA.....	113
4.1 Equivalència i identitat.....	113
4.2 Substituïbilitat definit-definició	124
4.3 Equivalència argumental.....	127
4.4 Equivalència actancial.....	142
4.5 Equivalència esdevenimental.....	162
4.6 Equivalència sincategoremàtica	175
5 CONDICIONS SOBRE LA DESCRIPCIÓ LEXICOGRÀFICA	187
5.1 Desviacions estructurals del vocabulari definidor: la circularitat	187
5.2 Desviacions metodològiques del vocabulari definidor: la redundància.....	202
5.3 Un exemple de vocabulari definidor: els vocabularis controlats	214
5.4 El vocabulari definidor i l'estructura dels diccionaris	224
5.5 Condicions per a l'establiment d'un vocabulari definidor	229
5.6 Un vocabulari definidor per al català.....	235
6 VOCABULARI DEFINIDOR I SISTEMA DE DEFINICIÓ	241
6.1 D'un vocabulari definidor a una metallengua de definició	241
6.2 Els elements primaris d'un sistema de definició.....	251
6.3 Les condicions d'un sistema de definició	262
7 UNA METALLENQUA DE DEFINICIÓ PER AL CATALÀ	267
7.1 El vocabulari proposat com a punt de partida.....	267
7.2 Processament del vocabulari de partida	273
7.3 Eliminació d'articles innecessaris	280

7.4	Selecció de sentits	283
7.5	Regularització de les definicions.....	287
7.6	Integració en un sistema de definició	291
8	METALLENGUA I SISTEMA DE DEFINICIÓ.....	299
8.1	Estructura i naturalesa de la metallengua	299
8.2	Elements complementaris de la metallengua	303
8.3	Condicions de descripció i nivells de descripció.....	307
8.4	Els nivells de descripció en la redacció lexicogràfica	312
8.5	Desambiguació de la metallengua i forma del diccionari	319
8.6	Un cas pràctic de redacció d'un article	323
9	CONCLUSIONS.....	335
	BIBLIOGRAFIA	341
	Obres de consulta	341
	Sigles i referències de diccionaris	369
	Altres sigles utilitzades en la redacció del treball.....	371
	APÈNDIXS	373
	APÈNDIX I	375
	I.1 Llistat A: lemes de freqüència més alta al DFDCC	377
	I.2 Llistat B: lemes de suma d'usos més alta al DFDCC	417
	I.3 Llistat C: lemes de freqüència més alta al DIEC	457
	I.4 Llistat D: lemes comuns a A i B	483
	I.5 Llistat D': lemes exclusius A	519
	I.6 Llistat D'': lemes exclusius B.....	523
	I.7 Llistat E: lemes comuns a C i D.....	527
	I.8 Llistat E': lemes comuns a C i D'	543
	I.9 Llistat E'': lemes comuns a C i D''	545
	I.10 Vocabulari de partida	547
	APÈNDIX II	565
	Vocabulari definidor de nivell 0.....	567

TAULA DE FIGURES

Figura 1.1. Activitats relacionades amb la descripció lexicogràfica.	10
Figura 2.1. Llista de codis gramaticals de lema amb la freqüència acumulada.....	28
Figura 2.2. Llista de categories amb ocurrències, lemes i relació ocurrències/lema.	29
Figura 2.3. Llista de categories lèxiques de primer nivell del DIEC.....	43
Figura 2.4. Llista de codis de segon nivell del DIEC.	44
Figura 2.5. Llista de categories d'EAGLES	45
Figura 3.1. Diferències entre els diccionaris convencionals i els lèxics electrònics.	85
Figura 3.2. Classificació dels noms a <i>WordNet</i> extreta de G. A. Miller (1998b).....	100
Figura 3.3. Representació convencional de les relacions jeràrquiques del lèxic.....	107
Figura 3.4. Representació multidimensional de les relacions semàntiques del lèxic.	109
Figura 3.5. Representació de les relacions del <i>Constitutive role</i> en el model de <i>SIMPLE</i>	111
Figura 4.1. Classificació componencial de les definicions de Martin (1983).	121
Figura 4.2. Proposta general de classificació de les definicions segons Martin (1983).	122
Figura 4.3. Mostra de concordances d' <i>amortallar+amb</i> i d' <i>ensucrar + amb</i>	131
Figura 4.4. Casos específics d' <i>engiponar</i> i d' <i>amonestar</i>	133
Figura 4.5. Mostra de concordances de <i>rodolar+per</i>	135
Figura 4.6. Mostra de concordances d' <i>incloure+dins/dintre</i>	139
Figura 4.7. Mostra de concordances d' <i>intervenir+en</i>	140
Figura 4.8. Classes aspectuals de Vendler (1967).	167
Figura 4.9. Adaptació al català del test de Dowty (1979) extreta de Marín (1999).	168
Figura 4.10. Tipus d'aspectualitat segons Miguel (1999).	169
Figura 4.11. Informacions lexicogràfiques del verb <i>dormir</i> i derivats.	171
Figura 4.12. Anàlisi aspectual del verb <i>dormir</i> i derivats.	172
Figura 5.1. Representació de la circularitat en els substantius abstractes.	196
Figura 5.2. Representació de la circularitat en la sèrie relacionada amb <i>calor</i>	200
Figura 5.3. Jerarquia dels mots relacionats amb sentits corporals.....	209
Figura 5.4. Representació de la sèrie dels sentits corporals (DIEC).	210
Figura 5.5. Representació de la sèrie dels sentits corporals amb vocabulari controlat.....	212
Figura 5.6. Mostra comparativa entre els vocabularis definidors del LDOCE i del CIDE	220
Figura 5.7. Rang dels lemes del CTILC.	233
Figura 7.1. Procés d'obtenció del vocabulari definidor.....	275
Figura 7.2. Representació d'un diccionari basat en vocabulari definidor controlat.	294
Figura 7.3. Representació d'un diccionari basat en un vocabulari definidor lliure.	295

Figura 7.4. Representació d'un diccionari basat en un sistema de definició estructurat en nivells de descripció.	297
Figura 8.1. Procés de redacció lexicogràfica.	317
Figura 8.2. Representació de la circularitat entre els elements del nivell 0.	324
Figura 8.3. Representació per nivells de la descripció de <i>teatre</i> (1).	328
Figura 8.4. Representació per nivells de la descripció de <i>teatre</i> (2).	329
Figura 8.5. Representació per nivells de la descripció de <i>teatre</i> (3).	330
Figura 8.6. Representació per nivells de la descripció de <i>teatre</i> (4).	331

Agraïments

La redacció d'aquest treball no hauria estat possible sense l'ajut i la col·laboració d'un bon nombre de col·legues que hi han contribuït amb aportacions molt valuoses. Agraïco la lectura de l'esborrany a Victòria Alsina, de la Universitat Pompeu Fabra, i a Josep Maria Domènech, de l'Institut d'Estudis Catalans. Tots dos van portar a terme una lectura acurada, atenta i pacient, i van fer-me un bon nombre de suggeriments de millora. Dec a Teresa Sadurní, de l'Institut d'Estudis Catalans, la resolució informàtica de diverses qüestions de processament de les dades que he utilitzat i produït, en especial el tractament de les dades obtingudes del vocabulari definidor del DIEC i les validacions estructurals de la metallengua de definició de nivell 0 que proposo. Dec a Roser Sanromà, de l'Institut d'Estudis Catalans, l'obtenció del conjunt de definicions del DIEC per al tractament posterior que n'he fet, i el seu ajut en l'establiment dels models de flexió per a la classificació de les unitats del nivell 0.

Agraïco també a Lluís Padró, de la Universitat Politècnica de Catalunya, l'etiquetatge morfosintàctic de les definicions del DIEC que m'ha permès aprofitar els seus valuosos treballs en el terreny del processament de corpus textuais. Per altra banda, les converses que vaig mantenir amb Joan Gómez, de la Universitat Politècnica de Catalunya, em van resultar molt útils per al plantejament de certes formulacions. Els consells professionals d'Antoni Escolà, de l'Institut d'Estudis Catalans, també m'han estat molt valuoses per a les qüestions relatives al format de presentació del treball.

També vull agrair d'una manera molt especial les aportacions, la paciència i la dedicació amb què Joaquim Rafel ha distingit aquest treball. No exagero gens ni mica si dic que la seva participació ha anat molt més enllà de la que s'espera d'un bon director de tesi.

No cal dir que totes les persones que he esmentat han fet contribucions decisives a través d'indicacions, esmenes, suggeriments, i col·laboracions en aspectes particulars. Per tant, tots ells han contribuït a potenciar els possibles aspectes positius d'aquest treball, mentre que les mancances, inexplicituds, i els errors que el lector hi trobi són exclusivament responsabilitat meva.

Cal també, finalment, que expressi la meva gratitud a l'Institut d'Estudis Catalans, i d'una manera especial a Joaquim Rafel, com a director del projecte "Diccionari del català contemporani" (DCC), per permetre'm d'accedir a dades i recursos lingüístics d'ús intern constituïts en el marc del DCC, sense el concurs dels quals la recerca que he portat a terme no hauria estat possible.

1 Introducció

1.1 La lexicografia: canvis recents

La lexicografia pot ser definida genèricament com la disciplina que té com a objecte principal d'anàlisi i de descripció el lèxic d'una llengua en els seus diferents aspectes: sintagmàtics o paradigmàtics, qualitatiu o quantitatiu. D'un temps cap aquí, la tècnica lexicogràfica ha experimentat canvis en els seus mètodes de treball que han motivat que els diccionaris siguin, ells mateixos, un objecte d'estudi teòric. Neix així la *lexicografia teòrica* (o *metalexicografia*) que en els darrers anys ha representat un corrent notable d'estudis sobre els diccionaris i la seva sistematització.

Una de les causes indiscutibles del progrés de la lexicografia com a disciplina és l'aparició d'una lingüística científica i formalitzada que es produeix en l'àmbit mundial a partir de l'aparició dels corrents estructuralistes. La definitiva consolidació i generalització d'aquests corrents, que es pot situar a mitjan segle XX, coincideix amb un període de gran projecció de la lexicografia tant pràctica com teòrica (Quemada, 1990).

Com a disciplina aplicada de la lingüística, la lexicografia presenta nombrosos punts de contacte amb la lingüística teòrica. Totes les qüestions que la lingüística es planteja pel que fa a l'explicitació, descripció i formalització del component lèxic són susceptibles de rebre un tractament més o menys específic en la descripció lexicogràfica. Així, la lingüística pot afavorir o accentuar tendències que modifiquen l'anàlisi que els diccionaris fan del lèxic; el resultat d'aquest procés és l'evolució de l'estructura i dels mecanismes de descripció lexicogràfica cap a models de diccionari més rigorosos i fonamentats lingüísticament.

Al costat d'aquesta influència de la lingüística com a factor de progrés metodològic, es pot assenyalar també un altre element decisiu que ha contribuït a configurar la lexicografia actual i que sens dubte contribuirà a determinar la lexicografia dels pròxims anys: la incorporació dels mitjans informàtics a totes les fases del procés d'elaboració dels diccionaris.

La utilització dels nous mitjans tecnològics ha provocat la introducció d'un nou concepte que no s'havia considerat mai en relació amb els diccionaris convencionals, i que és un important factor d'evolució tècnica: la *multifuncionalitat* dels diccionaris com a recursos lingüístics. El concepte de multifuncionalitat (Zampolli, 1991a, 1991b) s'ha formulat per a les bases de dades lèxiques, i fa referència a la diversitat d'aplicacions a què poden ser destinades. D'una manera cada cop més evident, el procés d'elaboració de diccionaris presenta grans similituds amb el procés de constitució d'aquests tipus de bases de dades o de lèxics electrònics. Per multifuncionalitat aplicada a la lexicografia entenc la tendència cada cop més consolidada a concebre l'elaboració de diccionaris com un procés de constitució de bases de dades orientades a la caracterització integral del lèxic (en els seus aspectes sintagmàtic, semàntic, morfològic, freqüencial, etc.), que condueix a l'elaboració de diferents tipus de recursos en tant que producte final.

Dit d'una altra manera, cada cop més l'activitat lexicogràfica es dirigeix cap a la producció de recursos que permeten diferents presentacions finals i diferents possibilitats d'utilització o de reutilització. L'aparició de les noves tècniques en l'elaboració i la presentació dels diccionaris no s'ha d'entendre com la mera utilització de mitjans més adequats i eficients per a la realització de les tasques lexicogràfiques més dures i pesades (habitualment introductores de manca de sistematicitat en els diccionaris). Molt més enllà d'això, la introducció dels nous mitjans representa un veritable canvi metodològic, amb grans repercussions en el procés d'elaboració de diccionaris i en la presentació final de les dades lexicogràfiques.

Considerats des d'un punt de vista global, els canvis suscitats pel desenvolupament tecnològic es poden resumir en els següents punts:

- a) Utilització dels corpus com a material de referència en la descripció lexicogràfica (reducció de la introspecció, augment del rigor descriptiu).
- b) Integració dels mitjans informàtics en el procés d'elaboració del diccionari (constitució de bases de dades lèxiques, mecanismes de validació, diferents plataformes de publicació).
- c) Possibilitats de reutilització eficient de dades lexicogràfiques ja constituïdes (explotació de diccionaris en línia, constitució de bases de dades lexicogràfiques, etc.)

Des d'un cert punt de vista, aquestes possibilitats —especialment la *b*)— permeten de concebre el procés d'elaboració d'un diccionari d'una forma que no és funció exclusiva de l'obra publicada, del diccionari com a producte. La influència d'aquests factors en conjunt és tan gran que no sols representa canvis en els diccionaris, sinó en la mateixa lexicografia teòrica, que haurà de ser cada cop més sensible a aquest panorama canviant. Fins ara, les diverses tipologies de diccionaris a què ha donat lloc la lexicografia teòrica —com per exemple la de Ščerba (veg. Farina, 1995) o les de Malkiel (1967), Rey (1970*b*) o Zgusta (1971)— es refereixen als diccionaris en tant que objectes finals, productes comercials acabats que corresponen a l'aplicació d'un mètode particular de selecció de dades, de criteris d'organització estructural, de tècniques de redacció, etc. Aquestes tipologies, però, no resulten adequades per a la caracterització de determinats projectes lexicogràfics actuals que s'identifiquen amb el corrent més avançat de la lexicografia, que es basen en la constitució de bases de dades multifuncionals de les quals es poden obtenir diferents tipus de diccionaris.

Distincions que fins ara han resultat un punt de referència per a la lexicografia teòrica, com ara la que estableix diferències entre diccionaris *semasiològics* i *onomasiològics* (veg. per exemple Svensén, 1993) són poc operatives si es consideren des del punt de vista a què em refereixo. El factor de distinció bàsic entre aquests dos tipus d'obres assenyalats recau en el fet que el consultant del diccionari parteix d'una forma per arribar a localitzar la seva descripció lingüística (diccionaris semasiològics), o bé parteix d'una informació semàntica considerada com a contingut per arribar a les *formes*

que l'actualitzen (diccionaris onomasiològics). Així doncs, no és la diferència entre l'objecte de descripció i d'anàlisi, sinó més aviat la diferència en la relació entre els tipus d'*accés* a la informació allò que permet establir aquesta tipologia. Plantejat en termes d'estructura de dades, l'*accés* no implica en si cap modalitat estructural concreta, cap disposició específica de les dades, sinó únicament el procediment a partir del qual recuperem informació en un format de resultats variable en funció de les necessitats específiques de cada consulta.

També s'ha assenyalat, especialment en l'àmbit de la lexicografia bilingüe, una qüestió relacionada més o menys indirectament amb les possibilitats d'accés a la informació dels diccionaris: es tracta de la doble funció codificadora i descodificadora dels diccionaris. En aquest terreny sovint s'estableix la distinció (Kroman & al., 1984) entre diccionaris *actius* (que, genèricament, van del sentit a la forma) i *passius* (de la forma al sentit). Aquesta distinció, però, no és exclusiva de la lexicografia bilingüe. És freqüent pensar (Tazawa, 1998) que un diccionari bilingüe que té com a objectiu principal l'ús actiu (codificador) de la llengua d'arribada ha de contenir forçosament molta més informació (sintagmàtica, gramatical, etc.) que un diccionari orientat a l'ús passiu (descodificador). La diferenciació objectiva de tots dos tipus d'obres recau més en raons de tipus pràctic (les necessitats prioritàries del diccionari o la seva viabilitat comercial) que en la metodologia de descripció lexicogràfica.

En l'àmbit de la lexicografia monolingüe es poden trobar també elements estructurals del diccionari que es relacionen amb una o altra funció. Fins i tot es pot considerar l'existència de diccionaris monolingües específicament orientats a la codificació —a l'ús actiu, per tant— com ara el COBUILD, el LDOCE, el CIDE i l'OALD.

Des d'una perspectiva actual que no resulta totalment coincident amb les anàlisis que s'han fet d'aquesta qüestió (Kroman & al., 1984; Tazawa, 1998), i tenint en compte els canvis a què em refereixo en aquesta secció, entenc que els diccionaris generals monolingües han de tenir un caràcter bifuncional. En termes generals, no es pot associar de manera exclusiva l'ús d'un diccionari a les tasques de codificació o de

descodificació. És cert que, en general, un parlant nadiu d'una llengua qualsevol usarà un diccionari per a la resolució de dubtes d'interpretació concrets que es refereixen a significats que ignora, mentre que un aprenent estranger l'utilitzarà en les seves tasques de construcció de textos, amb la finalitat de saber si tal o tal verb va amb tal o tal preposició, o per determinar una col·locació qualsevol; també és cert, però, que aquest ús està en funció de l'activitat a què dona suport un diccionari, i és precisament aquesta activitat la que determina el caràcter actiu o passiu d'aquesta mena d'obres.

Aquesta raó pràctica pot, encara, complementar-se amb arguments de tipus teòric. En semàntica no és cap idea nova el fet de pensar que l'estudi del significat lèxic no es pot entendre d'altra manera que a través de l'estudi dels contextos possibles d'aparició¹ o, dit d'una altra manera, el coneixement del significat d'un mot pot identificar-se amb el coneixement de l'ús que aquest mot té en els seus diferents contextos. La descripció d'un lexema, doncs, ha de presentar de manera estructurada: *a)* totes aquelles informacions que permeten la identificació del seu significat en un context determinat (descodificació), i *b)* les informacions sobre les seves possibilitats d'aparèixer en un context determinat i les característiques generals d'aquesta inserció (codificació).

En la seva formulació moderna, aquesta línia de pensament penetra en la semàntica filosòfica a partir dels treballs de Wittgenstein (veg. De Mauro, 1967), que identifica el

¹ Una de les formulacions més clares i precises del concepte de significat contextual en lingüística prové de Hjelmslev: "Ninguna de las entidades mínimas, ni los radicales, tiene existencia tan «independiente» que le pueda ser asignada una significación léxica. Pero desde el punto de vista básico que hemos adoptado —el análisis continuado sobre la base de las funciones que aparecen en el texto— no existen otras significaciones perceptibles que las contextuales; toda entidad, y por tanto todo signo, se define con carácter relativo, no absoluto, y sólo por el lugar que ocupa en el contexto. Desde este punto de vista carece de significado la distinción entre significaciones que aparecen solamente en el contexto y significaciones de las que podría suponerse que tienen existencia independiente, o —de acuerdo con los antiguos gramáticos chinos— entre palabras «vacías» y «llenas». Las llamadas significaciones léxicas de ciertos signos no son sino significaciones contextuales artificialmente aisladas, o paráfrasis artificiales de las mismas. Totalmente aislado, ningún signo tiene significación; toda significación del signo surge en el contexto, entendiendo por tal un contexto situacional o un contexto explícito." [Hjelmslev (1943: 69-70)]

significat dels elements lèxics amb el seu ús. Wittgenstein elimina de la seva concepció del significat tota referència externa al propi llenguatge: si els éssers humans són capaços d'aprendre una nova llengua a partir de definicions ostensives és perquè, d'alguna manera, ja tenen desenvolupada la facultat del llenguatge. El significat d'un mot, doncs, es defineix en relació al paper que el mot desenvolupa en el llenguatge; així, el significat no és la referència a l'objecte designat pels mots, sinó el conjunt de les propietats que determinen l'ús dels mots (veg. Brand, 1975).

Aquests principis han estat considerats (i assumits amb més o menys convicció segons el tractadista) com a element de reflexió per la semàntica de base lingüística. Sembla clar que un dels avantatges més destacats d'aquesta interpretació és l'adequació descriptiva que presenta, i que es fa palesa en formulacions com la d'Allan (1986):

it should be obvious that the assignment of meaning by ordinary use is philogenetically and ontogenetically prior to defined meaning: typically a child acquires the meanings of expressions from the way they are used by the people around him, and through other people's responses to his own use of them. On this evidence there seems little doubt that the meaning of a word is governed by its use. [Allan (1986: v. 1, 77)]

Una altra qüestió és, però, el fet que la identificació total entre significat i ús sigui teòricament defensable.² Allò que es tendeix a admetre universalment en la semàntica moderna és que l'ús d'un mot és a la base del seu significat i, per tant, tota descripció semàntica que tingui per objectiu la definició d'un significat conceptual ha de partir de les circumstàncies que determinen l'ús dels mots.

² El mateix Allan es mostra contrari a aquesta identificació: "To say that the meaning of a word (...) is determined by its use, is not to say that the meaning of a word IS its use. As we noted earlier, the meaning of a word can be given by definition before the word is used —though it is usage which will determine whether that definition stands or falls. In any case, it is absurd to seriously hold that the meaning of a word is equivalent to its use just because the meaning is determined by its use (...). We therefore conclude that Wittgenstein was wrong when he said that «the meaning of a word is its use in the language»." [Allan (1986:I, 78-79)]

Per a la lexicografia, aquesta assumptió representa l'establiment d'una correspondència directa entre el significat d'un lexema i la totalitat dels seus contextos d'aparició. La importància que han adquirit avui els corpus de referència en la descripció lexicogràfica es deu, almenys en part, a la rellevància que la semàntica dels darrers anys ha concedit a l'estudi dels contextos lingüístics.

L'article lexicogràfic, i la mateixa descripció lexicogràfica del significat, doncs, deixen de partir de la idea de la conceptualització absoluta dels sentits d'un mot a partir de la reunió d'enunciats descriptius. Els diccionaris actuals, d'una manera progressivament més notable, presten una importància decisiva al vessant sintagmàtic del mot. La importància que s'ha donat en els darrers anys als usos col·locatius dels mots, a llurs restriccions selectives, a l'expressió dels règims verbals, permet de constatar l'existència d'una tendència irreversible en aquest sentit.

Finalment, crec que val la pena de destacar, en aquest resum general dels canvis essencials experimentats per la lexicografia, el fet que aquesta disciplina ha hagut de diversificar els seus processos de descripció del lèxic i no limitar-los exclusivament als usuaris humans. Els lèxics electrònics tenen com a finalitat principal dotar els processos computacionals (d'anàlisi, reconeixement i generació d'estructures) de la informació corresponent als diferents aspectes que ofereix el lèxic. Una de les conseqüències fonamentals d'aquesta diversificació és la interacció que es produeix entre la lexicografia concebuda per a ús humà i la reutilització de les seves dades per al desenvolupament d'aquests nous sistemes d'accés, de tractament i de processament d'informació. Aquesta interacció és un dels elements més destacables per a la formalització i la normalització de la lexicografia actual.

Totes les relacions que he exposat sintèticament en aquesta secció poden veure's representades gràficament a la figura 1.1, que reflecteix els diferents components que intervenen en el treball lexicogràfic actual i la complexitat de les implicacions (bidireccionals o unidireccionals) que aquests components mantenen entre si.

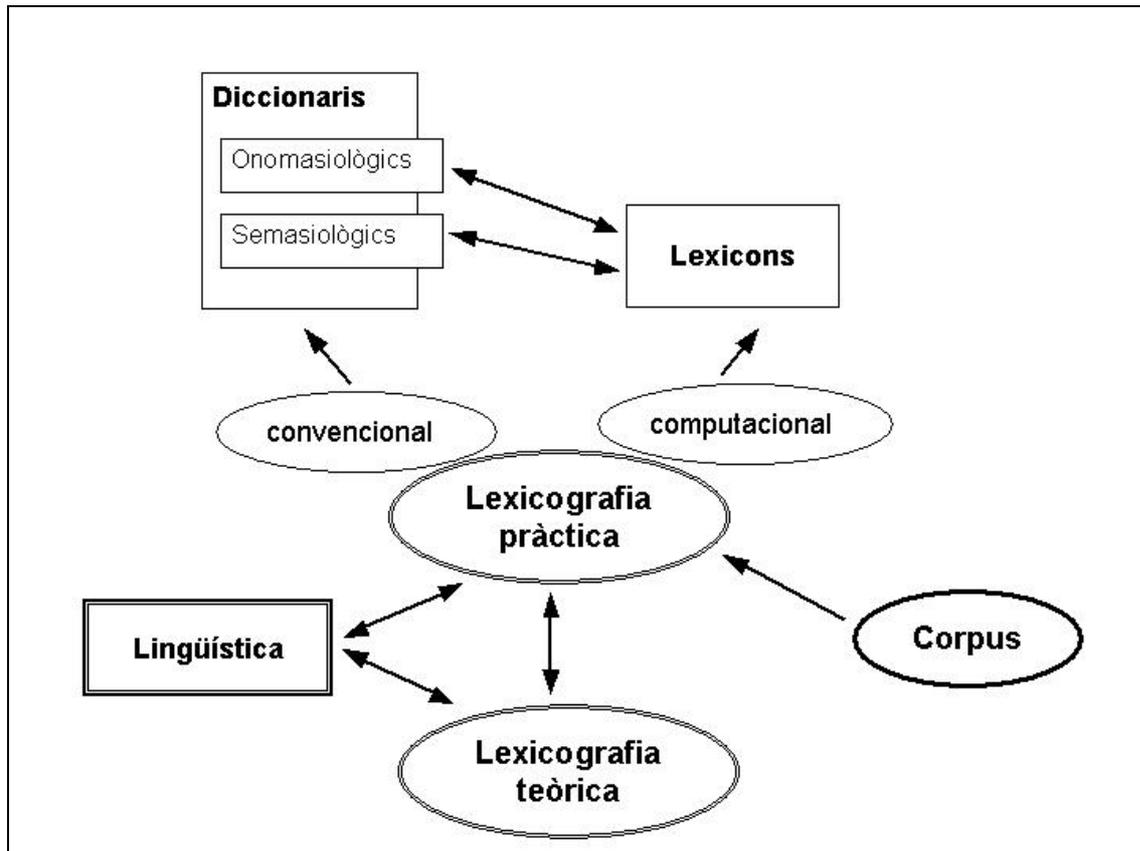


Figura 1.1. Activitats relacionades amb la descripció lexicogràfica.

Aquest treball s'inscriu de manera genèrica en aquest panorama canviant i es refereix de manera primordial a un dels elements estructurals del diccionari que presenta una problemàtica més sensible als canvis que es produeixen en la lexicografia: la definició.

1.2 Enfocament del treball

Entre tots els aspectes relacionats amb la lexicografia, aquest treball s'ocupa de la descripció del significat. Per la seva naturalesa, la definició lexicogràfica és, de tots els components del diccionari, el que concentra d'una manera més evident la majoria de les

qüestions fonamentals relatives a l'expressió del significat conceptual del terme definit. El conjunt de les definicions d'un diccionari pot ser, a més, interpretat com una estructura conceptual referida a la totalitat del lèxic que reflecteix els diversos tipus de relacions semàntiques que els elements lèxics mantenen entre si. Per aquest motiu, l'interès primordial d'aquest treball se centra no només en la definició lexicogràfica, sinó també en les seves projeccions estructurals.

Entre tots els tipus de diccionaris possibles, aquest treball es dedica als diccionaris generals monolingües, que són el tipus més característic de diccionaris (*el diccionari*, quan hom s'hi refereix sense cap altra especificació). En concret, les propostes d'aquest treball s'adrecen a la sistematització de l'expressió de les definicions en els diccionaris catalans. L'anàlisi de diccionaris existents, però, no està limitada al cas del català, ja que m'ha interessat d'avaluar les aportacions metodològiques que s'han produït en els darrers anys en la lexicografia feta en altres llengües.

A partir de les consideracions que he exposat a § 1.1, assumeixo que un diccionari que correspongui a l'actual nivell de desenvolupament de la tècnica lexicogràfica ha de donar resposta adequada a les qüestions que s'hi plantegen d'una manera general. En concret, crec que un diccionari que aprofiti els canvis produïts darrerament en la lexicografia ha de satisfer els següents requeriments:

- a) Ha de fonamentar les seves descripcions sobre bases empíriques adequades (corpus de referència).
- b) Ha de caracteritzar els elements lèxics des de les perspectives activa i passiva.
- c) Ha de ser coherent pel que fa a la seva descripció semàntica, i dotar-se dels mecanismes adients per assegurar al màxim les possibilitats de reutilització de la informació que conté.

Les necessitats de recursos que planteja el primer d'aquests tres aspectes estan cobertes, en el cas del català, pel "Corpus Textual Informatitzat de la Llengua Catalana" (CTILC) (Rafel, 1994; Rafel, 1996), que s'ha concebut com un corpus de referència, constituït

per l'Institut d'Estudis Catalans com a part integrant del seu programa de recerca "Diccionari del Català Contemporani". Aquest corpus, d'una extensió total de 52,3 milions de mots, ha donat lloc a la publicació del DFDC, que en recull els resultats estadístics. El CTILC, a més, és un corpus específicament orientat a la redacció de diccionaris, i actualment serveix de base per a la redacció del "Diccionari descriptiu de la llengua catalana" (DDL), ja que compta amb totes les eines d'exploració necessàries per a la recuperació d'informació lexicogràfica. Al llarg d'aquest treball em refereixo sovint a dades extretes del CTILC per donar suport a diversos aspectes de les idees que s'hi exposen, però sense analitzar amb detall les múltiples implicacions metodològiques que presenta la utilització de corpus en el treball lexicogràfic.

Les necessitats del caràcter bifuncional (actiu i passiu) dels diccionaris monolingües obliguen a plantejar-se la relació entre característiques d'inserció lèxica, estructura argumental, propietats col·locatives i comportament morfològic, d'una banda, i descripció semàntica, de l'altra. Tot i que les qüestions que conformen el primer grup no són tractades amb profunditat en aquest treball, hi faig referència en diverses ocasions per l'estret lligam que mantenen amb la definició lexicogràfica.

Pel fet que aquest treball s'ocupa especialment de l'expressió del significat, vol incidir de manera primordial en el tercer dels aspectes que acabo d'assenyalar. Em centro en l'anàlisi de les definicions lexicogràfiques, no amb voluntat de negar la importància de la resta d'elements estructurals del diccionari, sinó amb la idea que aquest aspecte concentra un dels punts fonamentals de l'actual recerca en lexicografia. La coherència dels diccionaris (veg. Pascual, 1996) és bàsica per a l'objectiu de la reutilització de la informació, i és un requisit que només es pot aconseguir a través de la sistematització del procés d'elaboració de diccionaris. Aquesta sistematització constitueix una de les línies essencials en el desenvolupament de la recerca lexicogràfica moderna.

Aquesta línia de recerca i de desenvolupament, d'altra banda, no és independent de la llengua que es descriu, ja que requereix la formulació d'una metallengua adequada per a la descripció lexicogràfica en els diccionaris monolingües. Aquesta relació de

dependència fa especialment necessari el treball sobre la metallengua de definició dels diccionaris catalans, amb l'objectiu que la lexicografia catalana se situï al nivell de desenvolupament que determinen els processos de canvi a què he al·ludit a § 1.1.

Aquest treball parteix del fet que un diccionari és un conjunt altament organitzat de dades interdependents. En certa mesura, aquesta interdependència és un correlat de les relacions que es produeixen en la caracterització semàntica de tots els elements lèxics d'una llengua. Tenint en compte això, una definició lexicogràfica no és una relació que s'esgota exclusivament en l'equivalència que manté respecte al terme definit, sinó que forma part d'una metallengua que és correlat de la llengua que es descriu, i els elements de la qual mantenen un gran nombre de relacions particulars que, en conjunt, configuren una xarxa semàntica complexa. En aquest sentit, es pot assumir que els diccionaris poden (o han de) reflectir l'estructura del lèxic pel que fa a la seva descripció semàntica.

Per raó de les relacions internes que manté la metallengua, les definicions lexicogràfiques solen implicar dos tipus de desviacions: la circularitat i la redundància. La circularitat és una desviació de tipus estructural, que prové *necessàriament* del fet que la metallengua de descripció és un subconjunt de la llengua general (és a dir, de l'objecte de descripció). Aquesta vinculació implica, en la lexicografia convencional, la necessitat lògica d'algun tipus de circularitat (és a dir, si *a* significa *b*, *b* significa *c*..., en la sèrie *a, b, c...*, *n* es produeix que *n* significa *a*).

A diferència de la circularitat, que és una desviació estructural associada al mètode de descripció i, per tant, *inevitable* sense la modificació substancial d'aquest mètode,³ la *redundància* (*a* significa *b* que *c*, *b* significa *e* que *d*, *d* significa *a* que *c*) representa la inclusió d'informació innecessària, ja present en la descripció d'algun dels elements de la definició, i es pot concebre com una desviació metodològica evitable.

³ De fet, l'única manera factible d'evitar la circularitat és el desenvolupament d'una metallengua de descripció totalment independent de l'objecte descrit. En aquest sentit s'han fet propostes (veg. més

La lexicografia moderna no ha prestat una atenció especial als fenòmens de la redundància i de la circularitat. Tot i així, determinades estratègies de descripció han contribuït a pal·liar els efectes d'aquestes desviacions, especialment de la circularitat. És el cas dels anomenats *vocabularis controlats*, aplicats per diccionaris com ara el LDOCE i el CIDE. Aquests diccionaris utilitzen com a metallengua de definició un conjunt explícit i limitat d'elements lèxics. Això condueix, d'una manera natural, a la limitació dels problemes estructurals del diccionari al conjunt establert com a metallengua. Tot i així, els vocabularis controlats no han estat ideats com a factor de reducció de les desviacions del diccionari, i a més han estat criticats per la limitació i la inexactitud que impliquen en alguns casos. A més, no suposen *per se* un factor de minimització de la redundància, i les circularitats són freqüents pel fet que es tracta de conjunts de descripció relativament grans (2000 mots o més).

Des d'un punt de vista metodològic, a més, també pot dir-se que els vocabularis controlats no apliquen amb claredat certes restriccions de tipus descriptiu. Entre les més importants d'aquestes mancances es pot destacar la inexplicitud (o l'escàs grau d'explicitud) que aquests vocabularis tenen respecte del sentit concret amb què són utilitzades les seves unitats i, en segon terme, la tendència a incorporar sentits diversos per a una mateixa unitat del vocabulari. Des del punt de vista metodològic, cal aclarir que un vocabulari d'aquest tipus no és una llista d'entrades, sinó d'unitats de la metallengua de descripció, que fonamenta l'adequació descriptiva i la coherència estructural del diccionari. El poc rigor amb què s'han resolt les condicions d'utilització dels vocabularis controlats constitueix un element de crítica fonamental a aquesta estratègia descriptiva.

endavant § 6.2) que desenvolupen un conjunt d'elements de tipus axiomàtic i universal que serveix per a la formulació de definicions. M'ocupo amb detall de la circularitat lexicogràfica a § 5.1.

1.3 Objectius i hipòtesis del treball

Aquest treball es proposa la superació dels vocabularis controlats a partir de la formulació d'una metallengua que, integrada en un sistema de definició, constitueixi un recurs per a l'elaboració de diccionaris més coherents i descriptivament adequats.

En l'enfocament del treball que he exposat a § 1.2 he avançat els problemes que comporta el mètode de descripció lexicogràfica, basada en l'explicitació del significat mitjançant una metallengua que s'identifica amb la llengua natural. També he fet notar les mancances dels sistemes per arribar a una descripció coherent que comporti la superació dels problemes estructurals i metodològics que sovint s'associen a la lexicografia. Alhora, cal tenir en compte que la utilització d'una metallengua de descripció basada en la llengua natural (veg. més enrere § 1.2) és un dels elements que d'una forma més característica identifiquen la tasca lexicogràfica i, per tant, se li ha de suposar un caràcter inherent des del punt de vista metodològic.

Tenint en compte aquest estat de coses, que assumeixo com a marc de treball, cal la proposició d'un sistema que:

- a) Minimitzi els efectes de la circularitat, mantenint el llenguatge natural com a base del sistema de descripció del significat.
- b) No estigui limitat *a priori* a un conjunt determinat d'elements, a fi d'evitar les mancances descriptives que provoca l'ús de vocabularis controlats.
- c) Estableixi amb rigor els criteris d'utilització de la metallengua de definició en un sistema de definició clar i estructurat.
- d) Elimini la redundància en la descripció lexicogràfica.

La formulació d'un sistema que reuneixi aquestes característiques ha de partir de certes anàlisis i reflexions teòriques sobre el mètode de descripció lexicogràfica, sobre les

seves mancances i sobre les necessitats de sistematització descriptiva. Les aportacions teòriques d'aquest treball a la metodologia de descripció lexicogràfica se centren en:

1. La sistematització dels criteris de delimitació dels objectes de descripció, és a dir, de les entrades de diccionari.
2. L'anàlisi dels diccionaris com a model estructurat del lèxic i la revisió d'algunes aportacions de la lexicografia computacional i dels lèxics electrònics.
3. L'estudi de les característiques semàntiques que incideixen o haurien d'incidir en la formulació de les paràfrasis de definició.
4. L'estudi de la definició lexicogràfica a partir de les contribucions de la lexicografia teòrica i la presentació de certs aspectes de l'equivalència definit-definició.

El principal objectiu d'aquest treball, però, no és estrictament d'índole teòrica sinó aplicada, i es pot formular de la manera següent:

Sistematitzar l'expressió del significat i millorar els aspectes estructurals de la descripció lexicogràfica en els diccionaris catalans, a través de la proposta d'una metallengua de descripció integrada en un sistema de formulació de les definicions.

Les hipòtesis de partida que em proposo de demostrar ordenadament per a la consecució d'aquest objectiu principal són les següents:

1. Es pot determinar un vocabulari controlat apte per a la lexicografia catalana, a partir de l'aplicació, sobre recursos lingüístics referits al català, de criteris de selecció similars als dels vocabularis controlats de l'anglès.
2. A partir del vocabulari obtingut a 1 i de la vinculació de cadascun dels seus elements amb la informació que un diccionari convencional aporta, es pot determinar un conjunt

de descripció anotat semànticament, encara més reduït i que, per tant, concentri en menys elements les desviacions estructurals i metodològiques.

3. Es pot descriure un sistema de definició adequat que, a partir del conjunt nuclear obtingut a 2, estableixi els mecanismes de projecció d'una metallengua de definició descriptivament adequada per a la redacció de diccionaris catalans.

El sistema establert com a resultat de la demostració ordenada d'aquestes tres hipòtesis ha de satisfer les condicions de l'objectiu principal del treball, pel que fa a l'adequació descriptiva i a la coherència estructural del diccionari. A més no ha de limitar *a priori* la metallengua total de definició del diccionari (pels inconvenients descriptius que això suposa) i, alhora, ha de reduir al mínim del conjunt nuclear les desviacions estructurals de tota la metallengua.

Del procés que condueix a la consecució de l'objectiu principal i a la demostració de les hipòtesis de partida se'n deriven algunes aportacions d'importància per a la lexicografia catalana, que es poden concretar en les dues següents:

1. Contribuir a l'elaboració de diccionaris catalans més rigorosos des del punt de vista de les seves possibilitats descriptives i, fins i tot, de reutilització per a finalitats no humanes.
2. Contribuir al desenvolupament de la lexicografia teòrica aplicada al català, a través de dotar-la d'eines de descripció adequades i específicament concebudes.

La consecució de l'objectiu principal d'aquest treball i la demostració de les hipòtesis que hi condueixen s'inscriuen en el marc general d'una concepció de la lexicografia metodològicament renovada, integrada en els aspectes teòrics, computacionals i aplicats de la descripció del lèxic, però alhora respectuosa amb la característica distintiva dels diccionaris convencionals: la utilització d'una metallengua “natural” de descripció.

1.4 Estructura del treball

A § 1.2 he introduït, sense aprofundir-hi, la idea que els diccionaris han de reflectir l'estructura del lèxic d'una llengua. L'exposició més detallada dels motius que donen lloc a aquesta afirmació s'aborda, com a tema general, a partir de dos pols d'organització temàtica: *a)* l'anàlisi del concepte d'entrada lèxica (§ 2) i *b)* la presentació de les similituds i de les divergències entre els models teòrics del lèxic (i les qüestions que es planteja la lexicologia sobre l'estructura del lèxic) i els diccionaris (§ 3).

Dedico § 2, doncs, a un dels conceptes fonamentals en lexicografia: l'entrada de diccionari. L'entrada constitueix l'objecte a partir del qual s'organitza el diccionari en la seva dimensió macroestructural i, tot i que és subsidiària del concepte de *lexema* establert per la lingüística, no és totalment coincident amb aquest. A més, sovint es produeixen, entre diccionaris diferents, interpretacions divergents d'una mateixa realitat lèxica. Aquestes divergències solen comportar una estructuració diferent: la informació que en un diccionari es troba sota una mateixa entrada, en un altre diccionari pot trobar-se repartida sota entrades diverses; això sol justificar-se amb arguments de tipus diacrònic, encara que les interpretacions divergents poden trobar-se també fàcilment en diccionaris que pretenen donar una descripció sincrònica del lèxic. Així, doncs, analitzo la problemàtica relacionada amb l'entrada de diccionari, tant des del punt de vista de la delimitació de les unitats d'anàlisi com dels criteris que es poden tenir en compte per al tractament de la macroestructura del diccionari des d'una perspectiva sincrònica, a fi d'establir una proposta explícita per a la delimitació de l'entrada de diccionari sobre bases estrictament formals i sincròniques.

Un cop establerts els principis per a l'organització macroestructural del diccionari, dedico § 3 a dues qüestions fonamentals per a la vinculació dels diccionaris amb l'estructura del lèxic:

a) el tractament que fan els diccionaris d'un problema bàsic de la lexicologia: la creativitat lèxica (§ 3.1), i

b) la reflexió sobre els punts de contacte que presenten els diccionaris respecte dels models teòrics del lèxic, des d'un punt de vista global (§ 3.2) i de la presentació particularitzada de diversos models de diccionaris teòrics que han tingut una especial incidència en la lexicografia (convencional o computacional) (§ 3.3 - § 3.5).

La resta del treball es dedica, d'una banda, a l'anàlisi de la incidència dels diferents aspectes de la descripció del significat en la definició lexicogràfica i, de l'altra, a la formulació d'una proposta per a la sistematització de la descripció semàntica basada en la utilització del llenguatge natural (el català, en aquest cas) com a metallengua de descripció. Així, § 4, § 5, § 6, § 7 i § 8 es refereixen de manera gairebé exclusiva a les qüestions relacionades amb la formulació de les definicions lexicogràfiques.

En primer lloc analitzo (§ 4), des d'un punt de vista lingüístic, el valor de la relació d'equivalència (§ 4.1) existent (en el millor dels casos, caldria dir) entre definit i paràfrasi de definició. L'assumpció que la definició lexicogràfica adquireix, respecte del definit, una relació de sinonímia és generalment acceptada, però presenta una sèrie d'implicacions que la matisen: la substituïbilitat entre definit i definició (§ 4.2) està sotmesa als condicionants derivats del fet que definit i definició no són, en general, elements de la mateixa naturalesa; mentre que el primer és un lexema, el segon sol ser una paràfrasi formada per una sèrie de constituents referits a diferents aspectes descriptius que es troben de forma unitària en el definit. Perquè la substituïbilitat sigui possible s'han de verificar una sèrie de condicions d'equivalència, que jo concentro en quatre, referides a diferents aspectes de la semàntica lèxica: l'equivalència argumental (§ 4.3), l'equivalència actancial (§ 4.4), l'equivalència esdevenimental (§ 4.5) i l'equivalència sincategoremàtica (§ 4.6).

Un altre dels aspectes que presenta una problemàtica específica en relació a la definició lexicogràfica és el vocabulari de definició. El vocabulari, analitzat en termes de les relacions entre els constituents de la definició i el definit, d'una banda, i de la definició

amb la resta de les definicions, de l'altra, concentra una part important dels problemes descriptius associats correntment als diccionaris. Dedico § 5 a l'anàlisi, d'una banda, dels problemes estructurals i metodològics vinculats a la lexicografia convencional i, de l'altra, a les possibilitats de minimització d'aquestes desviacions a través de la utilització conscient d'un vocabulari de definició que limita el nombre d'elements a utilitzar en la redacció de definicions. Finalment, estableixo la llista proposada com a vocabulari definidor per a un diccionari monolingüe del català basant-me en dades freqüencials.

A § 6 mostro que la mera proposta d'una llista de mots per a ser utilitzada com a vocabulari definidor és una condició necessària però no suficient per a l'establiment d'una metallengua de definició. En primer lloc, perquè cal que aquesta llista sigui contrastada respecte a les necessitats de descripció d'un diccionari. En segon lloc, perquè cal determinar quin és el significat o significats dels elements lèxics que s'utilitzen com a vocabulari de definició. I en tercer lloc, perquè cal que un vocabulari de definició s'integri en un sistema descriptiu més ampli: un sistema de definició. L'explicitació de les condicions generals que ha de reunir aquest sistema és una de les aportacions fonamentals de § 6.

Un cop establertes aquestes condicions, dedico § 7 a la formulació d'un sistema de definició per a un diccionari del català; el mètode de treball que he seguit en la formulació d'aquest sistema és la vinculació de la llista obtinguda a § 5 amb les definicions que figuren al DIEC referides a les unitats lèxiques establertes com a punt de partida de la proposta. Aquesta operació representa, de fet, la conversió de la llista de mots en una sèrie d'entrades de diccionari. Una sèrie d'operacions selectives sobre el conjunt d'aquestes entrades, amb la finalitat de determinar els sentits amb què cada unitat intervé en la metallengua de definició, donarà un conjunt d'unitats de metallengua (que anomeno *nivell 0*) integrable en un sistema complet de descripció del significat en lexicografia. Un cop establert aquest nivell 0, em proposo d'arribar a la formulació general d'un sistema de descripció del significat en lexicografia que compleixi els requisits generals que he esmentat a § 1.3.

A fi de completar i de detallar alguns aspectes del sistema de descripció formulat a § 7, dedico § 8 a la caracterització d'alguns aspectes de l'aplicació del sistema. Fonamentalment, em proposo de verificar fins a quin punt aquest sistema compleix les condicions de descripció assenyalades en aquest treball i, també, d'exposar les implicacions de la utilització del sistema en el procés de redacció d'un diccionari general monolingüe.

A les conclusions (§ 9) faig la recapitulació de tot el treball i em centro de manera especial en la contrastació de les hipòtesis plantejades a § 1.3 i en la presentació del grau de consecució dels objectius del treball.

2 Els diccionaris i el model del lèxic

2.1 *Del lexema a l'entrada de diccionari*

La lexicografia classifica la informació sobre el lèxic a partir d'un dels elements estructurals més característics del diccionari: l'entrada. A grans trets, aquesta presentació resulta coincident amb els models basats en la lingüística que interpreten el lèxic com un inventari d'unitats (entrades) que conjuntament formen el component lèxic (també anomenat *diccionari*) de la llengua. La lingüística concep el lèxic, doncs, com un conjunt nombrós d'elements d'una llengua sotmès a certs processos de regularitat variable determinats per la morfologia, que interactua amb la gramàtica.

L'entrada de diccionari és una entitat subsidiària del concepte de *mot* o, més precisament, del concepte lingüístic de *lexema* (Lyons, 1968). Els intents de definició del mot (veg. per exemple González Calvo, 1982) que s'han assajat en el marc de la lingüística teòrica del segle XX es basen, en major o menor mesura, en un nombre bastant limitat de criteris de segmentació (o de delimitació) possibles.

A banda dels criteris fonològics, que no poden considerar-se determinants per a totes les llengües, i que rarament passen de ser fenòmens que afecten només algun tipus de mot determinat, i dels criteris semàntics, que sempre s'han considerat pràcticament inoperants o decididament secundaris per a aquesta finalitat, s'han imposat altres criteris d'ordre funcional que poden ajudar a determinar formalment el concepte de mot.

González Calvo (1982: 393) assenyala quatre criteris d'aquest tipus, que sintetitzen les aportacions de la lingüística (especialment de la lingüística estructural) al problema de

la caracterització del mot. Aquests criteris són: *a)* l'autonomia oracional, *b)* la separabilitat, *c)* les possibilitats de permutació, i *d)* la funció sintàctica mínima.

El punt en comú de totes aquestes proves és que parteixen exclusivament de l'enunciat lingüístic a fi de donar una definició delimitativa del mot dins de la cadena. En realitat, però, cap d'aquests criteris no pot ser considerat autosuficient, ja que el seu grau d'interdependència és alt i els límits entre cadascun resulten en ocasions força imprecisos; per això cal entendre'ls com a estratègies diverses de delimitació d'unitats que cal que es complementin les unes amb les altres per obtenir resultats satisfactoris.

L'anàlisi delimitativa, però, permet únicament de caracteritzar un mot com una unitat pertanyent a un conjunt d'elements directament observables en la parla i, en tot cas, representa el punt de partida d'anàlisis més interpretatives (veg. per exemple Matthews, 1974) des del punt de vista lingüístic. Cal transcendir aquests criteris per a arribar a considerar un altre dels sentits que sol tenir l'expressió 'mot' en un dels seus usos més habituals, que coincideix amb el sentit amb què s'utilitza en la gramàtica escolar i tradicional.

El mot, considerat en aquests termes tradicionals (que es corresponen bastant fidelment a la idea de mot com a *entrada de diccionari*), consisteix en una unitat de rang més alt que la que prové d'una anàlisi merament delimitativa, i agrupa tota una sèrie flexional.

Aquesta idea de mot com a unitat més o menys *abstracta* (Lyons, 1968, 1977) no ha estat, com es pot suposar, compartida universalment per tots els corrents de la lingüística d'aquest segle, sinó que amb molta freqüència ha estat considerada un element extern als interessos d'aquesta ciència, o en tot cas una rèmora de la gramàtica escolar.⁴

⁴ Bloomfield (1933) fa notar que la tradició escolar manté que formes tals com *book*, *books*, són formes d'un mateix mot. Bloomfield qualifica d'inexacta aquesta idea, a causa de l'existència de diferències de forma i de sentit entre les formes, que corresponen a formes lingüístiques diferents i, doncs, a mots diferents.

Lyons (1968) sistematitza les principals diferències existents en els diversos sentits amb què s'ha usat el terme *mot*, i n'aïlla tres valors:

a) El mot fonològic (o ortogràfic), que es caracteritza per ser la *realització* d'un mot gramatical; per exemple, el mot *descriuen*.

b) El mot gramatical: unitat de funció gramatical i combinatòria; per exemple, la *tercera persona del plural del present d'indicatiu de "descriure"*.

c) El mot com a entitat abstracta, o *lexema*:⁵ unitat que agrupa les *formes* pertanyents a un mateix paradigma flexional. La gramàtica tradicional havia usat sovint *mot* en aquest sentit, que permet afirmar proposicions del tipus "*descriuen*" és una forma de "*descriure*", on *descriure* no és l'infinitiu del verb, sinó una entitat referencial que es representa convencionalment a través d'una de les seves formes o *expansions* flexionals.

El concepte de *lexema* projecta sobre la totalitat del lèxic la dissociació entre un nivell *abstracte* (el del lexema) i un de més *concret* (el de la forma). Es pot dir que el lèxic està format per un conjunt d'elements del primer tipus, i que la seva presència en el discurs es produeix a través d'elements del segon. Cal no oblidar, però, que aquesta projecció parteix d'alguns pressupòsits que convé assumir amb caràcter previ:

1. L'existència d'una diferenciació clara entre la morfologia flexional (que intervé en processos de formació de *formes de lexemes*) i la morfologia derivativa (que intervé en processos de formació de *lexemes*). Tal com assenyala Anderson (1992), hi ha alguns elements sobre els quals es pot basar aquesta distinció que, tot i això, de vegades no resulta del tot clara.

⁵ Lyons (1968) proposa, amb notable èxit, la denominació *lexema* per referir-se a aquest concepte. El terme havia estat ja utilitzat per Hockett (1958) amb un sentit que no té res a veure amb el que proposa Lyons. Alhora, Lyons objecta a la lingüística teòrica que no hagi continuat amb la pràctica tradicional de definir el mot com una unitat *abstracta*.

2. L'existència, en tots els casos, de la mateixa relació entre el lexema i la forma (o formes) que el representen en el discurs, malgrat que no totes les categories lèxiques són susceptibles de rebre marques de variació flexional.
3. L'existència de processos lèxics en què característiques primàriament flexionals passen a formar part inherent dels mots; per exemple, substantius que només es realitzen en singular o en plural, adjectius que només combinen amb noms masculins, etc.
4. L'assumpció que tots els mots, en tant que lexemes, tenen una categoria lèxica caracteritzable en termes d'una part de l'oració. Aquesta categorització és dubtosa o, en tot cas, intuïtiva en els mots lligats del tipus *antuvi* (en expressions com *d'antuvi*) o *pirandó* (en expressions com *tocar pirandó*).

Lyons no explicita quins són els límits del concepte abstracte anomenat *lexema*. De fet, no estableix cap diferència operativa entre els diferents tipus de mots: és a dir, entre els pertanyents a categories *tancades*, com ara preposició, article, etc.; i els pertanyents a categories *obertes*, com ara substantiu, adjectiu, verb, etc.

Els mots pertanyents a aquestes categories *tancades* —anomenats en ocasions *mots buits* i també *clítics*— es diferencien dels mots *plens* per la característica que no són significatius en si mateixos (Ullman, 1962: 51).

Dit d'una altra manera, els mots buits no tenen un significat objectivable independentment dels contextos en què poden aparèixer, ja que hom parteix de la idea que els mots buits només adquireixen significat quan es troben acompanyats d'altres mots.

Les dificultats metodològiques que planteja aquesta distinció, basada exclusivament en principis semàntics,⁶ s'han volgut relativitzar amb criteris d'ordre fonològic,⁷ o d'ordre

⁶ Ullman (1962: 52) assenyala que la lingüística té resistència al reconeixement de categories lèxiques

gramatical.⁸

Mel'čuk (1993: 229) determina sis grups principals de mots buits, que ell anomena *clitics*:

- Pronoms personals
- Articles
- Verbs auxiliars
- Preposicions
- Conjuncions
- Adverbis (=partícules)

Des del punt de vista quantitatiu, els clítics són elements molt recurrents en els textos i poden caracteritzar-se sobre bases freqüencials. Aquesta caracterització ve donada pel fet que els pertanyents a classes gramaticals *tancades* tendeixen a aparèixer amb una

sobre bases únicament semàntiques, i no sobre la base de l'expressió formal.

⁷ En l'ordre fonològic Ullman (1962) addueix, per exemple, que en anglès el so fricatiu linguodental inicial (representat ortogràficament per *th-*) es pronuncia sord en el cas dels mots plens (*thank, theft, thin, thorn, etc.*), mentre que la pronúncia és sonora en el cas dels mots buits (*than, the, those, etc.*). En el cas del català, es constata el fet que un cert nombre de mots buits són mots àtons i que, a més, segueixen comportaments específics pel que fa als seus enllaços fonètics amb altres elements de la cadena lingüística. Per a un conjunt de proves de tipus fonològic per a la determinació d'aquest tipus de mots en català, veg. Mascaró (1986).

⁸ En l'ordre gramatical, s'ha remarcat sovint l'existència d'una relació estreta entre el concepte de *categoria gramatical* i alguns mots buits. De fet, són elements que se situen entre el mot i el morfema (Mascaró, 1986; Mel'čuk, 1993). Fins i tot es pot adduir que certs mots buits tenen un comportament gairebé idèntic al de les marques flexionals (*vàrem* estudiar = estudiàrem), o que corresponen a l'expressió de relacions que tenen en certs casos solucions de tipus morfològic (*molt fosc* = fosquíssim; o, en anglès, *més fosc* = darker; *el més fosc* = darkest).

Altres comprovacions de tipus gramatical per a avaluar l'existència de factors de distinció en els mots plens consisteixen en observacions de tipus sintagmàtic. Alguns mots buits semblen evidenciar unes menors possibilitats de commutació dins de l'enunciat que no pas els mots plens. Aquesta limitació és absoluta en el cas dels anomenats pronoms febles, per exemple, en què no hi pot haver cap mot, ni ple ni buit, entre aquests i el verb (o perífrasi verbal) que complementen (“agafa-les”, “les agafa”, “les pot agafar”, “poder-les agafar”, “poder agafar-les”, etc. però **“agafa ara les”* o **“les cada dia agafa”*), cosa que no afecta els sintagmes que els podrien reemplaçar (“agafa ara mateix *les claus*”, “ell agafa cada dia *les claus*”, etc.).

freqüència més elevada que els mots plens i a més solen tenir un nombre reduït de representants per cada categoria gramatical. Aquest fet s'evidencia si es prenen en consideració alguns resultats extrets del DFDCC.

Categoria gramatical	Freqüència absoluta
a	4714242
af	1247
ai	1330685
aia	143267
aip	288
am	1602
amp	22
an	214410
ap	4350
ar	6378261
av	3240680
c	3377173
ct	1274420
f	5301932
fp	16554
fs	100
i	58847

l	41304
la	2576
m	5651909
mf	137924
mp	6284
ms	4876
nc	731
p	3981286
po	7180153
v	603040
va	194259
vi	1634661
via	655393
vip	168642
vp	58851
vt	1359972
vtp	1784582
vvp	1688466
TOTAL	51212989

Figura 2.1. Llista de codis gramaticals de lema amb la freqüència acumulada

La figura 2.1 recull les dades sobre la freqüència de cadascuna de les categories⁹ (codis de lema del CTILC) del DFDCC. Interpretades adequadament, aquestes dades permeten extreure conclusions d'interès sobre el comportament freqüencial de les diferents categories lèxiques.

A la figura 2.2 dono el resultat d'aquestes dades agrupades per categoria lèxica (adjectiu, adverb, etc.), seguida per la freqüència acumulada, i per una tercera columna

⁹ Mantinc l'ús de la classificació que fou aplicada en el procés de lematització del CTILC (i per tant, a les dades contingudes en el DFDCC). Aquesta classificació, seguint un criteri corrent en els diccionaris convencionals, combina l'ús d'etiquetes basades en la categoria lèxica amb d'altres etiquetes basades en informacions no exclusivament sintàctiques.

que expressa una magnitud de relació consistent en la divisió de la freqüència total de les ocurrencies pel nombre de lemes en què es distribueix. El resultat d'aquesta darrera operació mostra la mitjana d'ocurrencies per lema que té cada categoria. Per a una presentació més aclaridora d'aquestes dades, dono els resultats ordenats decreixentment a partir d'aquesta relació.

Les categories que mostren una relació més alta d'ocurrencies per lema pertanyen a la classe dels mots buits, mentre que en la resta de categories el valor d'aquesta relació és significativament baix. La categoria adverbial, que recull tant els adverbis anomenats *partícules* (*ara, aquí, darrere, sempre, etc.*) i els adverbis amb el formant *-ment* (considerats habitualment dins de la classe dels mots plens), constitueixen, des d'aquest punt de vista, una mena de categoria mixta. Aquesta observació pot fer-se extensiva a la categoria dels adjectius, ja que en l'etiquetatge morfosintàctic del CTILC s'aplica el mateix codi tant als adjectius qualificatius com als demostratius (*aquest, aquell, etc.*).

D'altres etiquetes del quadre que no corresponen estrictament a parts de l'oració (com ara *no codificat, mot lligat* o *locució al·lòctona*) no han de ser preses en consideració i no invaliden la delimitació de la diferència que s'estableix entre mots plens i buits.

	Ocurrencies	Lemes	Oc/lema
Article	6378261	3	2126087,00
Contracció	1274420	8	159302,50
Conjunció	3377173	41	82370,07
Pronom	3981286	55	72387,01
Preposició	7180153	151	47550,68
Adverbi	3240680	4313	751,37
Verb	8147866	10914	746,55
No codificat	731	3	243,66
Nom	11119579	25968	204,81
Mot lligat	41304	206	200,50
Adjectiu	6410113	39188	163,57
Interjecció	58847	1101	53,44
Locució al·lòctona	2576	500	5,15
	51212989		

Figura 2.2. Llista de categories amb ocurrencies, lemes i relació ocurrencies/lema.

La distinció lexicogràfica entre aquests dos grups de mots es verifica en els tipus d'informació que solen vehicular els representants d'una i d'altra classe. En els mots buits la *definició* lexicogràfica pròpiament dita és substituïda per una altra mena d'informació consistent en l'*explicació* del seu funcionament gramatical, i s'integren dins del text corresponent a l'article de diccionari informacions i fórmules que no segueixen el principi de substituïbilitat definit-definició.

La descripció lexicogràfica no estableix distincions excloents pel que fa a aquesta doble caracterització del lèxic, malgrat que alguns lingüistes han remarcat el caràcter no lèxic dels clítics. Coseriu (1987) restringeix el concepte de *lèxic* als mots que ell anomena *lexemàtics*, els quals structuren i representen la realitat extralingüística (*home, bosc, blanc, córrer*, etc.); la resta de mots que distingeix la classificació de Coseriu, és a dir, els *catagoremàtics* (pronoms) i els *morfemàtics* o *instrumentals* no pertanyen al domini del lèxic.¹⁰

Mel'čuk (1993: 326) distingeix dues classes de significació lexical: la semàntica i la sintàctica. La primera correspon als mots plens i la segona als mots buits. El diccionari com a component de la llengua els descriu tots dos com a lexemes (veg. Mel'čuk, 1993: 338). Aquesta pràctica coincideix amb la de la lexicografia convencional, que sol tractar tots dos tipus de mots sense exclusió. Les diferències entre una i altra classe s'evidencien d'una manera molt especial pel mètode a partir del qual la lexicografia en fa les respectives descripcions; especialment en la manera com es formula la definició dels mots pertanyents a una o a altra de les dues classes. Així, doncs, i atès que la descripció semàntica pròpiament dita només és possible en els mots lexemàtics, la lexicografia ha desenvolupat mètodes de descripció basats en l'explicitació del comportament i de les característiques funcionals dels mots no lexemàtics. Aquestes

¹⁰ Tal com Coseriu mateix diu: “sólo las palabras lexemáticas pertenecen con pleno derecho al léxico y, en consecuencia, al objeto propio de la lexicología”. [Coseriu (1987: 133)]

diferències en el mètode descriptiu són les que permeten, de fet, l'equiparació de totes dues classes i la concepció lexicogràfica del lèxic com el conjunt de mots de la llengua, prescindint del caràcter *buit* o *ple* dels elements que en formen part.

2.2 Delimitació formal de l'entrada de diccionari

La divisió del lèxic en entrades de diccionari és una operació que requereix l'aplicació d'una sèrie de criteris d'ordre lingüístic, sotmesos a un cert model teòric (o preteòric) del lèxic i de les seves unitats. Les decisions metodològiques del lexicògraf —els criteris que aplica en l'operació de dividir el conjunt en una sèrie finita— fan que aquesta representació convencional sigui més o menys estructurada i sistemàtica. Els criteris de divisió han de complir, a més, una condició necessària: la intel·ligibilitat per part del lector; per aquest fet, es pot dir que el lexicògraf ha de fonamentar les seves decisions sobre bases formals o, si més no, en el rigor en la determinació de casos.

Ja he mostrat la semblança que hi ha, en el cas dels mots lexemàtics, entre el concepte lingüístic de lexema i el d'entrada de diccionari. Per això, els problemes que es presenten en la delimitació i diferenciació de les entrades (Boulangier, 2000) són bàsicament anàlegs als que es poden produir en el pla de la lingüística en l'establiment de l'inventari de les unitats lèxiques. El caràcter bàsicament exhaustiu dels diccionaris pel que fa a l'inventari lèxic de la llengua fa que hagin de donar resposta específica a tots els problemes relacionats amb la diferenciació en entrades de tot el conjunt de mots d'una llengua (lexemàtics i no lexemàtics). A la pràctica, els diccionaris solen presentar vacil·lacions en aquest punt, que moltes vegades obeeixen a criteris lingüístics no prou explícits, o que són el resultat de la interacció del pla de descripció sincrònic i del pla de descripció diacrònic.

Intentaré aquí d'examinar els criteris que es poden tenir en compte per a la diferenciació de les entrades de diccionari basant-me en una anàlisi estrictament sincrònica.

A partir de l'anàlisi i la segmentació d'un corpus $L \neq \emptyset$ s'obté un conjunt, en forma de llista, de formes gràfiques diferents identificables en el text (mots d'una llengua natural aïllables en el discurs) que identificaré com una successió finita¹¹ $\{f_1, f_2, f_3, f_4 \dots f_n\}$ de manera que:

$$\{f_1, f_2, f_3, f_4 \dots f_n\} \subset L$$

Existeixen, allora:

- Un conjunt $C \neq \emptyset$ que conté com a elements totes les formes canòniques d'una llengua, i que identificaré com una successió finita $\{c_1, c_2, c_3, c_4 \dots c_m\}$,
- Un conjunt $G \neq \emptyset$ que conté com a elements totes les categories lèxiques (parts de l'oració) que es consideren en una llengua, i que identificaré com una successió finita $\{g_1, g_2, g_3, g_4 \dots g_k\}$,
- Un conjunt $P \neq \emptyset$ que conté com a elements tots els paradigmes de flexió diferenciables que es consideren en una llengua, i que identificaré com una successió finita $\{p_1, p_2, p_3, p_4 \dots p_q\}$,

De manera que:

$$\begin{aligned} \{c_1, c_2, c_3, c_4 \dots c_m\} &\subset C \\ \{g_1, g_2, g_3, g_4 \dots g_k\} &\subset G \\ \{p_1, p_2, p_3, p_4 \dots p_q\} &\subset P \end{aligned}$$

¹¹ Parteixo de la base que el conjunt del lèxic d'una llengua és una successió finita, malgrat l'existència de teories que el caracteritzen com un conjunt potencialment infinit (Corbin, 1987), a causa de l'existència de mecanismes morfològics de recursivitat. El model de lèxic de què parteixo aquí és de base referencial i no competencial, i pren com a element el conjunt de mots efectivament utilitzat en un corpus que serveix de base per a l'elaboració d'un diccionari.

Qualsevol element f_i ($1 \leq i \leq n$) $\in L$ està associat a una o més tríades escollides a partir de:

- una forma canònica c_a ($1 \leq a \leq m$) $\in C$
- una categoria lèxica g_b ($1 \leq b \leq k$) $\in G$
- un paradigma de flexió p_c ($1 \leq c \leq q$) $\in P$

De manera que:

$$\{c_a, g_b, p_c\} \in f_i$$

Existeix un nombre restringit gramaticalment de combinacions possibles de la tríada formada per elements de C , G i P , de manera que es pot determinar que el total d'entrades d'un diccionari està format per un subconjunt D de L que conté totes les combinacions possibles i diferenciades almenys en un dels seus elements de $\{c_a, g_b, p_c\}$; anomenaré $\{e_1, e_2, e_3, e_4 \dots e_p\}$ cadascuna d'aquestes combinacions diferenciades.

Una entrada de diccionari e_j , doncs, es pot caracteritzar a partir de la condició de reunir una combinació única d'elements $\{c_a, g_b, p_c\} \in f_i$.

Podem considerar $f_i\{c_a, g_b, p_c\}$ i $f_i\{c_{a'}, g_{b'}, p_{c'}\}$ com a pertanyents a una mateixa entrada de diccionari $e_j\{c_a, g_b, p_c\}$ si es compleix que tenen la mateixa combinació de valors per a $\{c, g, p\}$,

És a dir:

$$f_i\{c_a, g_b, p_c\} \wedge f_i\{c_{a'}, g_{b'}, p_{c'}\} \Rightarrow e_j\{c_a, g_b, p_c\} \Leftrightarrow (c_a = c_{a'}) \wedge (g_b = g_{b'}) \wedge (p_c = p_{c'})$$

Complementàriament, es pot considerar que si existeix una entrada de diccionari $e_i\{c_a, g_b, p_c\}$ no pot existir una entrada de diccionari $e_j\{c_{a'}, g_{b'}, p_{c'}\}$ tal que compleixi les següents tres igualtats:

$$(c_a = c_{a'}) \wedge (g_b = g_{b'}) \wedge (p_c = p_{c'})$$

Com a il·lustració d'aquesta formulació, es poden adduir alguns exemples de diferenciació d'entrades com ara:

a) **menjar** *v.* / **clau** *m.* que es diferencien en primer terme per la forma canònica, en segon terme per la categoria lèxica, i en tercer terme pel paradigma de flexió a què s'associen.

b) **menjar** *v.* / **pujar** *v.* que es diferencien exclusivament per la forma canònica.

c) **component** *m.* / **component** *adj.* (o **clau** *m.* / **clau** *f.*) que es diferencien exclusivament per la categoria lèxica.

d) **menjar** *v.* / **menjar** *m.* que es diferencien en primer terme per la categoria lèxica i en segon terme pel paradigma de flexió a què s'associen.

e) **cas** *m.* (pl. *casos*) / **cas** *m.* (pl. *cassos*) que es diferencien exclusivament pel paradigma de flexió.

A partir del sistema de diferenciació d'articles formulat aquí, es determina un diccionari sense entrades indiferenciades des del punt de vista dels elements que més correntment integren la capçalera d'un article lexicogràfic (forma gràfica de l'entrada, categoria gramatical, indicacions de flexió).

En general, els diccionaris existents no apliquen amb rigor aquests criteris, sinó que solen tractar de maneres diverses determinats casos específics, normalment a causa de la consideració de motivacions de tipus diacrònic o semàntic. Un sistema formalitzat de diferenciació d'articles lexicogràfics adquireix una importància determinant en l'establiment d'una nomenclatura de diccionari a partir d'un corpus amb anotació morfosintàctica. Siguin quins siguin els criteris de diferenciació de les unitats de referència d'aquest corpus (els lemes, per exemple) o del grau de detall de l'anotació morfosintàctica de les formes, cal dotar-se dels mecanismes necessaris per a la transposició de la informació textual a la informació lexicogràfica.

Cal, però, assenyalar aquí l'existència d'una qüestió important que determina, complementàriament a aquests criteris de diferenciació, la composició de la macroestructura d'un diccionari. Em refereixo a la manera com es tracten en lexicografia les variacions semofuncionals que es produeixen entre elements que poden concórrer en una mateixa tríada. Com examino amb detall en les seccions següents (veg. § 2.3, § 2.4 i § 2.5), davant de determinats fenòmens que afecten la forma canònica, la categoria gramatical, o la flexió, el lexicògraf ha de determinar si ha de tractar certes unitats com a entrades diferents (diferenciació macroestructural), o bé com a sentits diferents d'una mateixa entrada (diferenciació microestructural).

2.3 Forma gràfica de l'entrada

La tria de la forma que identifica convencionalment el lexema té molt a veure amb els hàbits establerts per la gramàtica tradicional, que també han estat adoptats per la lexicografia en l'explicitació de les seves unitats d'anàlisi. La forma assignada per a identificar una entrada de diccionari coincideix, en general d'una manera arbitrària, amb una de les variants flexionals del lexema que representa. Així, en la lexicografia de les llengües romàniques la forma que s'escull com a entrada de diccionari és l'infinitiu en el cas dels verbs, el singular en el cas dels noms, i el masculí singular en el cas dels adjectius. El fet possible i, fins a cert punt, demostrable que aquests criteris tinguin una base lingüística és una qüestió secundària, però de tota manera es podria fer observar com, en general, la forma del lexema correspon a la forma flexional *menys marcada* gramaticalment.¹²

¹² En català podem dir que el singular és menys marcat que el plural, a través de la constatació evident que el català no posseeix marques flexionals de singular i sí de plural. El mateix podem dir del masculí respecte del femení. I, fins i tot, en un terreny més neutre, encara es pot dir que l'infinitiu (sense marques de temps, ni de persona, i amb possibilitats sintàctiques de nominalització) és la forma *menys marcada* del verb.

La forma gràfica de l'entrada sol ser l'aspecte menys problemàtic de la tríada que constitueix l'entrada lexicogràfica. La majoria de les formes que apareixen en un diccionari solen estar fixades com a *normatives* des del punt de vista prescriptiu. Els conflictes que es presenten entre diferents diccionaris per la forma normativa de tal o tal mot són mínims; en general, aquests desacords es deuen a ajustaments fets amb l'objectiu de tenir una forma normativa més fonamentada lingüísticament i, a més, solen estar en funció de la naturalesa de la institució que promou el diccionari, o del prestigi social o acadèmic del seu autor o autors. En un diccionari que, a més, integri un volum important de lèxic *no normatiu* (com és el cas del DCVB, per exemple), el lexicògraf no tindrà greus problemes per establir la *forma canònica* de les entrades a partir de l'aplicació d'uns criteris ortogràfics que en general no seran objecte de discussió.

Alguns casos particulars presenten problemes específics en funció de les característiques del lèxic de cada llengua. Aquests problemes es concreten en els conflictes que hi pot haver entre els criteris de determinació del lexema i els de l'entrada de diccionari, i es produeixen quan, a partir de diferents formes canòniques, s'expandeixen formes flexionals idèntiques susceptibles de correspondre al mateix mot. En aquests casos paradoxals els criteris de determinació d'articles fixats per a cada diccionari esdevenen fonamentals per a l'organització final de la informació. Vegem-ho amb un exemple, la manera com es distribueixen en articles de diccionari els substantius masculins *trast* i *trasto* (que, amb significat idèntic, comparteixen la forma plural *trastos*) depèn dels criteris fixats per a cada diccionari; en aquest exemple, a més, s'hi produeix el fet que *trast* té altres significats que en dos dels tres diccionaris consultats donen lloc a entrades homògrafes:

trast¹ *m.* Qualsevol dels filets de metall (...) [DIEC]

trast² *m.* Conjunt de les feines d'una casa. [DIEC]

trast³ *m.* Moble, eina, estri, etc., inservible, que hi ha en una casa. | FIG. DESP.
Persona tinguda per inútil, que fa nosa. [DIEC]

trasto *m.* POP. Trast³. [DIEC]

trast *m.* Cadascun dels filets de metall (...) [DGLC]

trast *m.* Les feines d'una casa. [DGLC]

trast [fam. *trasto*] *m* Dit dels mobles, eines, estris, etc., vells, inservibles que hi ha en una casa. | Fig. Dit despectivament d'una persona al·ludint a la seva inutilitat, a la nosa que fa, etc. [DGLC]

trast *m* **1** (...) **2** [fam *trasto*] **a** Dit d'un moble, una eina, un estri, inservible i sense valor **b** *fig desp* Dit d'una persona tinguda per inútil, que només fa nosa. **3** Feines de la casa. [DLC]

Independentment de la coherència¹³ dels criteris de diferenciació dels articles aplicats en cada un dels diccionaris mostrats, es detecten dos tipus de tractament diferenciat: *trasto* com a variant d'una entrada (DGLC i DLC) i *trasto* com a entrada que forma part de la sèrie sinonímica d'una de les entrades *trast* (DIEC).

El tractament compartit pel DGLC i pel DLC no dóna a *trasto* la consideració d'entrada de diccionari (aquesta forma es recull en un altre article, no hi ha cap element de referenciació a la posició alfabètica de *trasto*, tipogràficament se la diferencia de la forma de l'entrada, s'introdueix per la marca d'ús *fam.* [= familiar], etc.); en certa manera, s'equipara la forma *trasto* a una variant que només es descriu com una modalitat del mot d'entrada. L'absència en la posició alfabètica d'una forma susceptible de tenir una entrada específica és una mancança pròpia dels diccionaris de nomenclatura simple.

El tractament del DIEC correspon a una altra de les solucions possibles en un diccionari de nomenclatura simple: s'incorpora una entrada que té com a definició sintètica el mot al qual referenciem la descripció. Des del punt de vista de l'estructura de diccionari, aquest tractament compleix més satisfactòriament el principi d'identificar una entrada lèxica amb la tríada que he esmentat més amunt. El tractament del DGLC i del DLC, en canvi, és inexplícit pel que fa a *trasto*, ja que opta per una presentació que implica el tractament sublèxic de la forma, equiparable a les irregularitats flexionals o a les especificitats pragmàtiques.

Els diccionaris, habitualment, no s'enfronten a aquest tipus de casuística amb uns criteris uniformes. Fins i tot es poden detectar desviacions respecte al tractament que acabem d'analitzar dins dels mateixos diccionaris:

morè -ena *adj.* Bru, esp. parlant del color de la pell. [DGLC]

moreno -na *adj.* Morè. [DGLC 4^a]

caldre [Quant a la flexió com *valer*] *v. intr.* Ésser necessari. (...) [DGLC]

caler *v. intr.* Caldre. [DGLC]

Des de l'òptica d'un diccionari descriptiu es poden plantejar, encara, algunes qüestions relacionades, com les alternances de formes canòniques que podem trobar en un corpus com el CTILC del tipus *pla/plan*, *germe/germen*, adjectius com *complex/complexe*, etc.

Una manera adequada de tractar aquests i altres fenòmens consisteix en l'establiment d'una nomenclatura complexa, que relaciona cada una de les entrades principals amb les entrades secundàries variants. El resultat són dues subnomenclatures: les entrades principals (que integren en la seva estructura d'article una relació de variants), i les entrades secundàries (que estableixen una crida o remissió a la seva corresponent principal).

La distinció de forma canònica es pot deure també a les diferents possibilitats de combinació amb un paradigma de flexió. Un exemple d'això el constitueixen les sèries d'entrades que, sobre la mateixa base morfològica, s'associen a paradigmes de flexió diferenciats pertanyents a una mateixa categoria lèxica, com en el cas de *fendir/fendre* (que tan sols tenen en comú l'imperfet d'indicatiu: *fendia*, etc. i la primera persona del perfet d'indicatiu *fendi*) o de *brandar/brandir* (que tan sols tenen en comú la primera persona del passat simple d'indicatiu: *brandí*). Els diccionaris solen registrar aquestes variacions com a sèries sinonímiques:

¹³ En el cas del DGLC i del DIEC la distinció entre tres entrades homògrafes no es correspon amb els criteris que he exposat a § 2.2.

fendir *v. tr.* Fendre. [DIEC]

fendre [quant a la flexió, com *prendre*] *v. tr.* Esberlar. | Tallar (l'aire, l'aigua, el vent, etc.), travessar subtilment i de pressa, un fluid. *L'àguila fenia l'aire batent suaument les ales.* || Amb una eina de vora fina però no tallant, fer un petit solc en una cartolina que facilita de doblegar-la correctament. [DIEC]

brandar *v. tr.* Portar (una arma) a la mà movent-la amenaçadorament, brandejar. *L'investí brandant l'espasa.* || Fer tocar (una campana). *Brandeu les campanes.* || *intr.* Estar animat d'un moviment oscil·latori vibratori. *Les xemeneies branden. Les campanes branden.* | Caminar balancejant-se, gronxant-se. *Caminava brandant brandant.* [DIEC]

brandir *v. tr.* Brandejar, brandar. [DIEC]¹⁴

Finalment, hi ha encara certs casos d'indiferenciació de la forma canònica de l'entrada que poden tenir realitzacions fonètiques diferents:

golf¹ *m.* Porció gran de mar voltada parcialment de terra. *Golf de Roses.* [DIEC]

golf² *m.* Joc que consisteix essencialment a anar ficant en una sèrie de forats una pilota petita que es llança donant-li cops amb un bastó adequat. [DIEC]

Segons la informació que recull el DOP, la pronúncia d'aquestes dues entrades homògrafes difereix en el timbre de la vocal. Aquest fet podria justificar sincrònicament el tractament en entrades separades d'aquestes unitats lèxiques, mitjançant la introducció d'un quart element (la pronúncia del mot) entre els que he determinat com a diferenciadors de l'article. Un tractament d'aquest tipus no està exempt de problemes, derivats bàsicament dels grans factors de variació observables en la pronúncia d'una mateixa unitat lèxica. Aquesta variació (diatòpica, diastràtica, etc.) té com a conseqüència directa l'absència de repertoris de referència (normativa o no) prou fiables per a la determinació de la pronúncia com a factor de diferenciació dels articles de diccionari. El mateix DOP a què m'he referit dona una sola pronúncia per a les següents dues unitats que, molt probablement, podrien ser distingibles sincrònicament a partir d'aquests mateixos criteris:

¹⁴ L'article corresponent al sinònim *brandejar* és el següent:

brandejar *v. tr.* Brandar (una arma). *Brandejant amb mà ferrenya l'espasa flamejant.* [DIEC]

dinar² *m.* Àpat principal del dia (el que es fa pels volts de migdia o a primeres hores de la tarda). *Un bon dinar. Un dinar abundant. Assistir a un dinar.* [DIEC]

dinar³ *m.* Moneda àrab d'or d'un pes aproximat de 4 grams, iniciada al segle VII. [DIEC]

Les diferències d'ordre fonètic sobre una mateixa forma gràfica no constitueixen, doncs, un paràmetre de distinció fiable en la divisió en articles del lèxic per a un diccionari català.

2.4 **Categorització gramatical**

Els diccionaris parteixen d'una llista definida de categories de classificació gramatical que corresponen en general a les anomenades *categories lèxiques* (o *parts de l'oració*, o *classes de mots*), i s'expressen en forma de marques abreujades entre els elements de la capçalera de l'article. Aquestes llistes no difereixen essencialment de les que pot donar una gramàtica tradicional.

Un dels principals problemes que planteja l'establiment d'aquests inventaris de categories és el nivell d'especificitat que han de tenir els elements que s'hi registren. Han de tractar diferenciadament fenòmens de tipus morfològic, com ara el caràcter invariable de gènere d'alguns adjectius?, o fenòmens de tipus sintàctic (transitivitat/intransitivitat, estructures argumentals, etc.)? o semosintàctic (comptabilitat dels noms, aspecte verbal, etc.)? Les solucions que la lexicografia convencional ha donat a aquesta qüestió varien segons l'època, el rigor metodològic, i la tradició de tractament gramatical en què es desenvolupen els diccionaris. Les possibilitats van des d'un marcatge absolutament simple a un marcatge molt complex, segons com es tracti la descripció dels diferents fenòmens dins de l'estructura del diccionari.

Una qüestió derivada de l'establiment d'un conjunt de marques de categoria és la interacció d'aquestes marques amb els criteris per a la diferenciació d'articles. És a dir, fins a quin grau de detall descriptiu cal arribar per a considerar l'existència d'un article diferenciat. Si un diccionari marca separadament els usos comptables i els no comptables dels substantius, quina incidència té això en el fet que un substantiu pugui aparèixer realitzat com a comptable i com a no comptable? O bé, si un diccionari integra dins de les categories verbals les marques de *transitiu/intransitiu*, què passa quan un verb pot aparèixer realitzat de totes dues maneres? I encara una altra qüestió, hi poden haver *compatibilitats categorials* entre marques diferents que comportin el fet de poder formar part d'un mateix article? (*adverbi i preposició*, etc.).

Els diccionaris convencionals solen utilitzar un doble marcatge: un que es produeix al nivell de l'article i un altre que es produeix al nivell de l'accepció. Determinades lleis d'herència implícites (o, més exactament, no explícites) regulen la projecció d'aquestes marques sobre els diferents components estructurals de l'article.

En un article com el d'*entendre* la marca *v.* actua com a categoria de l'article, mentre que les marques del tipus *intr.*, *trans.*, *pron.* es refereixen al nivell de l'accepció que codifiquen o sobre la qual es projecten:

entendre [quant a la flexió, com *atendre*] *v. intr.* Intervenir en un afer. *Ell no podia entendre en tals afers.* || *tr.* Tenir intenció (de fer alguna cosa). *Jo entenc donar-li la meitat de la meva fortuna.* || *tr.* Percebre amb la intel·ligència el sentit (d'alguna cosa), tenir la intel·ligència, una idea clara (d'alguna cosa). *Ell no entén el català. Entén el català, però no el parla. Perdoni: no l'he entès. És un home que no sap fer-se entendre. No us entenc: no sé el que voleu dir. Tu vas dir una cosa, i ell en va entendre una altra. Ja t'entenc: ja conec el que vols, el que intentes. Entendre les matemàtiques. Un zel mal entès. Amb mitja paraula ens entenem ell i jo.* | **entendre que** Jutjar que, opinar que. *Jo entenc que seria millor fer-ho ara (...)* || *pron.* ABS. Posar-se d'intel·ligència, anar d'acord. *Em vaig entendre amb el seu pare. Són dos que s'entenen. (...)* | POP. Dues persones, tenir relacions amoroses amagadament. || *intr.* **entendre en una cosa** Ésser-hi intel·ligent. *No hi entenc.* | *pron.* *No m'hi entenc.* [DIEC]

Aquest article correspon a aquesta estructura:

entendre v.

intr. Intervenir en un afer.

tr. Tenir intenció (de fer alguna cosa).

tr. Percebre amb la intel·ligència el sentit (d'alguna cosa), tenir la intel·ligència, una idea clara (d'alguna cosa).

tr. entendre que Jutjar que, opinar que. *Jo entenc que seria millor fer-ho ara.*

pron. ABS. Posar-se d'intel·ligència, anar d'acord

pron. POP. Dues persones, tenir relacions amoroses amagadament.

intr. entendre en una cosa Ésser-hi intel·ligent.

pron. No m'hi entenc.

La marca v. és la que intervé com a element diferenciador de l'article, mentre que la resta es poden considerar subespecificacions que concorren al nivell de l'accepció i que, en tot cas, determinen diferenciacions només en aquest nivell. Aquesta interpretació s'adequa a l'anàlisi del DIEC, i crec que essencialment també a la del DGLC i el DLC.¹⁵ Des d'aquest punt de vista, el primer dels dos nivells de classificació determina la naturalesa general dels articles i, per tant, intervé com a element distintiu en el procés de diferenciació de les entrades de diccionari. Les especificats semosintàctiques de cada un dels comportaments de les entrades corresponen així a una estructura subsidiària: la de l'accepció.

D'acord amb aquest marcatge en dos nivells, associats respectivament a un element estructural propi, es pot determinar que la llista de categories de primer nivell del DIEC correspon a la mostrada a la figura 2.3. El fet que hi hagi codis (com *m. pl.* en l'entrada *afores*) que, des del punt de vista tipogràfic i de la posició que ocupen en l'article, són susceptibles de ser considerats marques de primer nivell no invalida la classificació que proposo. Per altra banda, la utilització de codis que *a priori* són de primer nivell en les

¹⁵ Tot i que Lorente (1994) identifica, en el DGLC, la informació gramatical de la capçalera de l'article com una unitat d'informació i, així, distingeix entre codis principals (*v. tr.*, *v. intr.*, *v. pron.*) i codis secundaris (*Tr.*, *Intr.*, *Pron.*, *abs.*). Crec que aquesta interpretació és confusió, i no resol satisfactòriament la "repetició" de codis amb el mateix significat en tots dos nivells.

accepcions (*m.*, *f.*, etc.) només té per funció la restricció del conjunt de categories que es donen a la capçalera de l'article en una accepció concreta. Això tampoc no invalida l'anàlisi que proposo.

<i>adj.</i>	adjectiu
<i>adv.</i>	adverbi
<i>art.</i>	article
<i>conj.</i>	conjunció
<i>f.</i>	[substantiu] femení
<i>interj.</i>	interjecció
<i>m.</i>	[substantiu] masculí
<i>prep.</i>	preposició
<i>pron.</i>	pronom
<i>v.</i>	verb

Figura 2.3. Llista de categories lèxiques de primer nivell del DIEC

De fet, la presència d'un codi complex (que expressa conjuntament un marcatge de primer i de segon nivell) es deu al fet accidental que en certes entrades no concorren altres accepcions susceptibles de ser diferenciades per un codi de segon nivell; per aquest motiu, els diccionaris catalans solen associar-les —tipogràficament, però no conceptualment— a les informacions de capçalera de l'article. Considerat així, el DIEC reserva per al marcatge de segon nivell —el de l'accepció— els codis que recullo a la figura 2.4.

El conjunt d'elements del primer nivell conforma un inventari discret d'unitats que corresponen aproximadament a les categories lèxiques establertes per la gramàtica tradicional (nom, adjectiu, adverbi, preposició, verb, etc.). Hi ha factors, però, que estableixen clares diferenciacions respecte a aquest conjunt. Per exemple, en un diccionari d'una llengua com el català, els noms masculins i els femenins solen estar marcats distintament; per altra banda, l'existència de certes categories combinades com ara *adverbi* i *preposició* a l'encapçalament de l'article, sense que es tracti la combinació d'aquests dos codis com un factor per a la diferenciació d'articles, permet suposar que el lexicògraf considera l'existència implícita de categories *combinables*.

pl.	plural
intr.	intransitiu
pron.	[verb] pronominal
tr.	[verb] transitiu
tr. pron.	[verb] transitiu pronominal

Figura 2.4. Llista de codis de segon nivell del DIEC.

La informació de segon nivell, en el cas del DIEC (i també de la resta de diccionaris catalans, com el DGLC i el DLC), es refereix exclusivament a certes característiques de comportament verbal (transitivitat, intransitivitat, pronominalitat) i a les especificacions dels singulars o plurals lexicalitzats (*singularia tantum, pluralia tantum*). El caràcter hipoespecífic d'aquestes marques es fa evident, ja que un esquema de marcatge d'aquest tipus, si vol ser adequat des del punt de vista de la descripció lingüística, ha de contenir la informació necessària i suficient per a la caracterització global de la inserció sintàctica del mot d'entrada. Aquest és l'aspecte dels diccionaris en què més es posa de manifest el fet que la lexicografia convencional presenta en el seu mètode una sèrie de llacunes en la descripció que cal cobrir a partir del coneixement de la llengua que es pressuposa en el consultant de diccionaris.

Tenint en compte això, es palesa que el nivell més desatès tradicionalment per la lexicografia convencional és el de l'accepció. Una descripció lexicogràfica més adequada a les possibilitats i als avenços metodològics actuals hauria d'incloure en la seva microestructura informacions sobre qüestions diverses com ara:

- naturalesa semosintàctica dels substantius (comptable, massa, etc.)
- esquema de subcategorització de les categories predicatives (posicions sintàctiques, arguments, règims preposicionals, etc.)
- gradabilitat de l'adjectiu
- esquemes de predicació de l'adjectiu (*ser/estar*)
- estructures de subcategorització relacionades (alternances d'ergativitat, etc.)

Adjectiu
Adposició
Adverbi
Article
Classe de membre únic
Conjunció
Determinant
Interjecció
Nom
Numeral
Pronom
Residual
Verb

Figura 2.5. Llista de categories d'EAGLES

El fet de mantenir una distinció clara entre tots dos nivells de marcatge permet reduir la llista de codis del primer nivell i, sobretot, permet la diversificació potencial dels codis de segon nivell, d'acord amb el grau d'especificitat descriptiva que es vulgui assolir. Una aplicació regular d'aquests principis pot acostar la lexicografia convencional als tractaments que s'han determinat com a preferents per als lèxics electrònics, tal com s'ha establert a partir de les recomanacions d'EAGLES.¹⁶ EAGLES (1996) estableix una categorització en dos nivells semblant a la que aplica un diccionari convencional, però amb la particularitat que el segon nivell concentra una gran part de la caracterització morfosintàctica de l'entrada. El primer nivell aplica un etiquetatge molt semblant al que acabo de descriure per als diccionaris, tal com es pot veure a la figura 2.5. Les recomanacions d'EAGLES presenten l'avantatge afegit que han estat adaptades (d'acord amb el sistema d'especificacions del projecte), a un gran nombre de llengües europees, i que s'han convertit de fet en un estàndard en l'anotació de corpus i lèxics en format electrònic.

¹⁶ EAGLES (LRE1-6110) correspon a les sigles d'*Expert Advisory Group on Language Engineering Standards*. És una iniciativa de la Comissió Europea i ha reunit (entre 1993 i 1998) diversos experts de

D'acord amb els criteris de diferenciació d'entrades exposats, la categoria primària actua com a element de diferenciació d'articles en casos d'homografia entre formes canòniques.

Els diccionaris catalans no presenten, en aquest punt, coherència amb aquest principi, sinó que diferencien determinats casos en funció de la interpretació homonímica (*mel*¹, *mel*²) o polisèmica (*llum*) de cada cas:

llum *f.* Agent físic que és una de les formes d'energia, capaç d'impressionar els òrgans de la vista i que permet de veure els cossos. (...) || *m.* Estri o aparell que serveix per a fer llum. [DIEC]

mel¹ *m.* Os de la galta. [DIEC]

mel² *f.* Substància viscosa, dolça, elaborada per les abelles en una distensió de l'esòfag a partir del nèctar de les flors i que dipositen en les cavitats hexagonals de la bresca. [DIEC]

En la lexicografia convencional, la solució de tipus polisèmic assumeix implícitament dos principis, que són condició necessària per al tractament unitari de casos com el de *llum*: que els diferents sentits estan relacionats per algun tipus de vincle (sincrònic o diacrònic), i que hi ha categories “combinables” (nom masculí i nom femení, preposició i adverbi, etc.).

Un diccionari que pretengui la fixació dels criteris de delimitació d'articles des d'un punt de vista estrictament sincrònic ha de neutralitzar els aspectes de diferenciació diacrònica existents en casos com els de *mel* i *llum*, i partir de criteris de diferenciació estrictament funcionals. Alhora, ha d'analitzar l'existència de *categories lèxiques combinables* a partir de criteris que han de tenir en compte regularitats i fenòmens específics en el comportament sintàctic i semàntic de certes categories. Algunes d'aquestes regularitats es podrien concretar en casos com:

diferents àrees per al desenvolupament d'estàndards en l'àmbit del disseny, la creació i l'avaluació dels recursos lingüístics.

a) Els substantius que, sense tenir un referent de tipus sexuat, presenten tots dos gèneres (*amor, mar, etc.*)

b) Els substantius de referent sexuat que presenten una mateixa forma per a totes dues categories, o formes morfològicament relacionades (absència/presència d’-a) per a cada categoria

La lexicografia catalana convencional sol tractar tots dos tipus unitàriament: el tipus a), que és numèricament poc important, origina sempre un sol article que és categoritzat amb dos codis de primer nivell. El tipus b) sol ser tractat també unitàriament (ja sigui a través d’entrades complexes o simples), tret dels casos de noms d’animals amb formes diferenciades (*gos, gossa*), que originen articles diferents.

Una qüestió a tenir en compte des d’una òptica descriptiva (especialment en la lexicografia basada en corpus) són les dades aportades per un corpus que no concorden amb les solucions normatives de la llengua. El mot *calor*, per exemple, apareix en el CTILC un total de 3.472 vegades, 640 de les quals estan associades a formes en masculí. L’anàlisi del corpus ens ha de determinar si hi ha o no diferències objectivables entre una categoria i l’altra (condició *sine qua non* per a l’existència d’un nou article); si no hi ha diferències, el mot *calor* és susceptible de ser tractat, des d’una òptica estrictament descriptiva, com un cas de substantiu de doble gènere (com *amor, mar, etc.*).

Aquest tractament no està justificat en els casos en que la doble categoria porta implícita una diferència en el comportament sintàctic o en la descripció semàntica del mot.

El substantiu *sucre* ‘substància dolça’ presenta, en el CTILC, algun cas d’aparició en femení; pel que sembla, aquests casos són indiferenciables, des del punt de vista del seu contingut semàntic, dels que apareixen en masculí. Hi ha, però, un sentit de *sucre* ‘unitat monetària de l’Equador’ que no participa de l’ambivalència de gènere, sinó que té com a inherent la característica de masculí. Aquesta diferència de comportament pot

justificar (també des d'una òptica estrictament descriptiva i sincrònica) el tractament diferenciat d'aquest dos casos:

sucre *m., f.* Substància dolça (...).

sucre *m.* Unitat monetària de l'Equador.

Aquesta mena de factors que precisen els criteris de delimitació de l'article lexicogràfic són absolutament essencials per a la construcció de diccionaris més coherents des del punt de vista de la nomenclatura, i ajuden a precisar amb un grau de detall molt més alt tota la casuística del tractament lexicogràfic de les categories gramaticals (o dels grups de categories) de les entrades de diccionari.

2.5 Paradigma de flexió

Els diccionaris també incorporen, amb diversos graus d'explicitud, informació sobre el paradigma de flexió de cada lexema. Una qüestió bàsica i, en certa manera, prèvia al tractament d'aquesta informació és l'establiment dels límits precisos del concepte de flexió. Els aspectes que ha de tractar l'establiment d'aquests límits dins de la descripció lexicogràfica són de naturalesa diversa, com ara:

a) Quin tractament han de rebre els casos en què certes categories de flexió (com ara *singular/plural*) són susceptibles d'incorporar-se a elements lèxics per produir resultats que es poden considerar processos de diferenciació lèxica (*pluralia tantum, singularia tantum*)?

b) Quines són les categories de flexió que constitueixen l'objecte d'anàlisi: són fenòmens de flexió les diferències de cas en els pronoms personals (nominatiu *ell*, acusatiu *el*, datiu *li*, etc.)?

c) Quin tractament han de rebre els afixos no lèxics (diminutius, augmentatius, etc.) que s'adjunten a substantius i adjectius i que no solen ser considerats explícitament entrades de diccionari?

Totes aquestes qüestions han de ser determinades en les diverses opcions metodològiques que conformen un pla d'obra lexicogràfica; però tal vegada les que afecten més directament la problemàtica relativa a la macroestructura del diccionari són les del tipus a), és a dir, els casos de lexicalització de formes flexionals, tal com mostro més endavant en aquesta mateixa secció.

La flexió pot ser un factor determinant en la identificació dels elements macroestructurals del diccionari. En alguns casos es donen en català diferències entre mots que comparteixen forma gràfica i categoria gramatical, i que poden ser identificats com a unitats independents a causa del fet que cada un té una sèrie flexional diferenciada:

cas¹ *m.* Dels diversos esdeveniments, estats, circumstàncies, que es poden produir o donar, un d'ells, el que s'escau a produir-se, el que es dona. [DLC]

cas² *m.* Esquena d'un ganivet. [DLC]

cas³ [pl. *cassos*] *m.* Recipient de ferro, amb mànec, que s'utilitzava antigament per a fondre metall. [DLC]

En l'exemple precedent, les tres entrades tenen una mateixa forma gràfica i, a més, porten la mateixa categoria gramatical. Si ens atenim als criteris formals exposats a § 2.2, l'establiment de diferències objectives entre aquest grup d'entrades homògrafes justifica un tractament diferenciat de *cas¹* i *cas²*, per una banda, i de *cas³*, per l'altra. Aquest tipus de diferències han de ser tingudes en compte pel lexicògraf, i resulten l'element decisiu en la diferenciació d'un cert nombre d'entrades, com ara *do* 'nota musical' i *do* 'qualitat natural', o bé *fi* 'final' i *fi* 'lletra grega'.

Una qüestió més complexa relacionada amb la intervenció dels paradigmes de flexió en la diferenciació d'articles lexicogràfics és com es determinen els diferents tipus de models de flexió i quin tipus de fenòmens es consideren flexió i, en segon terme, quines

possibles interaccions hi ha entre l'esquema de models de flexió i l'esquema de caracterització gramatical d'un diccionari. Un cas particular d'aquesta qüestió el presenten els fenòmens d'assumpció amb caràcter lexical dels elements de flexió, que esdevenen inherents en determinats mots, com en el cas dels denominats *pluralia tantum* com ara *afores*, *acaballes*, etc.

Aquests casos solen rebre en els diccionaris catalans un codi gramatical (veg. més enrere § 2.4) que els marca com a representants d'una categoria gramatical amb la característica de plural inherent (*m. pl.* i *f. pl.*) i que s'associa a paradigmes de flexió incomplets. L'aplicació extensiva d'aquest tractament podria ser determinant per a justificar l'existència d'articles lexicogràfics diferenciats per a casos com *vacança/vacances*, *tisora/tisores*, etc. pel criteri de diferenciació bàsic exposat anteriorment (veg. més enrere § 2.2), que caracteritza l'entrada lexicogràfica diferenciada com la unió d'una forma canònica, una categoria gramatical i un model de flexió:

vacança *f.* Estat d'un càrrec, d'una plaça, etc., vacants. *La vacança d'una càtedra.*
|| *pl.* Temps de repòs durant el qual cessen els estudis d'una escola, els treballs d'un tribunal, etc.; temps de repòs concedit a un funcionari o a un treballador. *Les vacances de Nadal. Estar algú de vacances. Passar les vacances al camp.* [DIEC]
tisora *f.* Qualsevol de les dues peces d'unes tisores. (...) || *pl.* Instrument de tallar que consisteix en dues làmines tallants amb mànecs, entrecreuades i unides en el punt d'entrecreuament per un piu, al voltant del qual poden girar. [DIEC]

Des de l'òptica de la lexicografia convencional, aquest tractament presenta certes coincidències amb el que es dona als casos d'acumulació categorial en el gènere dels substantius que he analitzat anteriorment (veg. més enrere § 2.4). Els plurals lexicalitzats es tracten a la microestructura sempre que sigui possible d'associar-los a una entrada en forma canònica en singular; només en el cas que el mot en singular no sigui viable, la categoria de plural s'integra com una marca de primer nivell referida a l'entrada. La justificació metalexigràfica d'aquest tractament és, però, molt més convincent, i més adaptable a una lexicografia basada en les evidències que proporcionen els corpus.

Ara bé, els diccionaris catalans presenten vacil·lacions en l'aplicació d'aquests criteris. Podem veure els diferents aspectes de la problemàtica que comporta el tractament de les lexicalitzacions en plural de determinats mots a través d'una sèrie sinonímica complexa del tipus següent: *boda*, *casament*, *noces*, *núpcies*, *nuviances*. El DIEC dona per als mots d'aquesta sèrie les següents entrades i definicions:

casament *m.* Acció de casar-se; l'efecte; la cerimònia de contreure matrimoni. (..)

[DIEC]

boda *f.* Casament. | [usat generalment en pl.] Festa amb què se celebra un casament, noces. (...). [DIEC]

noces *f. pl.* Boda, núpcies. (...) [DIEC]

núpcies *f. pl.* Casament. (...) [DIEC]

nuviances *f. pl.* Noces. (...) [DIEC]

En general, la interpretació d'aquests casos és trivial: els *pluralia tantum* incorporen la marca *pl.* a la categoria gramatical; tret del cas de *boda*, que presenta una peculiaritat consistent a introduir una nota gramatical (“usat generalment en pl.”) en la segona accepció. El diccionari registra aquesta accepció sota de l'entrada *boda* seguint el criteri que hem vist per a *vacança/vacances*, però en aquest cas no es determina explícitament que només és possible l'aparició en plural de la segona accepció; la interpretació que sembla que cal donar a aquesta manera de representar les dades és doble:

a) que no hi ha distinció entre els sentits de singular i de plural (és a dir, que el plural i el singular poden designar un mateix referent), i

b) que la utilització del plural és més habitual (o preferible?) que la del singular.

Aquesta és la manera com tradicionalment la lexicografia catalana convencional resol els casos de coincidència d'un mot pertanyent a la classe dels *pluralia tantum* (*duia els bigotis bruts*, etc.) amb la denotació múltiple del plural del mateix mot (*dos bigotis postissos*, etc.); la fórmula “usat generalment en pl.”, però, té l'inconvenient que pressuposa, més enllà de l'equivalència significativa, una major freqüència d'aparició de la forma plural, cosa que no té per què produir-se necessàriament en aquest tipus de mots.

Una descripció lexicogràfica més coherent hauria de dissociar totes dues informacions, en el sentit que ha de determinar separatament la propietat lingüística d'aquests fenòmens i les dades de tipus extralingüístic. Pel que fa a les dades de tipus lingüístic, això es pot aconseguir per diverses vies:

a) donant entrada pròpia a *bodes* convertint-lo en una entrada més de la sèrie sinonímica.

b) assenyalant, al nivell de l'accepció corresponent de l'article *boda*, el fet que hi pot haver indistinció entre *boda* i el plural lexicalitzat *bodes*.

La solució a) no té precedents en la lexicografia catalana, i a més presenta l'inconvenient que disgrega la informació d'una manera no gaire aconsellable des del punt de vista de la praxi lexicogràfica (es pot localitzar una mateixa forma gràfica, corresponent a dues modalitats lèxiques molt relacionades semànticament, en dues entrades de diccionari).

Una lexicografia basada en corpus ha de tenir en compte, a més, certes dades complementàries que sovint aporten els textos. D'acord amb el DFCC, al CTILC apareixen 18 casos en singular lligats al lema **noces** *fp*, la qual cosa afegeix un nou aspecte al problema de l'establiment de la forma de les entrades, ja que una anàlisi descriptiva que no tingui en compte els aspectes normatius ha de determinar per a **noça/noces** (no reconegut normativament) un tractament anàleg al de **boda/bodes**.

En ocasions els *singularia tantum* (*gratitud, honestedat, sacerdoci, catalanitat*) s'han interpretat també com a fenòmens de paradigmes incomplets, i per tant s'han considerat lexemes associats a un model flexional defectiu (sense plural). Les capacitats de pluralització (en determinats contextos) d'aquest tipus de mots (Mascaró, 1986: 91) i el fet que no presenten cap especificitat de tipus morfològic que determini el fet que no puguin incorporar plural, fa que aquest fenomen s'hagi de considerar com un tipus de defectivitat semàntica i no pas morfològica (Mel'čuk, 1993).

La lexicografia convencional no tracta de cap manera específica aquests casos. El tractament generalitzat és la indiferenciació respecte dels mots que no presenten aquesta característica. En part, aquest tractament està justificat pels fets que acabo d'assenyalar (motivació semàntica de la defectivitat), i per l'extrema facilitat amb què es produeixen els canvis semàntics que conduirien a un replanteig constant de les informacions relatives a la nomenclatura del diccionari.

Un possible tractament lexicogràfic “objectiu” d'aquesta qüestió consisteix en la inclusió en el diccionari d'informació relativa a les freqüències d'aparició de cada una de les categories morfològiques del paradigma de flexió del mot. D'aquesta manera, per exemple, podríem determinar que mots com ara *gratitud*, *honestedat*, *sacerdocí*, *catalanitat*, etc. no presenten en el CTILC cap ocurrència de plural, però alhora poden ser associats a un paradigma de flexió que forma el plural a partir de l'afegiment de *-s* a la forma del singular.

Finalment, adduiré un cas força específic que mostra clarament la intervenció del paradigma de flexió en el procés de diferenciació d'articles. Es tracta de les modalitats de flexió que estan exclusivament lligades a qüestions de caràcter semàntic. Aquest fet, tot i que força infreqüent, pot trobar-se en català a partir de l'exemple d'un mot com *coure*:

coure² [ger. *coent*; p. p. *cuit*; ind. pr. *coc*, *cous*, *cou*, *coem*, *coeu*, *couen*; subj. pr. *cogui*, etc.; subj. imperf. *cogués*, etc.] v. tr. Preparar (un aliment) mitjançant l'acció del foc, sia bullint-lo o rostint-lo, socarrant-lo, etc.; fer apta o més apta per a l'alimentació (una substància) mitjançant l'acció del foc o de la calor. *Coure el dinar*, *el sopar*. *Coure carn*, *llegums*, *la vianda*. *Coure el pa*. (...) || [p. p. *cogut*] Produir una sensació anàloga a la que causa l'acció del foc, produir coïssor. *Els ulls em couen*.

L'accepció d'aquest verb descrita com “produir una sensació anàloga a la que causa l'acció del foc, produir coïssor” està, com es pot veure, marcada per la diferenciació de les formes de participi de passat (*cogut*) amb què es distingeix de les accepcions que utilitzen la forma *cuit*. D'acord amb els criteris de diferenciació d'articles, aquest cas és susceptible de ser tractat en dues entrades de diccionari diferents. Aquest tractament no

té precedents en cap diccionari català, però és conseqüència natural dels criteris que he exposat a § 2.2, ja que *coure* (participi de passat *cuit*) i *coure* (participi de passat *cogut*) són associables a dos paradigmes de flexió diferenciables i, a més, no es produeix cap mena de vinculació semàntica entre tots dos usos.

2.6 Entrades hipolexemàtiques

Els diccionaris solen tractar en la seva nomenclatura diferents tipus d'elements que no es corresponen estrictament amb el concepte de *lèxic* (Coseriu, 1987). La tipologia d'aquests elements és variada, però en especial m'interessa destacar certs tipus extrems que, més o menys habitualment, han estat objecte de tractament en els diccionaris convencionals. Aquests tipus són els següents:

- Clítics
- Afixos lèxics
- Irregularitats flexionals
- Marques de flexió quasilèxiques
- Categories de semantemes

Agrupo aquest conjunt de classes sota la denominació global d'*entrades hipolexemàtiques*, i al llarg d'aquesta secció faré una presentació particularitzada de cada una. El fet de tractar-les separatament està motivat perquè la presència dels representants de cada classe en els diccionaris correspon a diferents raons metodològiques i, també, perquè el seu tractament lexicogràfic és força diferenciat.

Les descripcions referides a les entrades hipolexemàtiques són, molt sovint, explícites a través d'un enunciat que no estableix entre definit i definició una relació de sinonímia, sinó que la definició sol fer el paper d'explicació de les propietats gramaticals del mot. Aquestes definicions estan expressades en allò que Rey-Debove (1971) anomena *metallengua de signe*, i es corresponen a allò que Martin (1977a) anomena definicions *metalingüístiques* (veg. més endavant § 4.1). Aquest tipus de definicions, que aplicades

a entrades lexemàtiques s’han vist com a desviacions metodològiques,¹⁷ no es basen en la descripció del contingut conceptual del definit, sinó que expliciten el seu funcionament gramatical, i integren elements que es refereixen al mot en tant que signe; expressions com *denota*, *dit de*, *és emprat*, etc., que no tenen per funció la descripció de la significació lèxica sinó, seguint Mel’čuk (1993), de la significació sintàctica.

El tractament lexicogràfic de cada un dels tipus exposats és diferent; aquestes diferències es deuen no només a la casuística que presenta cada fenomen, sinó també a les diferents tradicions lexicogràfiques (corresponents a les diferents llengües), que recullen aquestes especificitats de manera diversa.

2.6.1 Clítics

Els clítics i els mots buits s’incorporen en general a tots els diccionaris, normalment descrits a través d’un enunciat metalingüístic:

- ho** *pron.* És emprat: **1** Com a complement directe neutre, reemplaçant *açò*, *això* o *allò*. (...) **2** Reemplaçant un adjectiu o un substantiu indeterminat complements de *ésser*, *estar* o *semblar*. (...) [DIEC]

En alguns casos, aquestes entrades van acompanyades d’especificacions que descriuen aspectes del seu comportament en funció del context (com ara les formes que pot revestir en combinació amb d’altres unitats) o donen informacions complementàries a la descripció:

¹⁷ El fet que en el cas dels clítics hi hagi una motivació per a la inclusió d’elements metalingüístics en les definicions hauria de matisar la diferenciació que fa Seco (1978) entre definicions “pròpies” (parafràstiques, en la terminologia de Martin, 1977a), d’una banda, i definicions “impròpies” (o metalingüístiques), de l’altra. En el cas de certes classes gramaticals, la incorporació d’aquest tipus d’elements és l’única via possible per a la caracterització del mot definit, la qual cosa ens porta a la constatació que es poden distingir diversos graus d’impropietat en les definicions lexicogràfiques: mentre que en alguns casos la definició metalingüística està lligada a la naturalesa gramatical del mot, en d’altres

el² [pl. *els*; pot revestir la forma *l'*; precedit de les preposicions *a*, *de* i *per* dóna lloc a les formes contractes *al*, *del* i *pel*, i *als*, *dels* i *pels*, respectivament] *art.* Forma masculina de l'article definit (...) | En el llenguatge col·loquial, pot anteposar-se als noms propis de persona.(...) (Reemplaçat, però, normalment, en el català central, per *en* davant els noms començats en consonant.) [DIEC]

Algunes categories de mots buits (com els adverbis de partícula) poden tenir, en alguns casos, definicions sinonímiques, com la de l'exemple d'*ara*. En general, però, predominen les definicions metalingüístiques, com la de l'exemple d'*encara*.

ara² *adv.* En aquest moment. (...) [DIEC]

encara *adv.* Denota que en tal moment una cosa continua tenint lloc; i, amb negació, que en tal moment una cosa no ha tingut ja lloc. (...) Seguit d'un verb subordinat amb negació, denota que l'acció expressada pel verb principal s'esdevé immediatament després d'haver començat o d'haver-se acomplert la del verb subordinat. [DIEC]

Malgrat les objeccions que la lingüística teòrica hi pot plantejar (a causa del paral·lelisme que s'estableix entre els diccionaris i el conjunt lèxic de la llengua) la presència d'aquesta mena de mots no lexemàtics està motivada per factors que no només tenen a veure amb la tradició lexicogràfica, sinó també amb les necessitats d'una descripció adequada del lèxic.

En els articles referents als clítics, la definició lexicogràfica és l'únic element estructural de l'article que opera canvis fonamentals en la seva expressió, respecte de les entrades lexemàtiques. Cal veure aquest canvi com una adaptació al tipus d'entitat que es descriu. Un diccionari també s'ocupa d'altres aspectes de la caracterització lexicogràfica del mot (estructura argumental, col·locacions, fraseologia, etc.), que es mantenen sense canvis d'importància en els articles referits als clítics. La preposició *entre*, per exemple, té un esquema de selecció argumental (*entre X i Y*, o *entre dos X*) que pot ser descrit amb el mateix tipus de formalisme que el que utilitzem per a

(en els mots *plens*) la presència d'elements gramaticals en les definicions lexicogràfiques es pot interpretar com una distorsió.

qualsevol altra classe gramatical; l'adverbi *encara* forma part d'un cert nombre de locucions (*encara ara*, *encara que*, etc.), que tenen entrada al diccionari a través del seu article.

2.6.2 Afixos lèxics

En la lexicografia catalana és habitual la presència com a articles de diccionari de constituents de mot (prefixos, sufixos). En el DGLC, el DLC i el DIEC, es dona entrada a constituents lèxics de tipus culte que intervenen en la formació de mots en llenguatges d'especialitat o en processos més o menys productius. Això “permet” al lexicògraf l'exclusió d'un conjunt de lèxic que presenta un nombre d'unitats relativament elevat. El tipus d'elements incorporats no formen part, en general, de la llista de morfemes que intervenen en els processos habituals de formació de mots a partir de radicals patrimonials.¹⁸ El DGLC, per exemple, recull amb entrada pròpia casos com els següents:

al·lo-, forma prefixada del mot gr. *allos* altre; prefix usat algun cop en quím. per a designar l'un de dos isòmers. Ex.: *al·lomorf*, *al·lopatia*; *àcid al·locinàmic*. [DGLC]

amfi-, prefix que significa ambdós, de dues maneres, a ambdós costats, a l'entorn. [DGLC]

aminoazo-, prefix que denota la presència, en una mateixa molècula, dels dos grups -NH₂ i -N₂. Ex.: *aminoazobenzè*. [DGLC]

an-, forma secundària del prefix *a-* i del prefix *ana-*. [DGLC]

ana-, prefix grec que significa *contra*, *amunt*, *de nou*. [DGLC]

anglo-, prefix que significa anglès, relatiu a Anglaterra. Ex.: *anglòfil*, *anglòfob*, *angloamericà*, *anglonormand*, *anglosaxó*. [DGLC]

ante-, prefix que denota anterioritat quant al temps o a la posició. Ex.: *antediluvià*, *antepenúltim*. [DGLC]

-adelf, **-adelfa**, sufix usat en bot. per a significar que els estams, units per llurs filaments, són disposats formant un, dos, etc. manyocs o fascicles (*monadelf*, *diadelf*, etc.). [DGLC]

-àlgia, sufix que vol dir dolor. Ex. *cefalàlgia*, *neuràlgia*. [DGLC]

¹⁸ El GDLC constitueix una excepció a aquesta tendència, ja que incorpora a la seva nomenclatura afixos de tots els tipus.

-càrpic -a, sufix per a indicar, d'una planta, que té tal nombre o tal mena de fruits.

Ex.: *monocàrpic, apocàrpic*. [DGLC]

La presència de formants (especialment de formants savis) és característica de certes tradicions lexicogràfiques. En el cas de la lexicografia catalana, aquesta tendència es pot vincular a la tradició lexicogràfica francesa, que fa que els diccionaris actuals recullin encara aquests tipus d'elements. El DPR, per exemple, segueix aquest sistema però ho justifica amb la següent observació:

la nomenclature contient aussi des éléments de formation des mots savants. La présence de ces unités est destinée à expliquer la formation de termes récents (du dictionnaire ou hors du dictionnaire et non à repertorier les éléments (...) du français. [(DPR, *Préface*, p. XI)]

Malgrat les diferències que es poden establir entre les tradicions lexicogràfiques europees, s'observa, doncs, com la majoria de diccionaris conflueixen en la incorporació de radicals lligats de tipus savi que intervenen en els processos de formació de mots de la llengua. Altres diccionaris de diferents llengües i elaborats a partir de metodologies diverses com ara el VLI, el DEA, el DUE, el DRAE, el COBUILD, el LDOCE, etc. també recullen en la seva nomenclatura aquests elements.

La normalització metodològica de l'activitat lexicogràfica demana l'adopció de criteris clars i sistemàtics en aquest sentit. El fet d'incorporar a un diccionari, com a entrades, constituents que no tenen el funcionament autònom que caracteritza el lèxic no sembla prou adient, en el sentit que constitueix una barreja de plans de descripció.

Una opció metodològicament més acceptable consisteix a reservar el concepte de nomenclatura del diccionari al conjunt de mots (no pas de constituents de mot) que es descriuen i, explicitar fora de la nomenclatura pròpiament dita (en quadres específics o en estructures més complexes que ho permetin) els processos de formació de mots en què intervenen determinats constituents.

Un diccionari, per la seva pròpia naturalesa, ha d'assumir una aproximació *lexicalista* a la descripció de la llengua. De fet, aquesta és l'opció adoptada per obres de publicació recent, com el DMLC, que no recull en la seva nomenclatura aquest tipus d'entrades.

2.6.3 Irregularitats flexionals

En alguns casos els diccionaris recullen com a tipus d'articles específics algunes formes del paradigma de flexió d'una entrada. L'entrada està constituïda per la forma flexional del mot, i el contingut de l'article ens remet a una entrada que podem considerar principal. Aquesta pràctica és general en els diccionaris que s'adrecen a l'aprenentatge d'una llengua:

took (...) *past simple of TAKE*. [CIDE]

took (...) *past tense of TAKE*. [LDOCE]

took (...) is the past tense of **take**. [COBUILD]

En el cas de la llengua anglesa, aquest tractament es veu afavorit per la limitació que presenta el seu sistema de flexional, juntament amb la presència d'irregularitats molt específiques i de sistematització difícil.

En la tradició lexicogràfica catalana aquesta és una pràctica força inusual. Fins i tot en els diccionaris d'orientació escolar (*DIDAC, Primer Diccionari*), hom parteix de la base que les descripcions de la flexió de cada entrada ja descriuen les “irregularitats” en els models flexionals, i que la competència lingüística de l'usuari del diccionari permet obviar-ne el tractament com a entrades secundàries. El fet que el nombre d'irregularitats sigui més alt que en d'altres llengües, i la dificultat de fixació del concepte d'irregularitat (gran nombre de paradigmes de flexió, existència de subparadigmes, etc.) fan que en el cas del català cap diccionari manual no incorpori aquesta informació dins de la nomenclatura del diccionari.

Un cas que té punts de contacte amb les irregularitats flexionals és el de les variants de tipus contextual que donen lloc a formes de mots adaptades a l'entorn en què apareixen. El nombre d'aquest tipus de fenòmens, molt més reduït que el cas anterior, i les

especificitats que solen presentar en cada un dels casos fa que en els diccionaris catalans sigui habitual el tractament de cada una d'aquestes variants dins de la nomenclatura; les diferents formes que solen revestir els pronoms personals febles hi solen ser tractades així:

em *pron.* [pot revestir les formes *m'*, *'m* i *me*.] Acusatiu i datiu del pronom de primera persona *jo*. *Em sents? No em donis res.* [DIEC]

me *pron.* Forma que revesteixen l'acusatiu i el datiu del pronom de primera persona *jo*, quan, no seguint-lo immediatament algun dels dos pronoms febles *hi* i *ho*, es troba immediatament darrere un verb acabat en consonant o *u* o davant un altre pronom feble. *Escriu-me. Empeny-me. No me la prenguis. Dóna-me'l.* [DIEC]

m' *pron.* Forma que revesteixen l'acusatiu i el datiu del pronom de primera persona *jo* quan precedeix immediatament un verb començat en vocal o *h* o l'un dels dos pronoms *hi* i *ho*. *No m'escolta. M'ha donat la mà. No m'ho diguis.* [DIEC]

'm *pron.* Forma que revesteixen l'acusatiu i el datiu del pronom de primera persona *jo* quan, no seguint-lo un segon pronom feble, va immediatament darrere d'un verb acabat en vocal altra que una *u* subjuntiva o d'un pronom feble acabat en vocal. *Escolta'm. Dóna'm la mà. (Però: Dóna-me-la, dóna-m'ho.) Va posar-se'm malalt el menut.* [DIEC]

En algun cas similar que origina una entrada hipolexemàtica, aquesta es pot trobar associada a una entrada lexemàtica, com per exemple la forma contextual *bon*, que en el DIEC és registrada a part de l'adjectiu *bo*:

bon² *adj.* Forma que revesteix l'adjectiu *bo* quan precedeix un substantiu o un infinitiu. [DIEC]

Es poden trobar exemples de tractaments anàlegs en diccionaris anglesos:

-**'m** is a short form of 'am' that is used after the pronoun 'I' in spoken English, and in informal written English. [COBUILD]

-**'m** (...) *short form*: am. [LDOCE]

La resolució pràctica d'aquests darrers casos és, doncs, l'establiment d'una relació entre la variant i l'entrada considerada principal. Atès que en tots els casos (*em*, *m'*, ... , *bon*) es tracta de formes alternatives determinades pel context, aquesta solució, en línies generals, és coherent amb la realitat de la llengua. Cal fer notar, però, que en el cas de

l'exemple del pronom *em* tots els articles fan referència al pronom *jo*, i no a l'entrada *em*, que és la que recull les variants dins del seu article (i que per tant és considerada com la forma canònica de tota la sèrie). Tal vegada fóra més adequat establir la relació amb el pronom de primera persona únicament a través de l'element canònic *em* i tractar la resta com a variants contextuais d'aquest, sense la represa dels elements descriptius que suposa la referència a *jo*.

2.6.4 Marques de flexió quasilèxiques

Els diccionaris de vegades també recullen elements que pertanyen a dues *dimensions* diferents. D'una banda són els morfemes flexionals de temps, persona, etc. que caracteritzen la forma lèxica que acompanyen; de l'altra són elements que no formen part indissociable del mot com a morfemes flexionals, sinó que tenen un cert grau d'autonomia. En el cas del català podem trobar algun cas d'aquest tipus en les formes verbals auxiliars per a la formació del pretèrit perifràstic. Aquestes formes (*vaig, vas/vares, va, vam/vàrem, vau/vàreu, van/varen*) pertanyen a una sèrie verbal defectiva, sense forma d'infinitiu i que és única en el sentit que no es correspon amb la de cap altre verb.

La lexicografia pràctica no presenta una solució gaire satisfactòria ni universal en aquests casos, que es poden considerar problemes no estrictament lèxics, sinó de descripció morfològica de les llengües. En el cas dels diccionaris catalans, ni el DGLC ni el DLC recullen aquestes formes en la seva nomenclatura, però sí que ho fa el DIEC, que recull aquestes formes sota una única entrada *va*:

va¹ [*vaig, vas o vares, va, vam o vàrem, vau o vàreu, van o varen*] Auxiliar del pretèrit perfet perifràstic d'indicatiu. *Vaig caure.* || [*vagi, vagis, vagi, vàgim, vàgiu, vagin*] Auxiliar del pretèrit perfet perifràstic de subjuntiu. *Em va agradar que el vagis fer callar.* [DIEC]

En la majoria de diccionaris anglesos es produeix la inclusió, de forma sistemàtica, d'entrades que es poden relacionar amb fenòmens força comparables:

will (..) *v 3rd person sing. (...)* **1** (Used for expressing the simple future tense);
They say that it will rain tomorrow. [LDOCE]

will (...) **1** You use **will 1.1** to refer to something that is going to happen in the future. (...) **1.2** with ‘I’ and ‘we’ when you are saying that you intend to do something. [COBUILD]

2.6.5 Categories de semantemes

Altres casos, força inhabituals, d’entrades de diccionari que no constitueixen pròpiament lexemes són aquelles unitats que comporten un grau de significació prelexemàtica i que no són contínues des del punt de vista lineal, de manera similar a les *categories de semantemes*, que fan referència al simbolisme de la forma, exposades per Hjelmslev (1928):

t... c (var. “t... ch”, “t... g”) Grupo imitativo o expresivo de sonidos (consonante dental más palatal o velar), que se encuentra en un grupo de palabras que en unos casos significan golpe seco y, en otros, se aplican a objetos de forma maciza y tosca: ‘bitoque, bodoque, currutaco, estaca, pataca, patacón, pataqueta, potoco, retaco, tac, tac tac, taco, tacón, taque, taragallo, tarugo, tic, tic-tac, tocar, tocho, tocón, traca, trangallo, trueco’. ⇒ SONIDO expresivo, t... r... c [o tr... c], z... c. [DUE]

2.7 **Relacions entre macroestructura i microestructura**

Els criteris que he establert (veg. més enrere § 2.2) per a la diferenciació d’articles en la praxi lexicogràfica han estat aplicats pels diccionaris amb diferents graus de sistematicitat, de manera que és freqüent trobar casos de diferències de tractament en qualsevol dels tres atributs que caracteritzen l’entrada. La lexicografia ha de determinar en un sistema de descripció homogeni la caracterització del component lèxic de la llengua, i aquesta anàlisi adquireix un grau d’exhaustivitat quantitativa pel que fa al nombre d’unitats tractades que la lingüística teòrica no ha abordat mai.

Un diccionari ha de sistematitzar tots els aspectes relacionats amb la determinació de les unitats lèxiques des del punt de vista formal, i resoldre la problemàtica específica que presenta cadascuna de les propietats que determinen les entrades, és a dir:

- a) ha de presentar el lèxic com una llista de mots,
- b) ha de presentar l'esquema de categorització gramatical com una llista tancada i global de codis, i
- c) ha de presentar els diferents comportaments flexionals com una llista de models.

No és sorprenent, doncs, per raó de les especificitats que presenta cadascuna d'aquestes operacions, de trobar inconsistències entre els diccionaris, i fins i tot dins d'un mateix diccionari. Aquestes inconsistències estan en funció de la diferent distribució entre el tractament macroestructural i el microestructural de determinats casos que presenten diverses possibilitats d'anàlisi en la determinació d'aspectes gràfics, morfosintàctics o flexionals.

Els mecanismes morfològics (la derivació, la composició) que actuen sobre el lèxic pot condicionar també, des del punt de vista del diccionari, l'inventari de les unitats. L'existència d'aquests mecanismes regulars produeix en algunes ocasions que el lexicògraf hagi de determinar en quin eix d'organització del diccionari cal tractar aquests fenòmens. Així, una de les relacions més importants que es poden produir entre la macroestructura i la microestructura ve donada pel fet que el diccionari sigui de nomenclatura *simple* (sense tenir en compte la derivació; és a dir, amb un sol nivell d'organització) o *complexa* (tractant les regularitats derivatives al nivell de la microestructura de l'article; és a dir, amb dos nivells d'organització).

La macroestructura¹⁹ d'un diccionari actua en el pla del lèxic de la llengua, de l'inventari de les unitats de base. La microestructura²⁰ actua en el pla de la descripció de les característiques funcionals, semàntiques, sintagmàtiques, pragmàtiques, etc. de les entrades. El terreny en què més clarament es pot percebre la interacció entre macroestructura i microestructura és en la interpretació del fet que diversos significats tenen, en ocasions, formes significants coincidents entre si.

Qualsevol aproximació, teòrica o pràctica, al problema de les unitats del lèxic ha de resoldre aquesta correspondència múltiple a partir de dues possibilitats extremes de solució: considerar que la relació entre significants és d'*homonímia*, o bé considerar que la relació entre significats i significant és de *polisèmia*.²¹

Tradicionalment, l'homonímia s'interpreta com una relació que es produeix entre diferents mots (o lexemes), mentre que la polisèmia és una propietat dels lexemes en ells mateixos. Tal com fa notar Lyons (1977, 1995), l'existència de l'homonímia a nivell lexemàtic es fonamenta sobre tres requeriments: la nul·la relació de significat entre els lexemes homònims, la identitat de paradigma de flexió, i l'equivalència gramatical de totes llurs formes entre si.

Dubois & Dubois (1971) aborda la dimensió lexicogràfica d'aquesta qüestió assenyalant que el lexicògraf pot recórrer a dues hipòtesis per explicar els diversos tipus de relació

¹⁹ Crec que aquesta formulació inicial de Rey-Debove (1971) és prou clara, i que la suposada ampliació que, segons Martínez de Sousa (1995), en fa Haensch (1982b: 452) ja s'hi troba plenament desenvolupada: "On appellera MACROSTRUCTURE l'ensemble des entrées ordonnées, toujours soumise à une lecture verticale partielle lors du repérage de l'objet du message". [Rey-Debove (1971: 21)]

²⁰ "On appellera MICROSTRUCTURE l'ensemble des informations ordonnées de chaque article, réalisant un programme d'information constant pour tous les articles, et qui se lisent horizontalement à la suite de l'entrée (l'ordre des informations permet, au mieux, une consultation interne)". [Rey-Debove (1971: 21)]

²¹ El terme *polisèmia* fou introduït en lingüística per Bréal, per fer referència al fenomen diacrònic de l'addició de noves accepcions al sentit general d'un mot; el mateix Bréal, però, féu notar la dimensió sincrònica de la polisèmia en funció de la coexistència de diversos exponents en un sol terme.

existent entre significants i significats, i anomenen aquestes dues hipòtesis en funció del resultat que provoquen:²²

a) La hipòtesi polisèmica: és a dir, la identificació del mot-entrada i forma gràfica, de manera que cada entrada tingui una forma gràfica diferenciada; els sentits diferents (si n'hi ha) constitueixen el conjunt de paràfrasis de l'entrada-mot.

b) La hipòtesi homonímica: és a dir, la consideració que la descripció del mot d'entrada es basa en un únic sistema de paràfrasis sinonímiques (un sol sentit).²³

La determinació dels límits entre homonímia i polisèmia es correspon plenament a l'explicitació dels criteris d'organització macroestructural i microestructural d'una obra lexicogràfica. Una estructuració fortament macroestructural determinarà el predomini de l'homonímia sobre la polisèmia; mentre que una estructuració fonamentalment microestructural farà predominar l'anàlisi polisèmica.

Les discussions sobre la diferenciació entre l'homonímia i la polisèmia no han aportat conclusions de valor a la lexicografia,²⁴ i el tractament que es dona a aquests fenòmens en els diccionaris no reflecteix més que una posició metodològica que no té incidència

²² Dubois & Dubois (1971) recull això de la següent manera: “Lorsque deux entrées, distinctes sémantiquement sont identiques graphiquement, on dit qu'elles sont homonymes. Lorsque la même entrée recouvre des sens différents, on dit que le mot est polysémique. Il ne s'agit pas là simplement d'une disposition graphique différente, mais de deux conceptions du «mot» totalement distinctes. Dans la première conception, le mot est une *unité du discours* définie par son contexte: situation et distribution. Dans la deuxième conception, le mot est une *unité de la langue*, dont les réalisations dans le discours impliquent des variations de sens selon les contextes.” [Dubois & Dubois (1971: 67)]

²³ La realització pràctica d'aquesta segona possibilitat no és gaire freqüent, ja que hi ha tantes entrades com sentits: “on trouve cette analyse dans les TREVOUX du XVIII^e siècle, par exemple. Elle représente une forme extrême de correspondance exacte entre la distribution d'un terme et sa signification et a pour conséquence une multiplication des entrées homographes” [Dubois & Dubois (1971: 67)]

²⁴ D'acord amb Casas & Muñoz (1992), la distinció tradicional entre homonímia i polisèmia no té, en darrer terme, una justificació clara en una perspectiva sincrònica de la lingüística i, per tant, només és una distinció operativa en diacronia.

fora de l'àmbit lexicogràfic. La diferenciació d'homònims en funció de l'ètim no és metodològicament defensable,²⁵ si partim de la base que una obra lexicogràfica és (o ha de pretendre ser) una descripció sincrònica del lèxic. Aquesta diferenciació, a més, sol comportar diferències de tractament entre diccionaris homogenis, o fins i tot dins d'un mateix diccionari. S'ha intentat justificar la diferenciació entre homonímia i polisèmia a partir de la possibilitat d'establir relacions d'afinitat entre els significats. En principi, dos significats relacionables corresponen a una organització polisèmica, mentre que dos significats no relacionables corresponen a una organització homonímica. En molts casos l'explicitació d'aquest criteri no ha passat d'una formulació *preteòrica* (Lyons, 1977) que presenta nombroses dificultats d'aplicació: l'afinitat de significat és en darrer terme una qüestió de grau i rara vegada les intuïcions coincideixen plenament entre si.

Tot i la diferència de tractament que impliquen les dues hipòtesis assenyalades per Dubois & Dubois (1971), potser les diferències *reals* entre les capacitats descriptives d'una i altra no són tan radicals. Malgrat la diferència en el pla lingüístic, tant l'homonímia com la polisèmia no són sinó tipus d'ambigüitat; fins i tot es poden considerar ambigüitats no essencialment diferents des del punt de vista de la descripció que en poden donar els diccionaris.

En la distinció dels tipus d'ambigüitat, s'ha introduït la idea que en un lexema:

- poden concórrer dos o més sentits sense cap relació entre si, o bé
- que els diferents significats que pot adquirir un lexema són de fet manifestacions particulars d'un significat bàsic degudes als contextos en què apareix.

²⁵ No crec que valgui gaire la pena de discutir en profunditat aquesta qüestió; sobre la qual Lyons es mostra categòric: "Useful though it may be to have readily accessible in our standard dictionaries whatever information is available about the origins and history of particular words, this information is, or should be, irrelevant in the synchronic analysis of languages". [Lyons (1977: 551)]

Weinreich (1964) anomena el primer cas *contrastive ambiguity*, i el fa correspondre a la denominació general d'homonímia, mentre que es refereix al segon com a *complementary polysemies*.

Apresjan (1974) distingeix l'homonímia de la polisèmia en el següent sentit:

the existence in a given language of at least a pair of words with a given combination of meanings is considered to be an indisputable sign of polysemy (...) If the given combination is unique, that is if it occurs only once in the whole vocabulary of the language, then we are dealing not with a polysemantic word but with homonyms. (...) We can speak of different types of polysemy from the point of view of their relative remoteness to homonymy. Closest to homonymy are some types of metaphorically motivated polysemy (...) Metonymically and functionally motivated polysemy is, generally speaking, removed a step further from homonymy. [Apresjan (1974: 12)]

A partir d'aquesta diferenciació entre homonímia i polisèmia, Apresjan distingeix, encara, entre la polisèmia regular i la irregular:

Polysemy of a word A with the meanings a_i and a_j is called regular if, in the given language, there exist at least one other word B with the meanings b_i b_j , which are semantically distinguished from each other in exactly the same way as a_i and a_j and if a_i and b_i , a_j and b_j are nonsynonymous. (...) Polysemy is called irregular if the semantic distinction between a_i and a_j is not exemplified in any other word of the given language. [Apresjan (1974: 16)]

Sobre la base d'aquest tipus de distincions establertes per Weinreich i Apresjan, Pustejovsky formula la *complementary polisemy* i la *logical polysemy*:

There are two types of sense complementarity: (a) category preserving and (b) category changing. I will define *logical polysemy* as a complementary ambiguity where there is no change in lexical category, and the multiple senses of the word have overlapping, dependent, or shared. Hence, complementary polisemy is a slightly broader term than logical polisemy, since the former also describes how cross-categorical senses are related, for example with the use of *hammer* as both a noun and a verb. [Pustejovsky (1995a: 28)]

Segons Pustejovsky, allò que distingeix els sentits en un cas de polisèmia lògica del cas de l'ambigüitat contrastiva és la manera com es relacionen els sentits. La diferència fonamental és que la desambiguació pragmàtica o discursiva, tan important per caracteritzar l'ambigüitat contrastiva, sembla irrellevant en el cas de la polisèmia lògica, que es resol pel context lingüístic estricte. Un altre element de diferenciació és que mentre que l'ambigüitat contrastiva representa dos sentits mútuament excloents, la polisèmia sembla tenir una tolerància més alta a la possibilitat de coexistència:

Both senses of the noun seem relevant for the interpretation of the noun in the context, but one sense seems 'focused' for purposes of a particular context. [Pustejovsky (1995a: 32)]

Pustejovsky (1995a: 38) fa notar que, tot i l'existència de diferències entre l'ambigüitat contrastiva i la polisèmia, els sistemes de descripció del lèxic solen neutralitzar aquestes diferències pel que fa a l'organització general del lèxic a què donen lloc. Aquestes organitzacions no són essencialment diferents pel fet que totes dues són enumeratives i, per tant, es poden definir de manera molt similar (veg. més endavant § 3.5). Pustejovsky, però, no es refereix explícitament a la descripció lexicogràfica en fer aquestes observacions sobre els lèxics enumeratius. El caràcter enumeratiu d'un lèxic, per a Pustejovsky, resideix en la manca de mecanismes generatius que estableixin relacions entre significats (mecanismes que no es troben en els diccionaris convencionals). En lexicografia l'enumeració de sentits es produeix de manera natural a l'interior de l'article; però en alguns casos, com he mostrat, es pot projectar fins a la diferenciació de les entrades de diccionari.

L'aproximació de Pustejovsky a la qüestió de la polisèmia es refereix amb caràcter exclusiu a un cert tipus de polisèmia que es pot determinar sobre bases estrictament sincròniques, i en general té poc a veure amb les discussions que pretenen l'establiment de distincions entre homonímia i polisèmia que satisfacin alhora els principis sincrònics i els diacrònics. El plantejament de Pustejovsky se centra en l'anàlisi de la polisèmia sistemàtica, que Copestake & Briscoe (1996) subclassifiquen en *polisèmia construccional* (en què un mot amb un sol sentit bàsic té especialitzacions determinades

contextualment) i l'*extensió de significat* (que relaciona de manera predictable dos o més significats).

Buitelaar (1998) desenvolupa la idea de predictibilitat semàntica de la polisèmia i reserva aquest terme per als casos en què es poden preveure els canvis de significat. D'aquesta manera, postula la hipòtesi dels lèxics subespecificats que incorporen elements generatius que permeten la projecció semàntica dels lexemes. Així, representa les classes polisèmiques sistemàtiques a través de definicions de tipus semàntics subespecificats que prenen la forma d'estructures de *qualia* (veg. més endavant § 3.5).

La lexicografia convencional, per la seva banda, no ha reconegut en general l'existència de cap conflicte entre aquests tipus de polisèmia i l'homonímia. Les extensions de significat s'han resolt habitualment com a accepcions diferenciades d'un mateix article (sovint amb l'ús de marques valoratives que fan referència a aquest tipus de fenòmens (*figuradament, per extensió, etc.*). Els fenòmens de tipus metonímic que solen associar-se a la polisèmia construccional no han estat objecte, en general, de cap tractament descriptiu (Müller, 1992) que expliciti les relacions de predictibilitat que es poden establir entre els sentits.

Més que no pas dos models metodològics aplicables en sentit estricte, les denominades hipòtesis homonímica i polisèmica obeeixen a dues tendències màximament diferenciades de solució del mateix problema: l'ambigüïtat. Però rara vegada les realitzacions lexicogràfiques apliquen un d'aquests dos models en estat pur. Això és especialment cert en el cas de la hipòtesi homonímica, i ho seria semblantment per a la polisèmica si no fos per l'existència de criteris modificadors que, combinats de manera particular en la realització d'obres lexicogràfiques, poden conduir a organitzacions de la informació prou diferents. El LDOCE i el COBUILD constitueixen un exemple pràctic d'això; encara que tots dos diccionaris parteixen de la base que una entrada es pot realitzar en més d'un sentit, organitzen diferentment el tractament de certs casos particulars, com ara el del substantiu i verb *foil*.

El LDOCE dona tres entrades coincidents en la seva forma gràfica:

foil¹ (...) v [T (in)] to prevent (someone) from succeeding in (some plan); The thief was foiled in his attempt to enter the house. | We foiled his attempt to escape. [LDOCE]

foil² n 1 [U] (often in comb.) metal beaten or rolled into very thin paperlike sheets: Milk bottle tops are made of tin foils. | Wrap the chicken in foil before you cook it.

2 [U] paper covered with this: Cigarettes are wrapped in foil to keep them fresh. 3 [C (for, to)] a person or thing of a kind that makes the better or different quality of another more noticeable: In the play, a wicked old uncle acts as a foil to the noble young prince. [LDOCE]

foil³ n a light narrow sword used in FENCING -compare ÉPÉE, SABRE (2). [LDOCE]

Per altra banda, el COBUILD en dona només una:

<p>foil (...), foils, foiling, foiled. 1 Foil is metal in the form of a sheet as this as paper. It is used especially to wrap food in and keep it fresh. EG <i>Wrap cakes in foil before storing them... aluminium foil... the foil wrapper of a bar of chocolate.</i></p> <p>2 If you foil someone's plan or attempt at something you prevent them from being succesful. EG Their attempt to recapture Calais was foiled by a traitor... We had planned to leave at four, but he was foiled by Mrs Harlowe, who rang for tea.</p> <p>3 Something that is a foil for something else makes its good qualities more noticeable when the two things are experienced together, because of the great difference between them. EG <i>She had bronzed skin, for which her yellow swimsuit was a perfect foil... Cranberries are fairly sharp, but a good foil to fat meat.</i></p> <p>4 A foil is also a thin light sword used in fencing, which has a button on its tip to prevent injury. [COBUILD]</p>	<p>N UNCOUNT</p> <p>V + O</p> <p>= THWART N COUNT: IF + PREP THEN <i>for/to</i></p> <p>= CONTRAST, COMPLEMENT N COUNT</p>
--	---

En el LDOCE s'apliquen alguns criteris modificadors del model polisèmic que no són presents en el COBUILD, com el de la categoria lèxica com a factor de diferenciació d'entrades o l'aplicació de criteris diacrònics que determinen la diferenciació de *foil*² i *foil*³. El COBUILD, en canvi, reuneix tota la informació en un sol article, i no considera la categoria gramatical com a factor de diferenciació.

Un altre criteri diferenciador de l'homonímia-polisèmia consisteix a recollir com a entrades diferents mots que, malgrat la seva coincidència formal, presenten constituents

morfològics que poden ser considerats diferents i, per tant, tenen significats diferenciats i especialitzats. Aquest criteri és aplicat en el DIEC:

acceptador¹ -a *adj.* i *m.* i *f.* Que accepta. *L'acceptador d'una lletra.* | **acceptador de persones** Que usa parcialitat envers algú; que el tracta amb favoritisme o accepció. | **m.** i **f.** Acceptant.
acceptador² -a *adj.* Que pot ésser acceptat, acceptable.

Aquest criteri de diferenciació pot defensar-se sobre bases exclusivament sincròniques, ja que les característiques de cadascun dels dos processos derivatius són diferents i tenen una determinada productivitat lèxica. La consideració d'aquest fenomen com a diferenciador d'entrades, però, obvia el fet que la base lèxica de tots dos lexemes és la mateixa. A més, la incorporació al diccionari de dues entrades diferenciades no aporta claredat descriptiva al fenomen. En qualsevol cas, solucions d'aquest tipus traslladen el problema de la diferenciació entre homonímia i polisèmia als constituents del mot.

Tal com es pot veure, el diccionaris resolen i representen aquestes ambigüitats de manera diversa. La justificació de les diferències de tractament només pot trobar-se en la introducció de criteris d'ordre diacrònic que actuen com a factors de distorsió. L'única manera d'evitar aquestes distorsions és l'aplicació rigorosa dels principis de diferenciació d'entrades (veg. més enrere § 2.2) a partir dels tres elements de la tríada constituïda per la forma canònica, la categoria gramatical i el paradigma de flexió.

3 Els models del lèxic i el diccionari

3.1 Creativitat lèxica i diccionaris

La lingüística teòrica, en abordar el problema de la representació del lèxic, ha recorregut amb freqüència al model del diccionari. Els conceptes que la lingüística i la lexicografia tenen, respectivament, de diccionari s'han volgut aproximar i contraposar alhora. Així, per una banda, els diccionaris s'han presentat sovint com reculls que es poden identificar amb el cabal lèxic d'una llengua, o d'un parlant de la llengua; per aquest motiu s'han considerat presentacions més o menys sistemàtiques del lèxic. Per una altra banda, el principal factor de diferenciació entre els diccionaris i el component lèxic d'una llengua s'ha atribuït al caràcter dinàmic —en el sentit que ha d'afrontar la qüestió de la creativitat lèxica— del lèxic teòric en contraposició al caràcter estàtic dels diccionaris.

Per a la lingüística estructural el lèxic és un conjunt de límits indefinits, una *llista oberta*, en termes de Martinet (1960), dotada de mecanismes morfològics per a la formació de noves unitats. Els corrents generativistes, encara que van partir de models (Chomsky, 1957) en què el lèxic era tractat com un component secundari, han passat a concedir-li una importància creixent fins a assolir una posició principal (Scalise, 1984) en la teoria, sobretot amb la consolidació de l'anomenada hipòtesi lexicalista (Cabré, 1994a). Aquesta importància creixent ha coincidit amb l'èxit de les representacions del lèxic que, alhora que donen resposta a les qüestions relacionades amb el lèxic potencial i els mecanismes de creativitat, tenen el mot (Aronoff, 1976) com a unitat principal d'estructuració.

Malgrat que els diccionaris convencionals no solen estar dotats dels mecanismes de formació basats en la morfologia, el component lèxic utilitzat pel parlant *ideal* (idea recurrent en el generativisme) es denomina, per qüestions més pròximes a la intel·ligibilitat que a la semblança real, *diccionari mental*, *diccionari teòric* o *lexicó*.²⁶

La utilització de l'expressió *diccionari* per a referir-se al lèxic s'ha vist afavorida pel fet que, des d'algunes perspectives de base psicològica (Aitchison, 1987), es tendeix a afirmar que, malgrat la capacitat generadora del component lèxic d'una gramàtica, el nombre d'unitats emmagatzemades en el cervell humà està determinat per uns límits finits.²⁷ Cal, però, remarcar el fet que un diccionari mental no és equiparable a un inventari estàtic; tal com el mateix Aitchison afirma:

Turning to the mental lexicon, its content is by no means fixed. People add new words all the time, as well as altering the pronunciation and meaning of existing ones. Humans, however, do not just add on words from time to time, in between utterances. They often create new words and new meanings for words from moment to moment, while speech is in progress. [Aitchison (1987: 11)]

Un diccionari convencional, una obra lexicogràfica, és una llista necessàriament tancada i, en general, no dotada de cap component intern capaç de descriure adequadament allò que anomenem *creativitat lèxica*. Aquesta creativitat, a més, no es refereix només als mecanismes de creació de noves unitats a partir dels processos de la morfologia, sinó

²⁶ Encara que aquests tres termes solen usar-se com a sinònims, de fet, designen objectes diferents. La denominació *diccionari mental* fa referència al lèxic d'un parlant considerat com a model de la competència lingüística, mentre que *diccionari teòric* o *lexicó* són termes que fan referència a l'estructura del component lèxic de la gramàtica. Scalise (1984: 15 n.) fa notar l'ambigüitat amb què s'usa el terme *lexicó*, que en algunes ocasions es refereix a la totalitat del component lèxic de la gramàtica i en d'altres al diccionari, en tant que llista de mots, d'aquest component.

²⁷ Tal com assenyala Aitchison (1987); "The number of words known by an educated adult, then, is unlikely to be less than 50,000 and may be as high as 250,000. These high figures suggest that the mental lexicon is arranged on systematic basis." [Aitchison (1987: 7)]

que també implica la creació i la comprensió de nous significats, la capacitat d'introduir variacions en l'expressió dels mots, i la capacitat de rebutjar mots aparentment ben formats, però que no formen part del repertori lèxic d'una llengua. Si entenem, per tant, aquesta creativitat en un sentit ampli, podem concloure que un diccionari per a ús humà (o, fins i tot, un lèxic electrònic per a usos computacionals) té una capacitat necessàriament limitada de descriure tots aquests processos dinàmics per a tots els lexemes.

Aquesta diferència no pot ocultar, però, el fet que un diccionari s'adreça de manera primordial a la descripció del lèxic de la llengua. Malgrat el caràcter obert i dinàmic del lèxic com a conjunt, si el diccionari està basat en una mostra representativa de la llengua, pot oferir un inventari *representatiu* o, millor dit, *empíricament motivat* del lèxic de la llengua que descriu. Així, doncs, no és, en tant que estructura, directament equiparable ni amb la representació d'un *diccionari mental* o *teòric*, ni amb els processos que regulen la formació de paraules en una llengua; però sí que pot complir una funció complementària que no es produeix en el pla teòric sinó en el pla pràctic o, més precisament, en la dimensió *empírica* de la descripció lingüística.

Aquest darrer fet, que implica una clara especialització pel que fa als objectius del diccionari, s'acostuma a obviar en les discussions sobre les similituds i diferències que presenten ambdues estructures. El centre d'interès ha estat, més aviat, la comparació (amb una especial insistència a remarcar-ne les diferències) entre els aspectes parcials que presenten un i altre; en aquesta línia, Cabré & Rigau (1986) s'adonen d'aquest fet, però sense aprofundir-hi (el subratllat és meu):

Normalment un diccionari teòric vol ser més reduït que un de pràctic o imprès, però això no vol dir que no pugui contenir unitats que el diccionari pràctic no inclou. Aquest diccionari individual de cada parlant presenta característiques comunes i diferents de les del diccionari pràctic de llengua. *Les diferències són conseqüència lògica del seu abast i objectius.* Les identitats responen al seu caràcter de diccionari. [Cabré & Rigau (1986: 12)]

La possibilitats de comparació entre els diccionaris i el component lèxic es poden reorientar cap a terrenys més productius si prenem en consideració aquest fet, i possiblement s'eliminen els *quid pro quo* que, en certa manera, han originat la discussió i n'han fet una mena de lloc comú. Si partim de la base que un diccionari, un producte lexicogràfic, no té per missió la representació de tots els mots (i, per extensió, dels significats) *possibles* d'acord amb els mecanismes morfològics d'una llengua, sinó l'inventari dels mots *realment utilitzats* en els actes de comunicació d'aquesta llengua, la base fonamental de la discussió s'esvaeix o, en tot cas, se situa en uns límits més precisos que li confereixen sentit i interès metodològic.

La modelització del lèxic com un conjunt indefinit o, fins i tot, infinit (Corbin, 1983, 1987), no buida de sentit la tasca lexicogràfica, sinó que en fa una eina de descripció complementària, en el sentit en què ho expressa Rafel (1998):

Pour une bonne connaissance et une description fidèle de la langue, il est très important de pouvoir arriver à établir d'une façon détaillée l'ensemble complexe des règles qui sont à la base de sa capacité créative, mais il est aussi très important de connaître quelles sont, parmi les infinies combinaisons possibles, celles qu'une communauté linguistique actualise habituellement dans le discours. [Rafel (1998: 739)]

Aquesta opinió contrasta amb la de Corbin (1983: 47), que parteix de la base que la lexicografia té una doble contradicció a resoldre en el cas que vulgui enfrontar-se a l'explicitació dels mecanismes de creativitat lèxica: la contradicció ve donada pel fet que, d'una banda, ha de registrar en una llista finita un nombre potencialment infinit de mots construïts i, de l'altra, ha d'afrontar la qüestió que els mecanismes de descripció regular del significat es contraposen, de fet, a un nombre altíssim de comportaments idiosincràtics testimoniats en la llengua a través dels seus corpus de referència.²⁸

²⁸ Corbin (1983) parteix de l'anàlisi de diccionaris (concretament de les diverses edicions del DPR) per justificar la seva crítica a la lexicografia pel que fa a la representació dels mecanismes de creativitat lèxica: "La conclusion du linguiste sur ce point sera sévère pour le lexicographe: absence de théorie

L'objecció plantejada per D. Corbin als diccionaris es pot resumir en la pregunta següent: un diccionari ha de ser un model *competencial* o *referencial* del lèxic? Considerar la primera possibilitat obliga a dotar el diccionari d'uns components (regles de derivació i composició, llistes de radicals, idiosincràsies, etc.) que resulten poc habituals en la lexicografia convencional; a més, implica una descripció exhaustiva del lèxic que en aquests moments la morfologia potser no està en condicions de facilitar.

El model competencial, doncs, s'acosta més a la idea del diccionari teòric o del lexicó que no pas a la praxi lexicogràfica.

El model referencial, en canvi, constitueix l'evolució de la praxi lexicogràfica en la línia del seu perfeccionament metodològic. Les possibilitats actuals de tractament de la informació dels corpus textuais existents obren noves expectatives per a la descripció empírica del lèxic d'una llengua, tant en el seu aspecte macroestructural com en el microestructural.

El fet que el model referencial reculli millor les possibilitats de descripció de la lexicografia no implica l'assumpció que els diccionaris convencionals siguin coherents amb el principi de recollir el *lèxic real* i no el *lèxic virtual*. És ben cert (tal com assenyala Rafel, 1998) que els diccionaris, principalment els no basats en corpus, han tendit, tradicionalment, a incorporar lexemes que, tot i ser possibles des del punt de vista morfològic, no es troben avalats per cap evidència textual. Per altra banda, els diccionaris sovint han deixat de registrar mots que es poden considerar perfectament usuals. Si entenem el diccionari com a model referencial d'una llengua, aquesta qüestió s'ha de veure com una desviació metodològica. La majoria dels diccionaris no són prou exhaustius ni prou fiables perquè no recorren a les fonts d'informació adequades.

morphologique d'ensemble, contradictions internes, incohérences, redondances et inexactitudes diverses, voilà ce qui ressort de l'étude de la conception des régularités lexicales par le lexicographe. On ne peut évidemment pas en rester là." [Corbin (1983: 55)]

Dins de l'òptica de diccionari com a model bàsicament referencial hi ha, però, possibilitats de recollir certs mecanismes de creació lèxica, si més no els mecanismes de creació lèxica més regulars i que donen lloc a mots construïts d'una manera més sistemàtica.

La més habitual d'aquestes possibilitats és la que es coneix com *macroestructura complexa* (o, més impròpiament, *dobla macroestructura*), consistent a registrar certs derivats productius com a *entrades secundàries* o *subentrades*, sense desenvolupar un article complet corresponent a aquestes unitats, sinó integrant-les al diccionari com a part d'articles més complexos. Com que aquest procediment es basa en certes regularitats lèxiques que impliquen la regularitat semàntica dels mots construïts, permet l'estalvi d'un nombre força elevat de definicions que poden calcular-se com a funció directa del procés de derivació. Aquesta disminució del nombre de definicions *formulàries* (veg. Ilson, 1987) permet concentrar la descripció lexicogràfica en els processos de naturalesa idiosincràtica.

Els diccionaris convencionals que tracten la derivació solen incorporar a la macroestructura secundària els processos de tipus derivatiu que resulten més regulars (en el sentit que són reductibles a regles lèxiques estables) i productius (en el sentit que afecten un gran nombre de casos). Els processos descrits a partir d'aquest sistema no són, però, sempre els mateixos en els diferents diccionaris. La comparació entre un sistema i l'altre pot ajudar a fer comprendre fins a quin punt l'aplicació d'aquestes solucions representa, a part d'un acostament als plantejaments de la lingüística teòrica, una considerable evitació de la redundància entre els diferents articles del diccionari. Comparem, en aquest sentit, l'article *demagogue* del LDOCE i els seus corresponents del COBUILD:

demagogue (...) *n* [C] Someone who gives political speeches that try to persuade people by using emotional language rather than reason. – **demagogy, demagogery** *n* [U] – **demagogic** (...) *adj* – **demagogically** (...) *adv*. [LDOCE]
demagogic (...) If you say that someone is **demagogic**, you mean that they are typical of or like a demagogue; used showing disapproval. [COBUILD]

demagogue (...) A **demagogue** is a political leader who tries to win support by appealing to people's emotion rather than by rational arguments; used showing disapproval. [COBUILD]

demagogy (...) is a method of political rule which involves appealing to people's emotions rather than using rational arguments. [COBUILD]

A part de la diferència en el nombre de mots registrats (entrades i subentrades) d'un i altre diccionari —el LDOCE en recull 5 i el COBUILD només 3—, es pot veure en l'exemple adduït com el COBUILD ha de reprendre informació semàntica d'un article a l'altre pel fet que cada entrada dona lloc a un article estructuralment idèntic a la resta dels articles del diccionari.

Els diccionaris amb tractament dels fenòmens derivatius solen ser bastant restrictius pel que fa als processos considerats regulars. En certa mesura, es pot entendre que aquests tipus de tractaments no estan encara prou integrats en la tradició lexicogràfica i, a més, presenten alguns inconvenients d'ordre pràctic als quals els diccionaris són sensibles.

En molts casos, els processos derivatius situen el derivat en una posició alfabètica allunyada de la de la base lèxica (pensem, per exemple, en els casos de prefixació, com *feixista/antifeixista*), o bé introdueixen alternants al·lomorfemàtics en la forma del radical lèxic (*construir/construcció*). Aquests fets poden ser fàcilment resolts amb la inclusió de dos tipus d'entrada (una de principal i una de secundària per als derivats, amb una remissió a l'article que correspon a la seva base lèxica) ordenades alfabèticament en una mateixa llista. Aquesta solució, però, no sol ser habitual, ja que implica, a efectes exclusivament pràctics, una disminució considerable de l'espai que queda alliberat per efecte del tractament sistemàtic de la derivació.

El DUE tracta la derivació a partir d'un sistema similar al que acabo d'exposar. L'article corresponent a la base lèxica inclou una llista de la "família de paraules" que són tant els derivats predictibles com els no predictibles; en el cas dels primers, un article a la posició alfabètica corresponent indica que es tracta de mots de significat deduïble a partir del de la base lèxica, i indica la remissió a l'article que correspon:

demagogia (...) En lenguaje corriente, acción de gobernar con la preocupación dominante por agradar a la plebe.

Familia de palabras

demagógicamente

demagógico, -a

demagogo, -a

demagogo, -a (...) Político que lisonjea con sus discursos o actuación a la masa, para atraérsela.

De la familia de **demagogia**

demagógicamente; demagógico, -a (...) Derivados de significado deducible del de “demagogia”

De la familia de **demagogia**

L'estat actual de desenvolupament de la tècnica de descripció lexicogràfica permet suposar que el tractament integrat de la derivació en els diccionaris experimentarà en un futur pròxim un grau de generalització molt més alt del que té actualment, tant pel que fa a la sistematicitat de la seva forma lexicogràfica, com pel que fa a la quantitat i naturalesa dels processos derivatius susceptibles de ser tractats regularment. Aquesta suposició es basa en l'aparició i la generalització de suports de publicació electrònica (CD-Rom, diccionaris on-line, etc.) molt més adequats al tractament integrat de la derivació, ja que superen els problemes d'ordre pràctic que representa l'edició impresa d'un diccionari.

3.2 El diccionari i els lèxics electrònics

L'evolució experimentada per la lexicografia en els darrers anys (veg. Quemada, 1990, i també més enrere § 1.1) no ha estat un fenomen aïllat, sinó que ha coincidit amb el progrés experimentat per les disciplines a què recorre el lexicògraf en el seu treball. Una prova d'aquest canvi, que ha suposat un augment del rigor metodològic general dels diccionaris existents, és el fet que en els darrers decennis diverses iniciatives de recerca sobre el lèxic (fonamentalment basades en la lingüística computacional) han trobat en el processament de la informació dels diccionaris un cabal d'informació lèxica difícilment

substituïble. En aquest sentit, el fet que les *Guidelines* de “Text Encoding Initiative” (TEI) incloguin els diccionaris entre els tipus de textos per als quals es proposen estàndards de codificació en SGML²⁹ (i, posteriorment, en XML) és un símptoma més del fet que els diccionaris són cada cop més rigorosos pel que fa a seva estructura, a les propietats dels seus elements estructurals i a les relacions que mantenen entre si.

Els canvis experimentats pels diccionaris són de naturalesa diversa, i cal cercar-ne l’origen tant en la lingüística teòrica com en l’aplicada, i fins i tot en àrees multidisciplinàries en què el component lèxic esdevé un element important, com ara l’enginyeria lingüística o determinades àrees de la intel·ligència artificial.

La importància creixent adquirida pel lèxic en la lingüística, que ha passat d’estar centrada en els aspectes pròpiament gramaticals (i especialment sintàctics) a concentrar bona part dels seus esforços en l’estudi de les propietats dels mots, ha estat un factor determinant per a la millora dels diversos aspectes que caracteritzen la descripció lexicogràfica. Una mostra d’aquest canvi en el panorama de la lingüística teòrica el constitueixen el gran nombre de models d’aproximació lexicalista que han proporcionat, en molts casos, formalismes per a la implementació de lèxics computacionals.

Alguns d’aquests models del lèxic de base teòrica han estat aplicats en la constitució de repertoris lèxics de base computacional. Els sistemes de processament del llenguatge natural requereixen la concurrència de mòduls d’informació lèxica per a la construcció de sistemes d’anàlisi, comprensió, traducció o producció d’enunciats lingüístics. Aquesta necessitat ha fet que l’anomenada *lexicografia computacional* (Boguraev & Briscoe, eds., 1989) hagi adquirit una importància remarcable i que es pugui disposar de lèxics en format electrònic per a ús computacional, en molts casos obtinguts a partir del

²⁹ SGML (Standard Generalized Mark-up Language) és un esquema de marcatge autodeclaratiu definit per la norma ISO 8879. Fou adoptat pel consorci TEI (Text Encoding Initiative) com a base per al desenvolupament dels seus estàndards de codificació textual. Posteriorment, aquests estàndards s’han adaptat a XML (eXtended Markup Language), basat també en SGML, però específicament orientat a la codificació d’estructures de dades.

reprocessament de diccionaris convencionals. La lexicografia computacional ha estat determinant en les implementacions dels diversos models teòrics del lèxic, i ha contribuït a consolidar-los o a validar-ne l'adequació descriptiva a partir de les aplicacions pràctiques i de la concreció extensiva d'aquests models.

Al costat, doncs, de les aportacions de la lingüística teòrica, que ha proporcionat en molts casos les bases per a la progressió de la tècnica lexicogràfica, és determinant també la contribució dels corrents lingüístics sorgits de la lingüística aplicada. Ooi (1998) descriu l'aportació de la lingüística a la tasca lexicogràfica com una triple incidència, canalitzada a través de la lingüística teòrica, la lingüística computacional i la lingüística de corpus, respectivament:

Besides the lexicographer then, the linguist's contribution to the dictionary-making process can be subdivided into, at least, three roles. Firstly, the theoretical linguist can extend the notion of his/her theoretical dictionary, the lexicon. Secondly, the computational linguist can extend the notion of a lexicon by storing it in the computational lexical structure known as the lexical data/knowledge base. Thirdly, the corpus linguist can extend the notion of the lexicon by basing its description on an appropriate analysis of textual instances. [Ooi (1998: 2)]

Potser la interpretació d'Ooi no té prou en compte la subsidiarietat de la lexicografia respecte d'altres disciplines de la lingüística. Tot i que (en part pel seu plantejament tan sintètic) aquesta interpretació pot ser vista com a excessivament simplista, Ooi assenyala tres línies d'influència que aglutinen la major part de les aportacions que s'han fet a la lexicografia. Aquestes aportacions han permès, en el pla estrictament pràctic, una evident millora metodològica³⁰ dels diferents tipus de diccionaris elaborats

³⁰ Cal fer notar que això és una tendència, no pas un principi incontrovertible i sense excepcions. Hi ha, en especial a la lexicografia catalana, nombrosos exemples en què aquesta millora és inexistent, o almenys resulta absolutament relativa, o centrada únicament en els aspectes macroestructurals, etc. L'existència d'aquestes excepcions no invalida l'observació general, que es veu reforçada pel fet que les llengües amb un més gran dinamisme en la descripció lingüística són les que presenten una lexicografia més avançada des del punt de vista metodològic.

en els darrers decennis, que mostren, respecte a la lexicografia anterior, una sèrie d'aspectes que els caracteritzen d'una manera més o menys particular segons el tipus d'obra de què es tracta. La contribució d'aquestes aportacions al rigor formal, la preocupació pel vessant codificador, la coherència estructural i les possibilitats de reutilització dels diccionaris han estat factors determinants del progrés de la lexicografia.

També es poden considerar relacionades amb la lexicografia algunes línies de recerca de la intel·ligència artificial, en especial les desenvolupades a partir de la introducció de les *xarxes semàntiques* (Quillian, 1968), concebudes com un mecanisme d'expressió de les restriccions semàntiques (Grishman, 1986: 111). L'evident relació entre les xarxes conceptuals i l'estructura semàntica del lèxic va fer que la construcció d'aquest tipus de models estigués basada ja des del principi en definicions de conceptes extrems de diccionaris (Potts, 1994: 96), i orientades en general a la modelització del component lèxic. El desenvolupament d'aquestes estructures ha estat afavorit per la teoria de grafs, que ha permès la representació d'estructures cada cop més complexes que poden usar-se per a mostrar adequadament les relacions lèxiques, semàntiques i conceptuals que es produeixen entre els mots.

Algunes millores de la tècnica i del mètode lexicogràfics s'han incorporat a partir dels nombrosos punts de contacte que hi ha entre la lexicografia computacional i la lexicografia convencional. La reutilització de diccionaris per a la constitució de recursos lingüístics orientats a finalitats molt diverses ha creat un efecte de flux i reflux entre tots dos tipus de lexicografia.

En termes generals, podem dir que un lèxic electrònic (o computacional) és un component que dona informació pertinent a un sistema de processament del llenguatge natural o a una aplicació de l'enginyeria lingüística que requereixi el concurs d'informació lèxica. Per a Evens (ed.) (1988), els diferents models de lèxics computacionals són classificables en els termes següents:

Applications fall into two fundamental categories: computer implementations of relational models constructed by anthropologists, linguists, and psychologists in order to investigate the implications of their theories or simply to store and manipulate data effectively; and lexicons and knowledge bases built by computer scientists, borrowing ideas from everywhere, for information retrieval, natural language interfaces, or other natural language processing projects. [Evens (ed.) (1988: 17)]

Tot i que aquesta classificació correspon a un estat de coses constatable, pot ser qualificada com més o menys aïllacionista i, en certa manera, contrària a la tendència integradora i multifuncional dels recursos lingüístics.

Podem entendre els lèxics electrònics més desenvolupats com bases de dades multifuncionals (Zampolli, 1991*a*, 1991*b*) que poden proporcionar informacions fonològiques, morfològiques, sintàctiques i semàntiques que caracteritzen els lexemes i expliciten les relacions que contreen entre si.

Des del punt de vista pràctic, l'elevat cost de constitució de recursos lingüístics d'aquesta naturalesa i la dificultat que planteja el fet de trobar un formalisme de representació unitari fa que encara siguin escassos els lèxics que satisfan les exigències d'aquesta multifuncionalitat pel que fa al nombre d'entrades i als diferents aspectes descriptius. En el panorama europeu, la majoria d'aquests recursos han estat finançats a partir de programes de recerca de la Comissió Europea. Dos casos de formalismes que permeten aquesta multifuncionalitat són MULTILEX (programa ESPRIT 5304, 1990-93) i GENELEX (programa EUREKA E!-524, 1990-94) (veg. Goñi, 1998); aquest darrer ha donat cobertura als lèxics de PAROLE (programa LE2-4017, 1996-98) i SIMPLE (LE4-8346, 1998-2000), que s'han constituït per a un total de 12 llengües³¹ europees. Altres lèxics electrònics prioritzen un tipus de descripció per sobre de la resta; és el cas del lèxic de WordNet, que està fonamentalment dedicat a l'explicitació de la

³¹ Aquestes llengües són l'alemany, l'anglès, el català, el danès, el finès, el francès, el grec, l'italià, el neerlandès, el noruec, el portuguès i el suec.

informació semàntica, i ha constituït el model a seguir per als lèxics electrònics que, en el marc europeu, ha desenvolupat el projecte EuroWordNet (LE2-4003, 1996-99) per a diverses llengües³² europees.³³

Arribats en aquest punt, però, convé avaluar els punts de contacte i de divergència que es poden establir entre els tipus de lèxic que representen, respectivament, els diccionaris convencionals i els lèxics electrònics. Les diferències fonamentals entre uns i altres vénen donades, sobretot, per la diferent finalitat amb què han estat constituïts, i poden resumir-se en el quadre de la figura 3.1. Aquest quadre recull, en forma de característiques extremes, els principals factors de diferenciació de tots dos tipus d'obres. Cal entendre aquestes característiques com a tendències que poden tenir diversos graus d'aplicació en la realitat.

Diccionaris	Lèxics electrònics
No orientats a la descripció de cap àrea de coneixement en particular	Finalitat específica orientada a l'aplicació
Molta informació implícita (exemples, pressupòsits, etc.)	Informació totalment formalitzada i explícita
En general no segueixen estrictament cap model teòric	Estructures lligades a un model teòric
Tolerància a la manca de coherència estructural	Coherència estructural
No solen integrar components generatius. Adreçats a la descripció d'un conjunt de lèxic determinat	Poden integrar components generatius: lèxic virtual
Informació sintàctica implícita (arguments, règims preposicionals)	Informació sintàctica totalment desenvolupada a partir de la implementació d'un model teòric
L'explicitació del significat se centra en la paràfrasi definidora	La paràfrasi no és l'element principal de la descripció semàntica
Clara distinció entre l'orientació onomasiològica i la semasiològica	Indistinció entre onomasiologia i semasiologia

Figura 3.1. Diferències entre els diccionaris convencionals i els lèxics electrònics.

³² Aquestes llengües són: l'anglès, l'espanyol, el francès, l'italià i el neerlandès. Un lèxic electrònic català amb les mateixes característiques que els precedents fou desenvolupat, amb fons de la Generalitat de Catalunya, en paral·lel a l'execució del projecte EuroWordNet.

³³ Cito aquestes dues concrecions (PAROLE-SIMPLE i EuroWordNet) perquè en totes s'han desenvolupat recursos electrònics referits al català.

La llista de diferències de la figura 3.1, tot i que no és exhaustiva, recull els principals aspectes de divergència entre tots dos sistemes de descripció del lèxic. Convé destacar el fet, però, que aquestes divergències no constitueixen contraposicions descriptives, sinó que estan originades, en primer terme, per la finalitat amb què els lèxics electrònics es constitueixen (és a dir, el desenvolupament d'aplicacions computacionals) i, en segon terme, per determinats hàbits tradicionals en la lexicografia convencional.

Ja he assenyalat en aquesta mateixa secció que un dels punts de contacte més directe que, en la praxi, presenten la lexicografia convencional i la computacional es troba en la freqüent reutilització automatitzada de diccionaris per a la constitució de lèxics electrònics. El processament de diccionaris sobre suport magnètic ha estat a la base d'alguns projectes com ACQUILEX (ESPRIT BRA-3030, 1989-92), que han tingut com a objectiu fonamental la constitució de bases de dades lèxiques com a punt de partida per a la constitució de Bases de Coneixement Lèxic (veg. Boguraev & Briscoe, eds., 1989). Les bases de dades lèxiques per a les diferents llengües s'han construït a partir de l'adquisició automatitzada d'informació semàntica continguda en diccionaris convencionals sobre suport magnètic. En el procés de constitució d'altres lèxics electrònics, com WordNet o SIMPLE, també han utilitzat diccionaris com a referència de la codificació de la informació semàntica.

Les possibilitats de reutilització de diccionaris convencionals per a la constitució de lèxics electrònics ha canviat força des de l'observació de Gross (1991):

la diferencia de forma y de finalidad de los diccionarios electrónicos y de los diccionarios comunes tiene una consecuencia clara: la cinta de fotocomposición de un diccionario usual no sirve prácticamente de ayuda alguna a la hora de elaborar un diccionario electrónico.
[Gross (1991: 151)]

L'existència, que el mateix Gross (1991) reconeixia, de propostes d'interès per a l'establiment de vincles entre els diccionaris usuals i els diccionaris electrònics ha consolidat, finalment, l'estratègia del processament de diccionaris convencionals com a sistema més eficient per a l'adquisició de lèxics electrònics que integren un volum de dades satisfactori.

A partir d'aquesta constatació, podem pensar de cara a un futur immediat no només en la reutilització, sinó en possibilitats d'interacció molt més consolidades, que d'una manera natural hauran de portar al disseny d'estratègies per a la constitució paral·lela de tots dos recursos (Soler, 2000).

Per a la lexicografia computacional la representació del lèxic va estretament unida als aspectes relacionals que presenten els mots en les diferents vessants de la seva descripció. Les anàlisis que la lingüística ha aportat sobre estructures basades en relacions semàntiques (relacions entre sentits) i lèxiques (relacions entre mots) (Lyons, 1977, 1995; Cruse, 1986) ha possibilitat l'establiment d'estructures organitzatives basades en taxonomies que s'han volgut projectar a l'estructura del lèxic (Soler, 1992*a*, 1992*b*) a través de l'anàlisi de l'estructura dels diccionaris. Aquest fet per si sol ja ha contribuït, a través dels treballs sobre lexicografia, a cridar l'atenció sobre la necessitat de constituir diccionaris més coherents des del punt de vista estructural. La lexicografia computacional ha aplicat marcs teòrics per a la implementació de lèxics electrònics que permeten de reflectir en un grau satisfactori aquestes estructures.

La complexitat d'aquestes relacions i el caràcter essencial que prenen en la caracterització del lèxic és el que porta els models de lèxics computacionals al disseny de formalismes i d'estratègies capaces de descriure adequadament el lèxic. En les seccions següents (§ 3.3, § 3.4 i § 3.5) presento alguns aspectes de tres models de representació que han influït notablement en desenvolupaments de la lexicografia computacional.

Treballs com el de Fontenelle (1997*a*) assenyalen la importància de la "Meaning Text Theory" (MTT) (Mel'čuk & Zholkovsky, 1988), i especialment de les funcions lèxiques, com a model de representació de la informació lexicogràfica i per a la constitució de lèxics electrònics.

En segon terme, WordNet (veg. Fellbaum, ed., 1998), tot i que no es basa estrictament en la implementació de cap teoria particular del lèxic, constitueix un lèxic electrònic de

dimensions considerables completament desenvolupat, i representa un dels recursos més importants basat en les relacions semàntiques.

Finalment, una de les aportacions que més impacte han tingut sobre els models computacionals del lèxic és el *lèxic generatiu* (Pustejovsky, 1995a), que es basa en el principi que la descripció semàntica no consisteix únicament en relacions de tipus taxonòmic o jeràrquic (Cruse, 1986), sinó en un formalisme que ha de reflectir les relacions *multidimensionals* que es produeixen entre els lexemes, i que no poden ser captades per estructures *monodimensionals* basades exclusivament en relacions semàntiques com la hiponímia i la hiperonímia.

3.3 Un model de diccionari teòric: les funcions lèxiques

La MTT³⁴ constitueix un exemple destacat de les teories lingüístiques orientades a la descripció del lèxic. La MTT parteix de la idea que a cada signe lingüístic li correspon un *sintàctic* (*syntactique*), de manera que:

Y és un significat de Z

Z és un significant de Y

W és un *sintàctic* del parell <Y; Z>

³⁴ “Meaning Text Theory”: aquesta teoria fou formulada a Moscou l’any 1965, per Alexander Zholkovsky i Igor Mel’čuk; més tard s’hi afegí Juri Apresjan. La teoria s’orienta bàsicament a una lingüística basada en la descripció del lèxic, i els seus esforços es dirigeixen principalment a la producció d’obres lexicogràfiques anomenades *Diccionaris explicatius i combinatoris* (DEC), en curs d’elaboració per a diferents llengües. Per a una descripció del marc teòric general, veg. Mel’čuk (1984, 1988a,1993); Mel’čuk & al. (1995), i també Mel’čuk & Zholkovsky (1988).

Aquest *sintàctic* (Mel'čuk, 1993) és el conjunt d'informacions que descriu completament la combinatòria del signe lingüístic amb altres signes. Aquestes informacions són de diversos tipus, i recullen especificacions com ara la categoria lèxica del signe, el règim (sobretot per als verbs, però també del nom i de l'adjectiu) i, amb un paper destacat, les *funcions lèxiques* (FL) d'un signe lingüístic, que fan referència a les “relations sémantico-syntaxiques dont l'expression lexicale pour un lexème donné est phraséologiquement contrainte par ce dernier” (Mel'čuk, 1993: 117).

Mel'čuk tracta tant les relacions sintagmàtiques com les paradigmàtiques a partir de les FL, que cobreixen els diferents aspectes de la descripció lingüística del mot a partir d'un formalisme unitari i que pot ser fàcilment tractat des del punt de vista computacional.

Un dels punts de contacte més importants entre la MTT i la lexicografia pràctica és que la primera concreta la seva descripció del lèxic en els anomenats *diccionaris explicatius i combinatoris* (DEC) (Mel'čuk, 1984, 1988a). Els DEC segueixen un sistema fortament formalitzat de descripció del significat; i un dels principis (veg. més endavant § 6.1) en què es basa la seva metallengua de descripció és el *principi d'univocitat*, que consisteix a no utilitzar com a llengua de definició cap terme ambigu.³⁵

Les FL³⁶ conformen la secció *lexicocombinatòria* del model de descripció de la MTT, i

³⁵ Segons Mel'čuk (1988a: 171-172), aquest principi porta associades dues conseqüències: “First, each term used in the definitions of an ECD must have one, and only one, sense. Polysemous expressions that have to appear in a definition must thus be previously disambiguated with the help of distinctive numbers specifying the intended sense and introduced elsewhere in the dictionary. (...) Second, each distinct semantic component must always be represented by the same linguistic expression.” Mel'čuk assenyala que el primer d'aquests criteris no s'aplica sistemàticament a cap diccionari i que aquest sistema planteja, naturalment, el problema que només es poden etiquetar els mots ja definits pel propi diccionari. El segon criteri es pot exemplificar amb la prohibició de la diversificació d'expressions que introdueixen el mateix tipus d'especificacions com ara *emetre / produire / fer (un crit, un soroll, etc.)*; aquesta mena de dispersió és molt habitual en els diccionaris d'ús corrent.

³⁶ Per a una presentació especialment aplicada de les FL, veg. Mel'čuk (1984). Per a veure el grau de pertinença que presenten les FL en el problema de l'aspecte sintagmàtic del lèxic, veg. Benson & al. (1986: 256 i ss.)

poden ser de diversos tipus que, de fet, engloben propietats tant sintagmàtiques com paradigmàtiques.³⁷ En aquest treball no intento de fer una presentació sistematitzada³⁸ del sistema de les FL ni de la MTT, sinó simplement remarcar alguns aspectes d'aquestes funcions que poden guiar la descripció lexicogràfica. L'elevat nombre de FL i els nombrosos treballs que s'hi han referit o que les han adoptat com a marc de treball fan que sovint la delimitació de les FL reals –o, en termes de Mel'čuk (1984), *estàndard*– no puguin ser fixades amb facilitat; una cosa semblant passa amb les possibilitats de combinació entre FL simples per crear-ne de complexes. Em limito, doncs, a exposar de forma sintètica les FL que poden tenir un interès més destacable per a la descripció lexicogràfica.

Les FL de tipus paradigmàtic formalitzen relacions semàntiques com la sinonímia, la complementarietat, la hiponímia, la meronímia, etc. Altres FL fan referència a qüestions relatives a la morfologia lèxica o reflecteixen relacions de tipus (morfo)semàntic existents entre el lexema i altres categories gramaticals.

La vocation des fonctions lexicales est de fournir au locuteur la totalité des moyens lexicaux nécessaires à l'expression la plus riche, la plus variée et la plus complète de la pensée et, en même temps, de garantir la choix le plus précis de la formulation appropriée. (...) les FL (...) alimentent un système puissant de PARAPHRASAGE, qui est à la fois une raison d'être des FL et un outil fondamental de leur vérification. [Mel'čuk & al. (1995: 129)]

³⁷ Fontenelle (1997a: 98) classifica en 8 grups les FL segons que es refereixin a: 1. Morfologia derivativa; 2. Casos profunds (*Deep cases*); 3. Funcions circumstancials; 4. Distincions aspectuals; 5. Visió del món; 6. Quantificadors; 7. Estilística; 8. Funcions semàntiques estrictes. Prefereixo mantenir l'única classificació reconeguda per Mel'čuk, que agrupa les FL en funció del seu caràcter sintagmàtic o paradigmàtic.

³⁸ Un dels problemes que planteja la presentació sistemàtica de les FL és l'establiment de la llista completa de funcions *estàndard* que la MTT reconeix. Mel'čuk & al. (1995) n'inventaria 56, però aquest nombre s'incrementa fins a 64 a Mel'čuk (1996). El problema s'agreuja si tenim en compte que aquestes dades es refereixen únicament a FL simples, ja que determinades FL tenen la possibilitat de combinar-se amb altres per formar FL complexes, amb un menor grau d'estandardització fàctica.

Entre les FL paradigmàtiques que expressen relacions lèxiques de caràcter més marcadament semàntic hi ha les funcions **Syn** (sinònim de); **Conv_{ijk}** (relació conversa del mot clau de la FL), o **Anti** (antònim de). Altres funcions que vehiculen relacions semàntiques són:

- **Gener**: mot genèric que, combinat amb un derivat sintàctic del mot clau, dóna un sinònim:

gener(gas) = substància [gasosa]

gener(monarquia) = estat [monàrquic]

- **Figur**: metàfora codificada per la llengua que dóna una combinació sinònima del mot clau.

figur(passió) = foc [de la passió]

Pel que fa a les FL que descriuen relacions de tipus derivatiu o que estableixen diferents tipus de relació morfològica entre el mot clau i el valor de la funció, hi ha les següents:

- **S₀, V₀, A₀, Adv₀**: modalitats substantiva, verbal, adjectival o adverbial amb el mateix significat que el del mot clau::

S₀(cantar) = cant

V₀(anàlisi) = analitzar

- **S₁, S₂, S₃, ...**: Nom típic per al primer, segon, tercer... actant del mot clau:

S₁(crim) = criminal

S₂(crim) = víctima [d'un crim]

- **S_{instr}, S_{loc}, S_{med}, S_{mod}, S_{res}**: nom típic per a l'expressió de l'instrument, el mitjà, el mode i el resultat:

S_{instr}(parlar) = llengua

S_{loc}(parlar) = locutori

S_{med}(rentar) = sabó, detergent

S_{mod}(escriure) = escriptura

S_{res}(rentar) = bugada

- **A₁, A₂, A₃, ...**: modificador adjectival típic per al primer, segon, tercer... actant del mot clau:

A₁(indiferència) = ple [d'indiferència]

A₂(indiferència) = tractat [amb indiferència]

- **Able₁, Able₂, ...**: tal que pot, tal que es pot (derivat semàntic adjectiu potencial):

Able₁(infectar) = infecció

Able₂(admiració) = admirable, digne [d'admiració]

- **Qual₁, Qual₂, ...**: tal que està predisposat a (fer) l'acció denotada pel mot clau, que està predisposat a ser fàcilment objecte de l'acció denotada pel mot clau (derivat semàntic adjectiu virtual, que implica amb molta probabilitat **Able_i**):

Qual₁(enganyar) = mentider

Qual₂(enganyar) = ingenu

- **Sing**: un quàntum, una porció regular de:

Sing(flota) = vaixell

- **Mult**: conjunt regular de:

Mult(vaixell) = flota

L'altre grup bàsic de FL s'ocupa de les relacions sintagmàtiques, i serveixen com a patró de classificació de les col·locacions més usals dels lexemes.

Entre les FL sintagmàtiques, n'hi ha un grup que, prenent com a mot clau un substantiu, formulen la relació existent entre aquest substantiu i una expressió verbal amb la qual sol concórrer. Tres d'aquestes FL representen verbs semànticament buits (això és, que perden les característiques semàntiques pròpies quan ocorren amb el substantiu en qüestió), i en cada una el conjunt format pel mot clau (C₀) i els actants de la situació (1,

2) es distribueixen d'una manera determinada en relació als tres papers sintàctics: Subjecte gramatical (SG), Complement d'objecte principal (COP) i Complement d'objecte secundari (COS). Les funcions són les següents:

- **Oper₁**, **Oper₂**, ...: verb semànticament buit que pren 1, 2, ... com a subjecte gramatical i C₀ com a COP:

Oper₁(consell) = donar

Oper₂(consell) = rebre

És a dir, en una frase com *En Pau dóna consell a en Pere*, en la representació de la qual *Pau* és l'actant 1 i *Pere* l'actant 2 d'una mateixa situació:

En Pau	dóna	consell
[actant 1 - SG]		[C ₀ - COP]
En Pere	rep	consell
[actant 2 - SG]		[C ₀ - COP]

- **Func₀**, **Func₁**, **Func₂**, ...: verb semànticament buit que pren C₀ com a SG i 1, 2 (si n'hi ha) com a COP.

Func₀(vent) = bufar

Func₁(ajuda) = venir, provenir [d'algú]

Func₂(llista) = contenir, comprendre [alguna cosa]

- **Labor_{ij}**: verb semànticament buit que pren *i* com a SG, *j* com a COP, i C₀ com a COS:

Labor₁₂(llista) = inscriure, posar [alguna cosa en una llista]

- **Real₁**, **Real₂**, ...: verb que té el sentit de 'realitzar', i pren C₀ com a COP i el nom de 1, 2,... com a SG:

Real₁(temptació) = caure en

Real₂(ordre) = executar, complir

- **Fact₀, Fact₁, Fact₂, ...**: verb que té el sentit de ‘realitzar-se’, i pren C₀ com a SG i 1, 2... (si hi són) com a COP:

Fact₀(desig) = acomplir-se

Fact₂(memòria) = guardar, retenir [alguna cosa a]

- **Labreal_{ij}**: verb que té el sentit de ‘realitzar’. Pren els dos actants de la situació com a SG i el seu COP respectivament, i C₀ com a COS:

Labreal₁₂(memòria) = conservar [alguna cosa a la memòria]

- **Involv**: verb que pren C₀ com a SG i el nom de l’objecte que pateix l’acció de la situació sense participar-hi, com a COP. La FL **Involv** recull els casos en què la situació no té actants pròpiament dits, però en què es produeix la presència predictable d’un complement d’objecte:

Involv(olor) = omplia [l’estança]

Altres FL de tipus sintagmàtic són de semàntica més específica:

- **Magn**: molt, en un grau elevat, intensament:

Magn(memòria) = prodigiosa, d’elefant

Magn(soroll) = infernal, ensordidor

- **Manif**: ‘manifestar-se en alguna cosa’. C₀ és el SG. Sovint apareix associada a la FL **Caus**

Manif(joia) = esclatar

Caus₁Manif(excusa) = presentar

- **Degrad**: ‘degradar-se, fer-se malbé’

Degrad(vi) = agrir

- **Excess**: ‘funcionar de manera excessiva’

Excess(cotxe) = embalar-se

- **Son**: ‘emetre el so típic’

Son(gos) = bordar

Finalment, altres FL de tipus sintagmàtic es refereixen a aspectes del significat verbal i tenen en general un caràcter complementari que fa que molt sovint es combinin amb d’altres FL:

- **Incep, Cont, Fin**: ‘començar, continuar, cessar’.³⁹

IncepOper₁(forma) = prendre

- **Caus, Liqu, Perm** (causar, liquidar, permetre).⁴⁰

LiquFunc₀(assemblea) = (dissoldre)

No cal insistir gaire en l’interès que tenen per al lexicografia —fins i tot per a la lexicografia convencional— aproximacions com la de la MTT. Els diccionaris solen recollir, de vegades de manera no gaire sistemàtica, mostres que donen idea del comportament sintagmàtic del mot, però la lexicografia convencional o comercial no ha explorat encara la possibilitat de dotar-se d’un sistema de classificació i descripció formals d’aquesta informació. Un sistema com el que acabem de veure podria ser un

³⁹ Les relacions semàntiques entre aquestes tres funcions es poden descriure de la següent manera:

$FinP = IncepnoP;$

$ContP = noFinP = noIncepnoP;$

Ha cessat de llegir: ‘Ha començat a no llegir’; *Continua llegint*: ‘No cessa de llegir’

⁴⁰ Anàlogament al cas anterior, les relacions semàntiques entre aquestes tres FL es poden descriure de la següent manera:

$LiquP = CausnoP;$

$PermP = noLiquP = noCausnoP$

punt de partida per a l'expressió estructurada de moltes característiques de gran interès per a la descripció adequada del lèxic.

Hi ha alguns factors, però, que limiten les possibilitats d'una aplicació directa de les FL a la lexicografia convencional. El principal d'aquests factors és la complexitat que les FL suposen per a un usuari de diccionaris mitjà, tant per a la seva codificació com per a la seva interpretació.

En la seva concreció formal, els DEC es diferencien notablement dels diccionaris convencionals. Ja hem vist que la descripció de les relacions sintagmàtiques i paradigmàtiques es basa en les FL. El *règim* dels lexemes s'especifica en un quadre per a cada accepció on s'assignen el nombre d'actants de la situació i s'indica per quina preposició han de ser introduïts i/o quina categoria gramatical han de tenir. Les FL van associades a un sentit i a un règim, i sovint es presenten en combinació o amb especificacions en les quals ara no entraré. Potser serà suficient, per fer-se'n una idea, de transcriure aquí algunes FL associades al sentit **1b** del lexema nominal *aide*, corresponents al *DECFC (Dictionnaire explicatif et combinatoire du français contemporain)* per a la caracterització dels aspectes sintagmàtics que presenta aquest mot (veg. Mel'čuk & al., 1984: s. v. *aide*):

Magn ^{temp}	: constante
AntiMagn ^{temp}	: provisoire
Magn ^{quant}	: collective, publique
Ver [= A ₁ Fact ₀]	: efficace
Bon	: précieuse
Bon ₁	: désintéressée
AntiBon ₁	: intéressée
Oper ₁	: apporter [ART _ à N]; prêter [ART _ à N] plutôt A.(a); fournir, accorder [ART _ à N] plutôt A.(b) [<i>accorder une aide généreuse, mais *Paul est venu accorder son aide, on dit Paul est venu prêter son aide</i>]
IncepOper ₁	: venir [en _ à N]
de nouveau IncepOper ₁	: rétablir [ART _]
FinOper ₁	: retirer [ART _ à N]
Oper ₂	: recevoir [ART/ de l' _ de N]

ContFunc ₀	: se poursuivre
F ₁ = tenter de	
Caus ₂ Func ₀	: demander [ART/de l' _ à N]; implorer [l' _ de N]
dans une situation de danger ou d'urgence pour X, F ₁	: appeler [N à l' _] // appeler [N] à la rescousse
Func ₁	: venir, provenir [de N]
Func ₂	: parvenir [à N]
Func ₄	: consister [en N]
Real ₂	: bénéficiaire [de ART _]
Adv ₂ Real ₂ l'aide de Jean]	: avec [ART _] C ₁ ' L [Il a obtenu ce poste avec l'aide de Jean]
dans une situation de danger	
Imper	à l' _ !

Com es pot veure, el grau de detall i d'exhaustivitat descriptiva d'un DEC és notable. Aquest és probablement un dels motius fonamentals perquè aquestes obres no arribin, pel que fa al nombre de mots descrits, a passar de l'estadi de mostra de diccionari. A més, el caràcter altament formalitzat de la descripció sembla orientar més els DEC cap a una utilització computacional que no pas a un ús humà. Tot i així, cal tenir en compte el grau de sistematicitat amb què la MTT ha tractat el problema de la caracterització del lèxic. Això fa que sigui un dels models que han estat utilitzats per a la constitució de bases de dades lèxiques a partir de diccionaris existents, ja que (Fontenelle, 1998) les FL fan explícites, a partir d'un formalisme conceptualment simple i sistemàtic pel que fa al tractament dels casos, una gran part de les informacions que un diccionari sol recollir de manera implícita (Heid, 1996) en diversos elements estructurals de l'article.

3.4 Un model de lèxic basat en les relacions semàntiques: WordNet

Els plantejaments conceptuals basats en la idea de *relació semàntica*, i en les estructures de significat que projecten, han tingut un gran pes en la semàntica de les darreres dècades. A partir d'aquesta idea s'han desenvolupat lèxics de base relacional. El més important és WordNet, que està orientat de manera preferent a la representació relacional dels lexemes, de manera que els aspectes de descripció sintàctica queden supeditats a l'organització conceptual que és representable en termes de xarxa de relacions. WordNet és un lèxic desenvolupat a la Universitat de Princeton (Miller & al. 1990) d'ençà de 1985.

WordNet organitza la seva estructura a partir de la divisió del lèxic en categories gramaticals (noms, verbs, adjectius i adverbis). En alguns aspectes, WordNet s'assembla a un diccionari convencional (G. A. Miller 1998a: 9), en el sentit que dóna definicions lexicogràfiques i exemples per a les seves entrades. La seva estructura interna, però, s'aproxima més a la d'un *Thesaurus* que no pas a la d'un diccionari general. Aquesta aproximació entre semasiologia i onomasiologia és molt característica en els lèxics electrònics (Chen & Chang, 1999) i és un dels punts més destacables de les noves plataformes de publicació electrònica dels diccionaris convencionals, que permeten l'accés a la informació a través de diferents elements estructurals del diccionari.

La relació semàntica bàsica de WordNet és la sinonímia, i els grups de sinònims (*synset*) són la base d'organització del lèxic per a cada categoria. Cal dir, però, que tal com G. A. Miller (1998a) assenyala, la noció de sinonímia adoptada en WordNet no implica la substituïbilitat en tots els contextos. En aquest plantejament es deixen de banda certs aspectes sintagmàtics del significat per concentrar-se en els estrictament conceptuals. La unitat de descripció és el mot, però la unitat de recompte i organització

és el *synset*, de base conceptual. Una de les conseqüències directes d'aquesta organització d'orientació onomasiològica és que un mateix mot pot formar part de diferents *synsets*, en funció dels diferents significats que se li poden atribuir, com fa notar G. A. Miller (1998b):

dictionary entries for polysemous words (...) have several different glosses, whereas a synset has only a single gloss. Thus, a dictionary entry can contain semantic information that, in WordNet, would be distributed over several distinct synsets, one of each meaning. It is convenient to think of a synset as representing a lexicalised concept of English. That is to say, a lexicalised concept is represented in WordNet by the set of synonyms that can be used (in an appropriate context) to express that concept. [G. A. Miller (1998b: 24)]

Les relacions a WordNet no es produeixen, doncs, entre lexemes sinó entre *synsets*. La projecció d'aquestes relacions permet el coneixement més aprofundit de les estructures del lèxic en termes de xarxa semàntica. Cada *synset* és susceptible de contreure amb altres *synsets* diferents relacions semàntiques verificables en la categoria gramatical que els és pròpia.

Els *synsets* nominals en WordNet (G. A. Miller, 1998b) contenen informació sobre les relacions semàntiques que contreuen amb altres *synsets*. Així, es codifiquen les relacions d'antonímia, hiponímia, hiperonímia, meronímia i holonímia, la qual cosa permet la projecció de les diferents relacions semàntiques en el lèxic i l'establiment d'estructures de relació jeràrquica entre els diferents *synsets*. Amb la finalitat de donar coherència a aquestes organitzacions, i d'una manera especial a les relacions de tipus jeràrquic, com la hiperonímia i l'holonímia, WordNet té, per als noms, un conjunt de 25 *unique beginners* (que encara es poden reduir a 11) que constitueixen el nivell màxim d'organització de les estructures jeràrquiques. Així, qualsevol *synset* pertany (per relació directa o inferida a través de l'herència de propietats) a una de les classes que es mostren organitzadament a la figura 3.2, que reproduïxo de G. A. Miller (1998b: 30), en què apareixen els 25 *unique beginners*, i en cursiva els 11 a què es poden reduir. WordNet també codifica un cert nombre d'*atributs* de base lèxica que no estan predefinitos ni limitats en nombre; l'associació dels noms a aquests atributs té,

aproximadament, la forma d'una funció lèxica del tipus COLOR(canari)=groc, SIZE(canari)=petit. Com pot veure's aquest marcatge d'atributs posa en relació les propietats dels noms amb els adjectius que les expressen.

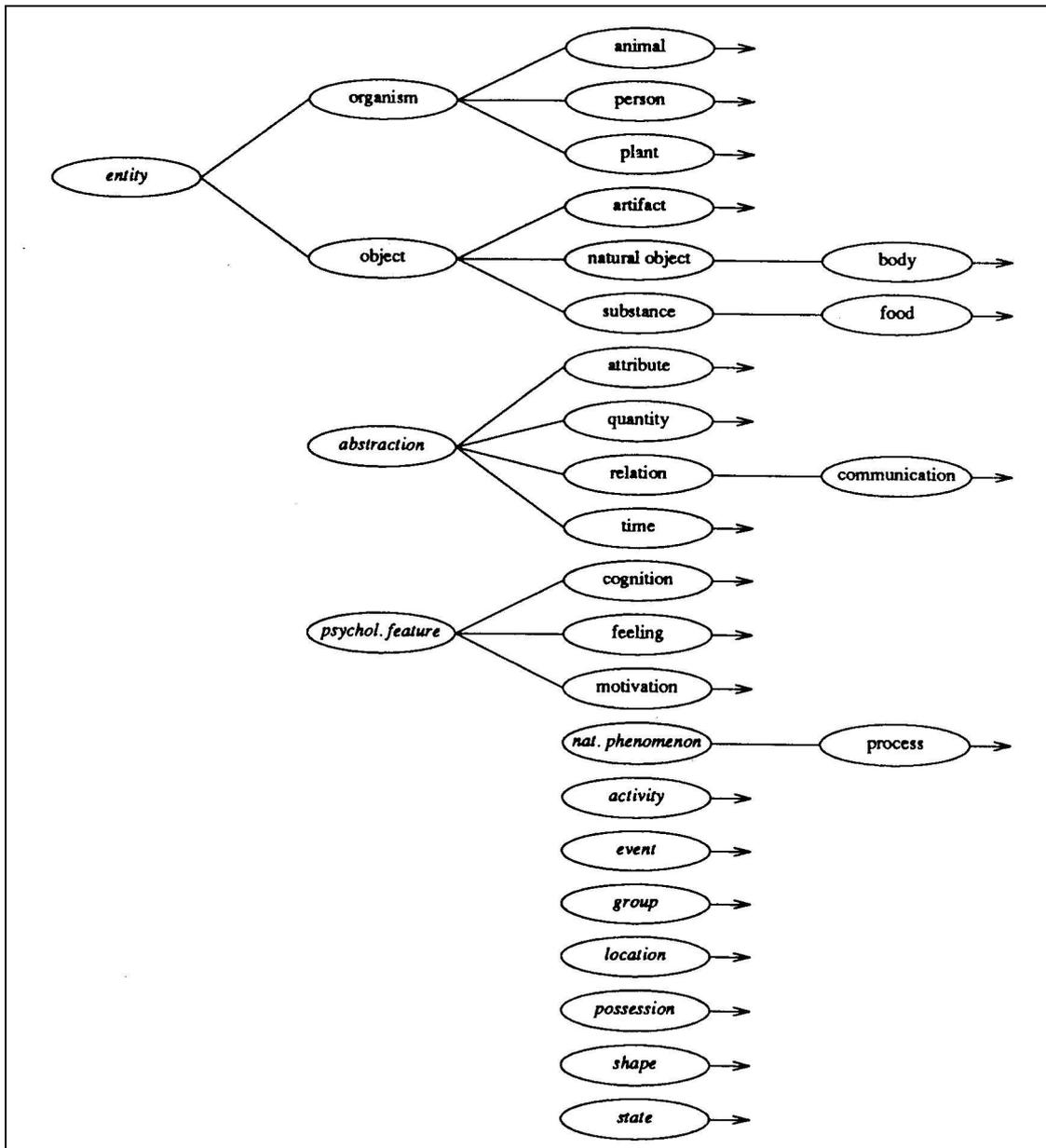


Figura 3.2. Classificació dels noms a *WordNet* extreta de G. A. Miller (1998b).

Els *synsets* verbals (Fellbaum, 1998) estan organitzats en classes semàntiques. En un primer nivell d'organització es distingeix entre els verbs que denoten accions i esdeveniments, d'una banda, i els que denoten estats, de l'altra. El primer d'aquests grups se subdivideix, al seu torn, en 14 subclasses més específiques: verbs de moviment, percepció, contacte, comunicació, competició, canvi, etc. Aquestes divisions actuen com a correlat dels *unique beginners* dels *synsets* nominals. El tipus de relacions que es codifiquen en els verbs són les d'implicació (*entailment*), del tipus *roncar* > *dormir*; les de *troponímia*⁴¹ i d'hiperonímia, del tipus, *mormolar* > *parlar*, *parlar* > *mormolar*, respectivament. L'existència d'aquestes relacions permet l'establiment de projeccions jeràrquiques en la classe dels verbs. També es codifiquen l'antonímia i la relació de causativitat, és a dir, els corresponents causals dels verbs que presenten alternança d'ergativitat (del tipus, *X bull l'arròs/l'arròs bull*; o bé *X tanca la porta/la porta es tanca*).

WordNet divideix la categoria dels *synsets* dels adjectius en dues classes bàsiques (K. J. Miller, 1998); els adjectius *descriptius* (*descriptive adjectives*) que corresponen a la classe més general dels adjectius (qualificatius, adjectius de color, quantificadors, etc.), i els *relacionals* (*relational adjectives*), que corresponen als adjectius que podem relacionar semànticament (i morfològicament o no) amb noms,⁴² com ara *dental*<*dent*, o bé *agrari*<*camp*. El primer grup s'organitza a través dels diferents *clusters*⁴³ de sentit que reflecteixen oposicions de tipus antonímic, del tipus *fast/slow*. Per als adjectius de tipus relacional s'indica quin és el nom amb què semànticament es relacionen.

⁴¹ Fellbaum (1998: 79) defineix la troponímia de la següent manera: “the troponymy relation between two verbs can be expressed by the formula: *To V₁ is to V₂ in some particular manner*”. Es tracta, doncs, de la relació corresponent a la hiponímia en el cas dels noms.

⁴² Es tracta del tipus d'adjectius que els diccionaris solen definir amb fórmules del tipus *relatiu o pertanyent a X*.

⁴³ Segons K. J. Miller: “adjectives are organised in clusters of synsets associated by semantic similarity to a focal adjective that relates the cluster to a contrasting cluster at the opposite pole of the attribute”. [K. J. Miller (1998: 50)]

Pel que fa als adverbis, l'organització en WordNet és simple i no es projecta estructuralment com en la resta de categories; WordNet no estableix de fet classes entre els adverbis i, pel fet que la gran majoria estan formats a partir de processos derivatius, només els codifica en referència a l'adjectiu de què provenen.

Cada tipus de lèxic electrònic està orientat a la representació del lèxic a partir d'un model que satisfà les necessitats de determinades aplicacions computacionals. WordNet està concebut com una representació de tipus mentalista i per aquest motiu orienta la descripció de totes les seves unitats sobre la base de les relacions semàntiques que contreuen. La informació continguda a WordNet s'ha obtingut a través del tractament manual de diccionaris existents, això explica les similituds que manté amb aquest tipus d'obres; de fet, les relacions lèxiques i semàntiques codificades poden verificar-se de manera implícita en diferents elements estructurals del diccionari i, de manera molt especial, en la definició lexicogràfica. Un dels valors afegits de WordNet és la seva extensió: la versió 1.5 conté 60.000 synsets nominals, 16.500 synsets adjectivals, i 11.500 synsets verbals.

En general els lèxics electrònics no presenten graus d'exhaustivitat tan importants, o bé no es refereixen al lèxic en general, sinó a dominis específics lligats al disseny d'aplicacions concretes. Aquesta particularitat ha fet que WordNet es convertís, de fet, en un estàndard *de facto* per a l'elaboració de lèxics electrònics com els concebuts en el projecte EuroWordNet (veg. Vossen, ed., 1998), que adopten els *synsets* establerts per a l'anglès com a elements de referència comuna per a l'elaboració d'una estructura multilingüe. En paral·lel a l'elaboració dels lèxics d'EuroWordNet, i seguint criteris de constitució idèntics, s'ha constituït el WordNet català, que conté 31.199 synsets nominals, 5.239 synsets adjectivals i 2.751 synsets verbals.⁴⁴

⁴⁴ La font d'informació que consulto és la pàgina web del Centre de Llenguatge i Computació de la Universitat de Barcelona: <http://clic.fil.ub.es/recursos/wordnet.shtml> (accés: 19 d'agost de 2003).

Hi ha dues característiques distintives de WordNet, però, que suposen una certa limitació de la capacitat descriptiva d'aquest lèxic pel que fa a la caracterització dels seus elements i del lèxic en general:

a) No conté relacions de tipus sintagmàtic entre mots de diferents categories gramaticals.

b) El tractament independent de les diferents categories gramaticals (noms, verbs, adjectius, adverbis) fa que les estructures projectades només integrin elements de la pròpia categoria i que, a més, no es reflecteixi cap interrelació entre elements d'una categoria i d'una altra.

Tot i aquestes limitacions, WordNet és un dels recursos que d'una manera més eficient han aconseguit tractar el lèxic a partir de les seves relacions. WordNet, que ha estat constituït a partir de la informació que contenen diferents diccionaris convencionals, és es especialment adequat per a ser utilitzat com a *feedback* en l'elaboració de diccionaris que tinguin la pretensió de presentar un alt nivell de coherència estructural en l'expressió de les seves relacions semàntiques.

3.5 Representació de sentits relacionats: el lèxic generatiu

Pustejovsky (1995a) constitueix una de les aportacions més influents a les actuals teories del lèxic, i ha estat un referent important en els treballs sobre semàntica i lexicografia computacionals. A partir de la utilització dels formalismes d'unificació i de les estructures subespecificades (veg. més endavant § 4.6), Pustejovsky introdueix la idea de *lèxic generatiu*, que es caracteritza per ser un component dinàmic pel que fa a les propietats dels lexemes. Per a Pustejovsky (1995a), el fet de no considerar les relacions de tipus generatiu comporta la modelització del lèxic com un conjunt d'elements que semànticament estan caracteritzats al mateix nivell. Això correspon a allò que ell anomena *Sense Enumeration Lexicon*, i que formula de la manera següent:

A lexicon L is a *Sense Enumeration Lexicon* if and only if for every word w in L , having multiple senses s_1, \dots, s_n associated with that word, then:

if s_1, \dots, s_n are contrastive senses, the lexical entries expressing these senses are stored as w_{s_1}, \dots, w_{s_n}

if s_1, \dots, s_n are complementary senses, the lexical entry expressing these senses are stored as $w_{\{s_1, \dots, s_n\}}$. [Pustejovsky (1995a: 38)]

Un lèxic generatiu es distingeix d'un lèxic enumeratiu pel fet que no especifica linealment els diferents sentits d'un mot, sinó que explica la polisèmia de què són objecte els mots a partir de mecanismes de generació de nous significats (Copestake & Briscoe, 1996).

Els principals elements caracteritzadors de les entrades, presos com a nivells de representació, del model de lèxic generatiu de Pustejovsky són els següents:

1. ARGUMENT STRUCTURE: specification of number and type of logical arguments, and how they are realized syntactically.
2. EVENT STRUCTURE: Definition of the event type of a lexical item and a phrase. Sorts include STATE, PROCESS and TRANSITION, and events may have subeventual structure.
3. QUALIA STRUCTURE: Modes of explanation, composed of FORMAL, CONSTITUTIVE, TELIC and AGENTIVE roles.
4. LEXICAL INHERITANCE STRUCTURE: Identification of how a lexical structure is related to other structures in the type lattice, and its contribution to the global organization of the lexicon. [Pustejovsky (1995a: 61)]

Complementàriament a aquests components hi ha una sèrie de mecanismes anomenats generatius que estableixen relacions entre aquests quatre nivells de descripció, i que proporcionen la interpretació componencial dels mots. Aquests mecanismes tenen la forma de transformacions semàntiques que estableixen les condicions de gramaticalitat de les combinacions de tipus:

TYPE COERCION: where a lexical item or phrase is coerced to a semantic interpretation by a governing item in the phrase, without change of its syntactic type.

SELECTIVE BINDING: where a lexical item or phrase operates specifically on the substructure of a phrase, without changing the overall type in the composition.

CO-COMPOSITION: where multiple elements within a phrase behave as functors, generating new non-lexicalized senses for the words composition. This also includes cases of underspecified semantic forms becoming contextually enriched, such as *manner co-composition*, *feature transcription*, and *light verb specification*. [Pustejovsky (1995a: 61-62)]

Les estructures de *qualia*, juntament amb les estructures d'*argument* i les d'*event* constitueixen tres nivells de descripció que interactuen per a construir un llenguatge formal per a la representació de la informació semàntica. Així, es crea un formalisme que consta de dues parts: una jerarquia de tipus sobre cada un dels tres nivells de descripció, i un sistema de restriccions que opera sobre els tipus. D'acord amb Pustejovsky:

What I would like to do is to propose a new way of viewing decomposition, looking more at the generative or *compositional* aspects of lexical semantics, rather than decomposition into a specified *number* of primitives. Briefly, a generative lexicon can be characterized as a system involving at least four levels of semantic representations. These include the notion of *argument structure*, which specifies the number and type of arguments that a lexical item carries; an *event structure* of sufficient richness to characterize not only the basic event type of a lexical item, but also internal, subeventual structure; a *qualia* structure, representing the different modes of predication possible with a lexical item; and a *lexical inheritance structure*, which identifies how a lexical structure is related to other structures in the dictionary, however it is constructed. [Pustejovsky (1995a: 58)]

La principal aportació a la semàntica lèxica del model generatiu de Pustejovsky està en les estructures de *qualia*, que es refereixen a quatre aspectes del significat del mot (constitutiú, formal, tèlic i agentiu). Es pot definir cadascun d'aquests *papers de qualia* de la següent manera:

- a) **Constitutiu**: la relació entre un objecte i els seus constituents, o parts que li són pròpies; assenyala en general el material, pes, parts i components del referent.
- b) **Formal**: el que distingeix l'objecte dins d'un domini més ampli; assenyala en general l'orientació, magnitud, aspecte, dimensions, color i posició del referent.
- c) **Tèlic**: propòsit i funció de l'objecte; assenyala el propòsit que té un agent quan realitza un acte, la funció prototípica o la finalitat que tenen certes activitats.
- d) **Agentiu**: factors relacionats amb l'origen d'un objecte; assenyala en general el creador, l'artefacte o la classe natural que causen el referent.

Pustejovsky caracteritza l'estructura de *qualia* de la següent manera:

What qualia structure tells us about a concept is the set of semantic constraints by which we understand a word when embedded within the language. The *mode* of explanation that characterizes a word as denoting a particular concept is potentially distinct from the manner in which that word is used in the language. That is, definition and word meaning need not have anything to do with grammaticalization or grammatical behavior.

(...)

The qualia provide the structural template over which semantic transformations may apply to alter the denotation of a lexical item or phrase. These transformations are the generative devices such as type coercion, selective binding, and co-composition, which formally map the expression to a new meaning. These operations apply only by virtue of lexical governance relations; that is, their application is conditioned by the syntactic and semantic environment within the phrase appears. For example, when we combine the qualia structure of an NP with that of a governing verb, we begin to see a richer notion of compositionality emerging, one which captures the creative use of words and the “sense in the context” phenomena. [Pustejovsky (1995b: 86)]

L'aportació que fa l'estructura de *qualia* al significat *conceptual* dels lexemes no és, en molts casos, una mera compleció externa o accessòria, establerta per a predir els canvis

que en el llenguatge obeeixen a motivacions de tipus semàntic⁴⁵ (veg. § 4.6). Sovint els tipus recollits en aquesta estructura es refereixen a constituents semàntics caracteritzadors del significat conceptual. Això es pot comprovar a través de les definicions lexicogràfiques (Ruimy & al., 2001) que molt sovint incorporen elements (*fet de, que serveix per a, en forma de, produït per*) que es poden associar a papers de l'estructura de *qualia*, com en els següents articles del DIEC:

beçull *m.* Cassó fet de (CONSTITUTIU) suro que serveix per a (TÈLIC) beure aigua.

palmatòria *f.* Canelobre baix amb peu en forma de (FORMAL) platet i nansa o mànec.

vomitada *f.* Vòmit abundant produït per (AGENTIU) una indigestió.

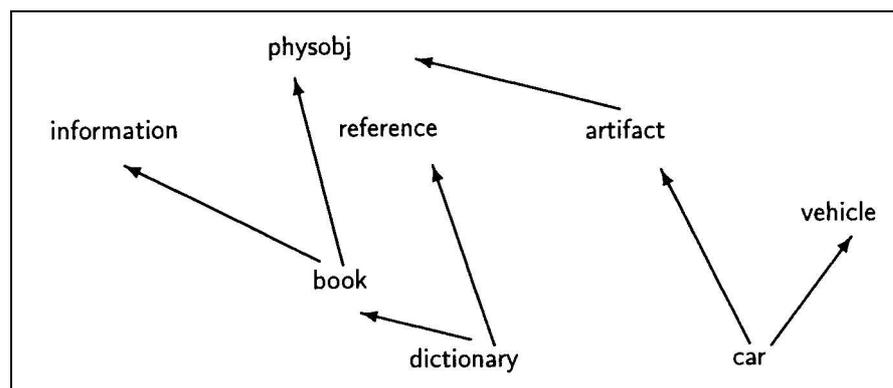


Figura 3.3. Representació convencional de les relacions jeràrquiques del lèxic.

Una de les aplicacions que es poden donar a l'estructura de *qualia* per a la lexicografia és que permeten plantejar de manera diferent les estructures basades en relacions dels lexemes amb els seus tipus semàntics. La representació convencional de les estructures lèxiques basades en la relació d'hiperonímia (o la representació de les associacions dels

⁴⁵ Per a una crítica dels principals aspectes del lèxic generatiu, veg. Fodor & Lepore (2001). Veg. també les respostes a aquesta crítica de Pustejovsky (2001) i de Wilks (2001).

lexemes als seus tipus semàntics) segueix el model representat a la figura 3.3 (que prenc de Pustejovsky, 1995a). En general la representació d'estructures jeràrquiques té una forma semblant a aquesta figura, és a dir, un model de base enumerativa i indistint pel que fa a la naturalesa de les relacions. Així, si volem recollir en aquest tipus d'estructures que un DICCIONARI és alhora una entitat de REFERÈNCIA i un LLIBRE; i que un LLIBRE és, alhora, un OBJECTE físic i una entitat caracteritzada pel fet de contenir INFORMACIÓ, hem de recórrer a una doble vinculació de cadascun d'aquests dos conceptes com el que es veu a la figura 3.3.

Una representació com la de la figura 3.3 planteja certs problemes de tipus teòric, com ara el fet que un DICCIONARI no hereta necessàriament de la mateixa manera que altres tipus de LLIBRE totes les característiques semàntiques d'un LLIBRE com a entitat que conté INFORMACIÓ; això es veu si comparem DICCIONARI amb altres tipus de LLIBRE, com per exemple NOVEL·LA: *llegim* una novel·la, però *consultem* un diccionari; *comencem* (a *llegir*) una novel·la, però no *comencem* un diccionari.

Si ens atenim a un tipus de representació de les estructures jeràrquiques que tingui en compte l'estructura de *qualia* dels lexemes, l'estructura presentada a la figura 3.3 pot convertir-se en la que es representa a la figura 3.4. Com pot veure's, en aquesta representació alternativa els vincles jeràrquics estan etiquetats per referència al tipus de *qualia* amb què poden associar-se (A: agentiu; F: formal; T: tèlic).

Com ja he mostrat en aquesta mateixa secció, l'estructura de *qualia* pot veure's reflectida en la definició lexicogràfica convencional, ja que ambdues es refereixen a aspectes determinants del significat conceptual dels lexemes. Conseqüentment, es pot considerar que els components de la definició que integren aquest tipus d'informacions representen unitats d'informació fonamentals en la caracterització semàntica del mot (veg. Labropoulou & al., 2000). Des d'un cert punt de vista, per tant, podem establir l'existència, en una definició lexicogràfica, de components sobre els quals pot resultar més negatiu l'impacte de desviacions com ara la circularitat. Més endavant (§ 5.1) m'ocupo dels aspectes pràctics que aquests problemes poden tenir en lexicografia.

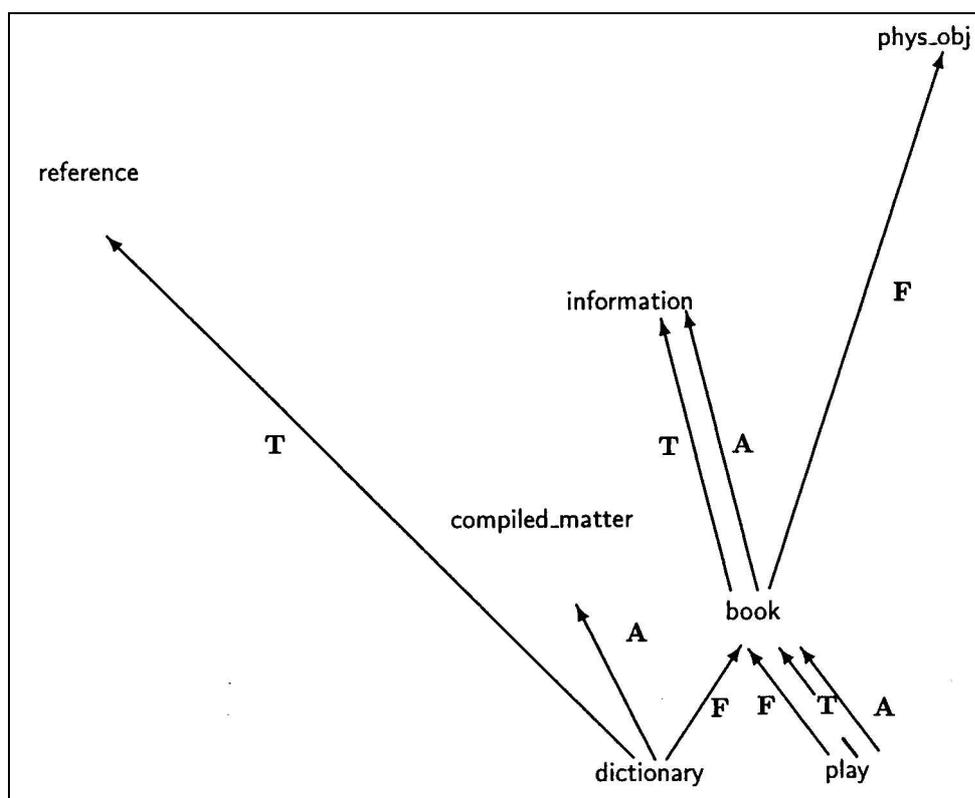


Figura 3.4. Representació multidimensional de les relacions semàntiques del lèxic.

Les estructures de *qualia* s'han implementat en els lèxics electrònics constituïts en el marc del projecte europeu SIMPLE (LE4-8346), entre els quals hi ha el lèxic català, desenvolupat per l'Institut d'Estudis Catalans. SIMPLE (Busa & al., 2001) es basa en el projecte previ PAROLE (LE2-4017), que va crear lèxics computacionals harmonitzats a partir d'un model comú per a 12 llengües europees (veg. més enrere nota 31), entre les quals hi ha el català. Els lèxics de PAROLE contenen informació referida als nivells morfològic i sintàctic de 20.000 entrades per llengua. SIMPLE afegeix el nivell semàntic a una part dels lèxics de PAROLE, i codifica 10.000 sentits (7.000 per a noms, 2.000 per a verbs i 1.000 per a adjectius) seleccionats a partir de criteris freqüencials. A fi de superar els inconvenients descriptius de les jerarquies monodimensionals, SIMPLE codifica per a les descripcions semàntiques de les seves entrades una sèrie de relacions semàntiques associades als diferents *qualia*.

Cadascun dels quatre papers de *qualia* es representa com una relació que, al seu torn, es descompon en relacions més específiques, que representen subtipus d'informació d'un determinat *quale*. Aquestes jerarquies de relacions referides als quatre papers de *qualia* conformen allò que en el model de SIMPLE s'anomena *extended Qualia set*. A la figura 3.5 es representa una d'aquestes jerarquies de relacions, concretament la referida al paper *constitutive*. En negreta assenyalo els nodes formals de la jerarquia que representen agrupacions de relacions, però no relacions en si; en aquest quadre això correspon als tipus *location* i *property*, que són subtipus de *constitutive*. Aquestes relacions en el model de SIMPLE s'estableixen entre unitats semàntiques (és a dir, *sentits* o *accepcions*, en terminologia lexicogràfica) de les entrades lèxiques. L'expressió d'aquestes relacions segueix la presentació següent:

Relació (<Unitat_semàntica_1>, <Unitat_semàntica_2>)

La <Unitat_semàntica_1> és el primer terme i, per tant, el subjecte de la relació; mentre que la <Unitat_semàntica_2> és el terme relacionat, o l'objecte de la relació.

Heus ací alguns exemples de relacions constitutives referides a diferents unitats semàntiques, representades aquí pels mots a què pertanyen els sentits relacionats:

Is_a_member_of (<diputat>, <parlament>)

Has_as_member (<bandada>, <ocell>)

Is_a_follower_of (<marxista>, <marxisme>)

Is_a_part_of (<cap>, <cos>)

Has_as_part (<avió>, <ala>)

Made_of (<pa>, <farina>)

Is_in (<oasi>, <desert>)

Lives_in (<italià>, <itàlia>)

Has_as_colour (<llimona>, <groc>)

Produces (<ocell>, <ou>)

Produced_by (<mel>, <abella>)

Property_of (<intel·ligència>, <intel·ligent>)

Concerns (<hepatitis>, <fetge>)

Contains (<llibre>, <informació>)

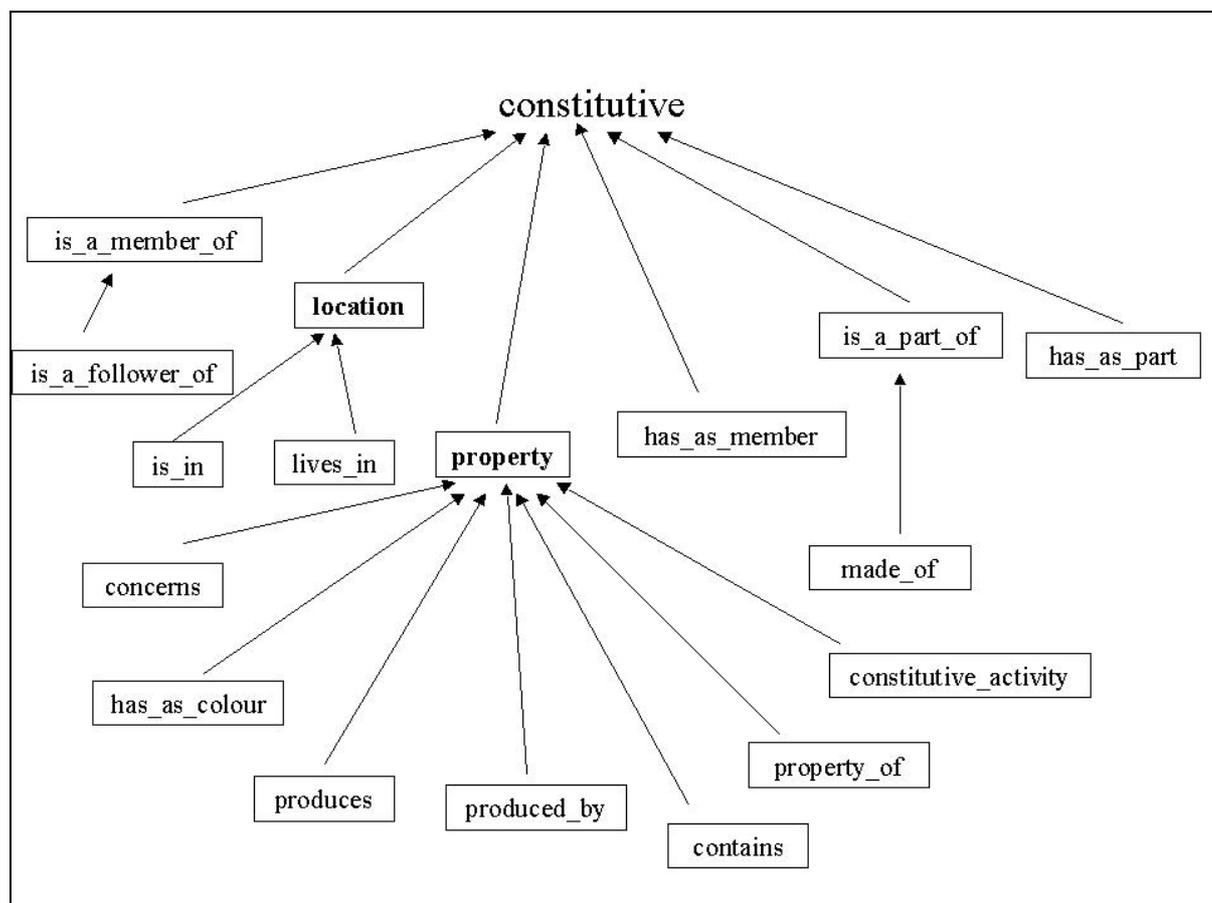


Figura 3.5. Representació de les relacions del *Constitutive role* en el model de *SIMPLE*.

La codificació d'aquestes informacions per a cada lexema permet la caracterització de certs fenòmens d'importància en la descripció lexicogràfica, com ara la modificació semàntica selectiva (veg. Soler, 2001) que tracto amb més detall a § 4.6.

El model de lèxic generatiu, i d'una manera molt especial les estructures de *qualia*, aplicades a l'anàlisi de l'estructura dels diccionaris, suposen aportacions d'interès per a la representació de les projeccions conceptuals que es produeixen en la descripció

lexicogràfica. La multidimensionalitat d'aquestes estructures està en funció de la complexitat de les relacions que contrauen els lexemes entre si. En la definició lexicogràfica es verifiquen diversos tipus de relacions que estableixen vincles entre els elements que són difícils d'expressar a través d'estructures monodimensionals, o a partir d'un nombre restringit de relacions lèxiques.

4 Definició lexicogràfica i equivalència

4.1 Equivalència i identitat

Entre els diferents *tipus de significat*⁴⁶ que es poden establir en l'anàlisi semàntica del lèxic, la *denotació* o *significat conceptual* és el que tradicionalment s'ha pres com el principal objectiu de la descripció lexicogràfica. El significat conceptual també ha estat el principal objecte d'estudi de la semàntica lèxica, tant des de les aproximacions basades en el componencialisme com des de les aproximacions basades en les relacions semàntiques (Lyons, 1968, 1977, 1995; Cruse, 1986).

La definició lexicogràfica, que es basa fonamentalment en la relació de sinonímia, té com a objectiu principal la caracterització de la denotació del definit. De tota manera, si provem de determinar quin és l'objecte estricte de descripció en lexicografia, cal posar de manifest dues qüestions relacionades amb la identificació entre l'explicitació del significat i la descripció lexicogràfica:

a) D'una banda, quina importància tenen els factors considerats no pròpiament semàntics des del punt de vista del significat conceptual, sinó que tenen a veure amb les propietats sintagmàtiques dels mots, i amb les seves característiques esdevenimentals, aspectuals, subcategorials, etc.

b) D'una altra banda, quina importància tenen, com a objecte de descripció lexicogràfica, la resta de *significats* aïllats per Leech (1971), en especial els

⁴⁶ És especialment interessant la classificació de Leech (1971), que distingeix entre set tipus de significat, i que segueixo en aquest punt concret del treball.

significats *associatius* (el significat *connotatiu*, el *social*, l'*afectiu*, el *reflex*, i el *col·locatiu*).

La incidència d'aquestes dues qüestions que, sense pertànyer estrictament al significat conceptual, es relacionen amb el significat lèxic, és diversa, i en determinats casos, si no es tracten adequadament, són susceptibles d'introduir elements de distorsió en la relació sinonímica definit-definició.

Les possibilitats de tractament, dins del diccionari, dels aspectes recollits en *a)* i en *b)* també són diverses. En general, l'explicitació o la integració de *b)* en el diccionari es resol a través del seu tractament en diferents elements estructurals (l'explicitació de les propietats sintàctiques, de les marques valoratives, de les col·locacions, etc.). La integració d'*a)*, però, pot exigir la formulació de certs ajustaments en la definició lexicogràfica, ja que incideix d'una manera més directa en l'expressió del significat conceptual del lexema.

El principi que se segueix en la descripció lexicogràfica per a avaluar l'adequació descriptiva de la definició és molt simple, però alhora difícil de caracteritzar en el pla teòric: una definició lexicogràfica és en termes generals una paràfrasi que, com he assenyalat, manté amb el mot definit una relació de sinonímia.⁴⁷ L'adequació d'una paràfrasi al terme definit es verifica a través del test de substituïbilitat de l'un per l'altra en els contextos que donen contingut a una accepció determinada.

La determinació de la naturalesa de la relació subjacent a aquesta condició de substituïbilitat no és una qüestió trivial. Fuchs (1982) fa una crítica del concepte

⁴⁷ La relació de sinonímia és un concepte no exempt de certs problemes de caracterització (Casas, 1999); aquí el prenc d'una manera acrítica, per fer referència a la relació d'equivalència significativocconceptual existent entre un terme definit i una expressió definidora. Fuchs (1982) tracta els problemes teòrics i metodològics que planteja la paràfrasi en lingüística, i s'ocupa amb preferència dels aspectes referencials i situacionals que validen la relació entre paràfrasis. Aquí tractaré sobre algunes qüestions relacionades amb les de Fuchs però des d'un altre punt de vista, que crec més pròpiament lingüístic i més allunyat de la filosofia del llenguatge.

d'*identitat* com a denominador de la relació sinonímica entre paràfrasis, i el situa en les aproximacions fonamentals amb què s'hi ha referit la lingüística teòrica:

Ainsi se trouve instaurée une nouvelle problématique: constat de l'impossibilité d'établir en langue une relation d'identité sémantique totale (de synonymie absolue) et insistance sur les différences sémantiques entre les séquences candidates à la paraphrase. A partir de là, deux positions sont possibles, et se sont, de fait, trouvées attestées: ou bien un désintérêt à l'égard de la paraphrase (il n'y aurait, en définitive, pas de relation paraphrastique en langue, puisqu'à chaque fois que l'on croit en tenir une, il est possible de montrer qu'en réalité il existe des différences sémantiques), ou bien au contraire un regain d'intérêt pour la paraphrase (nécessité d'abandonner cette idée "peu subtile" de l'identité de sens, et de pousser plus avant l'analyse de la paraphrase en termes de ressemblance, de proximité, d'équivalence sémantique). La première conclusion est celle qui, de fait, semble avoir prévalu chez la plupart des transformationalistes: l'enthousiasme pour la paraphrase a très largement décliné depuis les années 70 – le constat de l'impossible synonymie absolue paraît avoir balayé, du même coup, la paraphrase elle-même. La seconde conclusion a été tirée par les tenants d'une approche globale (c'est à dire syntaxico-sémantique) de la paraphrase. [Fuchs (1982: 52)]

La noció d'*equivalència*, d'una forma més precisa que la noció d'*identitat*, fa referència, segons Fuchs, a les capacitats d'intercanvi en determinats contextos. De manera que un enunciat *a* és equivalent a un enunciat *b* si l'un i l'altre poden vehicular el mateix contingut semàntic en un nombre *n* de contextos. Des d'una perspectiva no formalitzada, la identitat fa referència a la possibilitat d'associar a dos enunciats unes totals possibilitats de substitució sense l'alteració ni la modificació de cap dels components significatius lingüísticament rellevants. En paraules de Fuchs:

Si l'on renonce à caractériser la paraphrase comme relation d'identité sémantique totale (synonymie absolue), il est possible d'essayer de l'appréhender comme une relation d'équivalence sémantique, fondée sur l'existence d'un noyau sémantique commun sur lequel se greffent des différences sémantiques secondaires (cf. les notions de "pseudo-", "quasi-" ou "para-" synonymie, en matière de lexique). Entre deux phrases déclarées paraphrastiques, il y aurait donc à la fois du "pareil" et du "pas pareil": tout le travail du linguiste consisterait, dans une telle optique, à définir la nature des sémantismes communs et des sémantismes différentiels, et à établir des degrés d'équivalence entre phrases, selon le type et le nombre des éléments sémantiques communs à deux paraphrases. [Fuchs (1982: 53)]

Tot i que pel que sembla la distinció de Fuchs és bàsicament terminològica,⁴⁸ i està exempta de fonament formal, crec que la manera adequada de referir-se a la relació existent entre definit i definició és amb l'expressió *equivalència* (tot i els usos diferents que sol tenir aquest mot en lexicografia),⁴⁹ més que no pas amb la d'*identitat* o la d'*igualtat*, seguint la descripció que fa el DPR d'*equivalència* i dels mots que s'hi relacionen, especialment la que es recull com a sentit d'*equivalència* en lògica:

équivalence (...) n. f.

1♦ Qualité de ce qui est équivalent. ⇒ **adéquation, égalité, homologie, identité.**

Les jacqueries «mettent en avant un principe d'équivalence, vie contre vie» (Camus). (...)

2♦ Math. Relation d'équivalence sur un ensemble E: relation binaire sur E, réflexive, symétrique et transitive. Le parallélisme est une relation d'équivalence sur l'ensemble des droites du plan. Classes d'équivalence. — **Log.** Opérateur d'une proposition complexe (notée \Leftrightarrow , \equiv ou \sim), par lequel cette proposition est vraie si les propositions élémentaires sont toutes deux vraies ou toutes deux fausses (⇒ **congruence, modulo**). — **Phys.** Principe de l'équivalence mécanique de la chaleur. ⇒ **2. équivalent.**

⁴⁸ A més de les observacions de Fuchs (1982), es pot fer referència al fet que el concepte d'*identitat*, és subjacent al predicat lògic que sol caracteritzar-se amb el signe “=”. Proposicions del tipus “*a* és igual a *b*” o “*a* és el mateix que *b*” es poden formalitzar com a “*a* = *b*”. L'estabilitat d'aquest concepte lògic no es manté en certes expressions del llenguatge corrent, fora de la lògica i el càlcul matemàtic. Quan diem *aquest vas i aquest altre són iguals* parlem, inequívocament, de dos vasos diferents, mentre que quan diem *hem fet servir el mateix vas*, parlem d'un sol vas. Crec que aquest motiu pot ser considerat complementari per a l'elecció del terme *equivalència* en lloc d'*identitat* per referir-me a la paràfrasi de definició.

⁴⁹ Adopto aquest terme per referir-me a la relació entre definit i definició malgrat que, en lexicografia, *equivalència* sol utilitzar-se per referir-se a la relació entre entrada i correspondència als diccionaris bilingües. Martínez de Sousa (1995: 193) recull aquesta tendència i afirma: “Debe distinguirse entre voces sinónimas i voces equivalentes: los sinónimos pertenecen a una sola lengua (...), mientras que los equivalentes se refieren a los mismos términos en dos o más lenguas y no son intercambiables”. Malgrat aquesta tendència, crec que *equivalència* és la denominació més adequada per a referir-se genèricament a aquest tipus de relació en la lexicografia monolingüe.

équivalent, ente (...) adj.

1 ♦ Dont la quantité a la même valeur. ⇒ **égal**. Leurs parts d'héritage sont équivalentes. «J'inflige aux trois maîtres [...] une punition équivalente, huit jours d'arrêt» (Loti). Indemnité équivalente au dommage. (...)

2 ♦ Qui a la même valeur ou fonction. *Ces deux expressions sont équivalentes.* ⇒ **synonyme**. *Faits équivalents*, dont la portée est semblable ou identique. ⇒ **comparable**. *Concepts équivalents.* ⇒ **identique, similaire**. «*Toutes les phrases de son livre [l'Étranger, de Camus] sont équivalentes, comme sont équivalentes toutes les expériences de l'homme absurde*» (Sartre).

équivaloir (...) v. tr. ind. (...)

1 ♦ Avoir la même valeur en quantité que. En valeur nutritive, deux cents grammes de poisson équivalent à cent grammes de viande. Pour une somme équivalant à la moitié du prix. ⇒ **1. équivalent**. — Pronom. Les deux choses s'équivalent. Cela s'équivaut.

2 ♦ Avoir la même valeur ou fonction que. ⇒ **signifier**. Cette réponse équivalait à un refus. ⇒ **revenir** (à). «Pour Gilieth la vie d'un homme équivalait à celle d'un lapin» (Mac Orlan). ⇒ **valoir**. «l'inflexion était si douce, si compatissante, si timide, qu'elle équivalait au plus tendre aveu» (Martin du Gard).

La relació d'equivalència entre un mot i la seva definició lexicogràfica (que s'ha denominat *equació sèmica*)⁵⁰ és una relació implícita entre el definit *A* i la seva definició *B*. A aquesta relació, no expressada explícitament en el diccionari però subjacent en el seu valor semiòtic (veg. Rey-Debove, 1969, 1971), se li poden conferir valors del tipus: *A significa B*, *A designa B*, *A denota B*, etc.

L'anàlisi i la caracterització de les definicions i de llurs tipus constitueix un tema recurrent en l'estudi de la lexicografia⁵¹ i, més concretament, en els estudis referits als

⁵⁰ Per a una presentació del concepte d'*equació sèmica* i una revisió crítica de les principals aportacions teòriques aplicades a la lexicografia, veg. Lara (1997).

⁵¹ Robinson (1954), amb referències indirectes a la lexicografia, estableix set classes de definició, que després han estat adoptades en un cert nombre de treballs sobre classificació de les definicions lexicogràfiques. Aquestes classes, que ell anomena *methods of definition*, són:

a) El mètode sinonímic: "This is de commonest method in dictionaries, especially in two-language dictionaries and in very small one-language dictionaries." [Robinson (1954: 95)]

diccionaris monolingües. El nivell de profunditat assolit en la caracterització de les definicions⁵² no ha anat, però, més enllà de l'establiment de tipologies no compartides generalment i de la determinació d'algunes característiques que reuneixen les definicions descriptivament adequades. Els tractaments sobre diccionaris convencionals en format electrònic que ha fet la lingüística computacional per a l'extracció de coneixement lèxic han donat resultats més tangibles (veg. per exemple Hurk & Meijs, 1986) des del punt de vista quantitatiu, però d'un valor metodològic no gaire més destacable.

b) El mètode analític: “To indicate what specific object a word means, name a bigger class within which that object falls, and then name something that distinguishes it from the rest of the class.” [Robinson (1954: 96)]

c) El mètode sintètic: “The thing meant is assigned to its place in a system of relations, synthesized into a whole with other things. Whereas the analytic method indicates the thing meant by showing it as a whole of parts, the synthetic does so by showing it as part of a whole.” [Robinson (1954: 98)]

d) El mètode implicatiu: “We may also call it the ‘contextual’ method, because it puts the word in a context which determines its sense.” [Robinson (1954: 107)]

e) El mètode denotatiu: “consists in mentioning some part of the denotation of the word, some one or more of the things to which the word is applied.” [Robinson (1954: 109)]

f) El mètode ostensiu: “There must therefore be at least one more method of defining a word, and it must be a method that can dispense with words altogether (except the word being defined).” [Robinson (1954: 117)]

g) El mètode de formulació de regles (*rule-giving method*); “The rule of the word ‘I’ is that it is to be used by each utterer to indicate himself.” [Robinson (1954: 126)]

Aquests diferents mètodes es poden distingir a partir de diferents criteris. Un dels més importants és potser que els cinc primers utilitzen mots per a la definició de mots, mentre que *f*) i *g*) estableixen relacions amb elements no lingüístics, tot i que de fet els diccionaris no són estranys al mètode *g*), que s'utilitza habitualment en la definició de mots gramaticals.

⁵² La meua intenció no és ocupar-me, en aquest treball, de fer una crítica de les classificacions i de les tipologies de la definició lexicogràfica que s'han fet fins ara, ja que simplement m'interessa aquí de determinar el tipus de definició que en lexicografia es pot considerar genèricament com *substituible*. Per a una anàlisi crítica i comparativa de les diferents tipologies que s'han fet en aquest sentit, veg. Rafel (2002).

Una de les particularitats més característiques de la definició lexicogràfica és que s'expressa a través d'un enunciat lingüístic. Els problemes de la utilització del llenguatge convencional com a *discurs metalingüístic* propi dels diccionaris s'ha tractat amb un cert detall, tant des del punt de vista semiòtic com estructural.⁵³ Les idees fonamentals sobre la teoria del discurs lexicogràfic es basen en el fet que ja he assenyalat: aquest discurs sol tenir la forma d'una proposició que manté respecte al mot que encapçala l'entrada una relació d'implicació bilateral molt similar a la sinonímia. Martin (1977a) anomena *parafràstiques* aquest tipus de definicions, mentre que reserva el terme de definicions *metalingüístiques* a les expressades en allò que Rey-Debove (1971) anomena *metallengua de signe*, és a dir, no basades en la descripció del significat, sinó adreçades al *signe* en tant que significant i significat.

A partir de la distinció precedent, es pot parlar, com fa Seco (1978), de la diferenciació bàsica que s'estableix entre allò que podem anomenar definicions “pròpies” (parafràstiques, en la terminologia de Martin), d'una banda, i definicions “impròpies” (o més exactament, amb elements impropis, o *metalingüístiques*), de l'altra.⁵⁴

⁵³ La noció de metallenguatge aplicada al discurs de la definició lexicogràfica i a la seva caracterització ha estat una idea treballada amb profunditat per Rey-Debove (1967, 1969, 1971). Aquesta autora parteix de la base que la descripció lexicogràfica s'expressa a partir de dos eixos metalingüístics fonamentals: la metallengua de signe (referida al mot com a significant i significat) i la metallengua de contingut (referida al significat conceptual del mot).

⁵⁴ No cal entrar a fons en aquesta qüestió (veg. també més enrere nota 17), ja que per raons òbvies m'interessen només les definicions pròpies. Però en tot cas potser val la pena d'assenyalar que, almenys fins a cert punt, es poden distingir diversos graus d'impropietat en les definicions lexicogràfiques. Sembla que les definicions “impròpies” són l'única possibilitat de descripció de determinades classes gramaticals (articles, pronoms, preposicions, alguns adverbis, etc.), mentre que en alguns casos la incorporació d'elements de metallengua de signe en les definicions lexicogràfiques de substantius, adjectius i verbs es pot interpretar com una distorsió.

Les definicions pròpies s'han classificat a partir de diferents criteris (simples-múltiples,⁵⁵ parafràstiques-sintètiques), però tal vegada una de les diferenciacions de classe més operatives que s'han establert entre les definicions és la que Rey-Debove (1967) introdueix sota la denominació de definicions *substancials* i definicions *relacionals*. Les definicions substancials es basen sempre (per inclusió o per exclusió) en un mot de la mateixa categoria gramatical que el definit, el qual actua com a descriptor. En les definicions relacionals l'equivalència categorial definit-definició es produeix en el nivell sintàctic, ja que la parafrasi definidora és un constituent de frase complex que és la projecció categorial del definit. La definició substancial és l'únic tipus possible en el cas dels substantius i verbs, mentre que tots dos tipus poden ser usats en definicions d'adjectius i d'adverbis.

Diversos treballs, com el de Martin (1983), intenten establir una classificació tipològica global dels diferents tipus de definició lexicogràfica. Martin parteix de l'esquema que mostra la figura 4.1 (que prenc de Martin, 1983: 63) com a classificació componencial de les definicions per arribar a la integració d'una proposta general de classificació de les definicions que reproduïxo a la figura 4.2.

⁵⁵ La distinció entre definicions simples i definicions múltiples no és rellevant per als propòsits d'aquest treball. L'acumulació, en l'espai corresponent a l'element estructural definició, de diverses definicions que concorren a donar un sentit d'un mot pot ser fàcilment vista com una desviació metodològica i, per tant, no considerable com un paràmetre de classificació. Entre els comentaris sobre aquest tipus de definicions es poden adduir: Rey-Debove (1971: 208); “Il y aurait avantage à abandonner cette tradition de définition multiple et redondante dans les dictionnaires, qui introduit un élément supplémentaire de confusion.”; Martin (1977a) qualifica la definició múltiple de “une des plaies de la lexicographie”, i Rafel (1988: 64) diu que “és una mostra de la incapacitat del lexicògraf per a trobar la definició adequada, que faci explícit d'una manera suficient el significat de què es tracta”, i que “en una activitat lexicogràfica coherent i normalitzada aquesta manera de procedir no pot ésser admesa”.

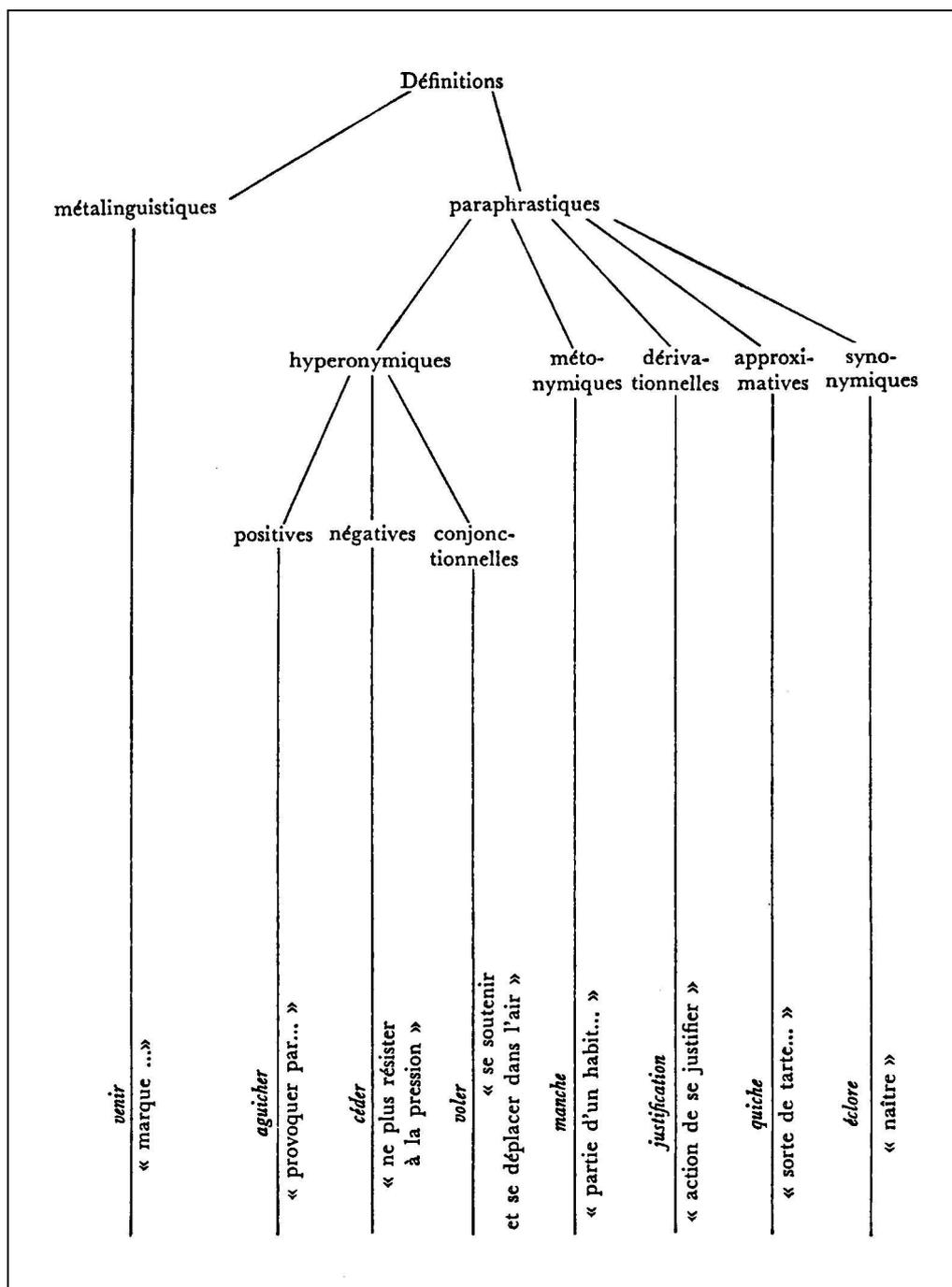


Figura 4.1. Classificació componencial de les definicions de Martin (1983).

La motivació principal de Martin no és, però, l'anàlisi de les condicions de formulació de les definicions sinó el tractament de les propietats lògiques dels diferents tipus

d'enunciats de definició. L'establiment d'una classificació general per a *tots* els tipus de definició lexicogràfica sol presentar l'inconvenient que, correntment, un mateix enunciat de definició pot pertànyer a més d'un tipus. A més, aquestes classificacions solen presentar l'inconvenient metodològic d'estar basades en criteris de classificació extralingüístics (semiòtics, en general, pel que fa sobretot a la tradició francesa) i no tenen en compte l'anàlisi lingüística dels constituents de la definició.

Per a les finalitats d'aquest treball, aquest tipus de classificacions presenten l'inconvenient de no caracteritzar globalment, des del punt de vista lingüístic, allò que podríem anomenar la *definició lexicogràfica substituïble*. Les possibilitats que una mateixa classificació tipològica global de la definició lexicogràfica compleixi aquest requisit és un objectiu que cau fora de l'àmbit d'aquest treball, de manera que, en principi, crec més adequat constatar l'existència de diversos criteris de classificació (veg. Rafel, 2002) que estableixen grups de definicions que es comporten d'una o altra manera en relació a tal o tal criteri.

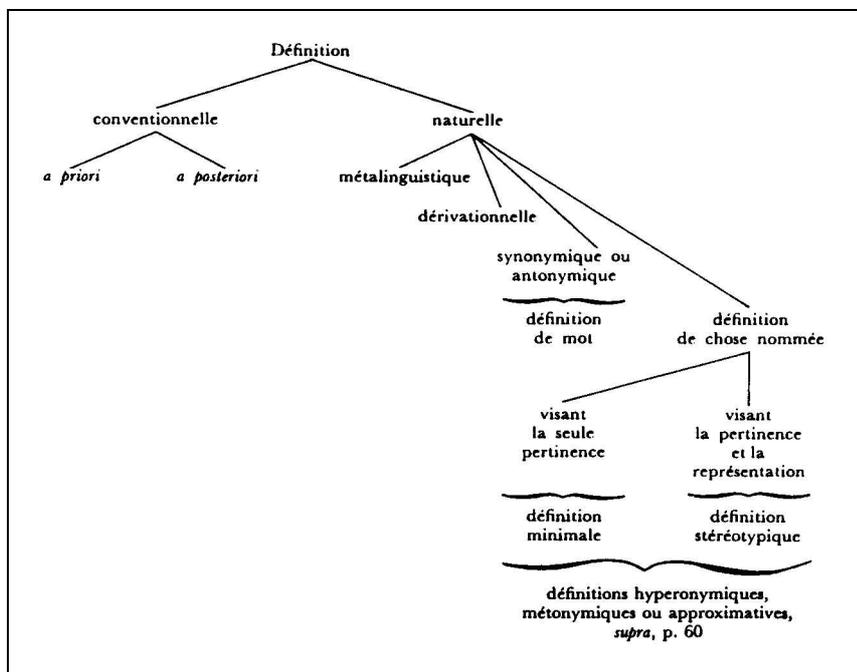


Figura 4.2. Proposta general de classificació de les definicions segons Martin (1983).

Una de les aportacions més interessants, basada exclusivament en criteris lingüístics, per a la caracterització de grup de les definicions substituïbles és la que fa Ilson (1987), que parteix de la base següent:

Definition is the de-lexicalisation of a lexical unit into semantic and syntactic components which are then presented in a single phrase whose content characterises the definendum semantically while its form characterises the definendum syntactically the definendum syntactically. [Ilson (1987: 71)]

Així, una definició substituïble es caracteritza per donar els quatre tipus d'informacions següents:

Syn₁: la categorització sintàctica del definit: *nom, verb, etc.*

Syn₂: la subcategorització sintàctica del definit: *transitiu, intransitiu, etc.* per als verbs.

Sem₁: la categorització semàntica del definit: “gener és un *mes*”.

Sem₂: la subcategorització semàntica: “gener és el *primer mes*”.

En una definició lexicogràfica aquests quatre elements poden aparèixer de forma aïllada o combinada. La manera com aquests quatre elements es combinen dins d'un enunciat de definició pot establir diversos tipus de definicions substituïbles: en una definició sintètica (o sinonímica) els **Syn₁**, **Syn₂**, **Sem₁** i **Sem₂** es troben concentrats en un sol constituent (el sinònim):

silent *adj.* Silenciós.

En aquest article de diccionari, la definició, que té un únic component, concentra en aquest component tots els elements de descripció assenyalats per Ilson (1987), és a dir:

silent: [silenciós]_{Syn, Sem}

En les definicions no sintètiques (perifràstiques), però, aquests elements es distribueixen d'acord amb la manera com es combinen els diferents constituents de la definició. Si la combinació d'aquests elements manté concentrats en un sol constituent el **Syn₁** i el **Sem₁**, aleshores es trobem davant d'una definició *analítica* (definició per assignació de *genus* i *differentia*) que sovint s'ha anomenat *hiperonímica*:

solatge: [Pòsit]_{Syn, Sem1} [que deixa un líquid]_{Sem2}

fisonomia: [Aspecte]_{Syn1, Sem1} [exterior]_{Sem2} [d'alguna cosa]_{Syn2}.

També pot passar que la informació principal sintàctica (Syn₁) i la semàntica (Sem₁) es presentin en grups diferents; quan es dóna aquest cas, Ilson (1987) anomena aquestes definicions *formulàries* (*formulaic*):

fertilitzar: [Fer]_{Syn} [fèrtil]_{Sem}

inflexió: [Acció de]_{Syn} [corbar]_{Sem}

mental: [Relatiu a]_{Syn} [la ment]_{Sem}

L'anàlisi d'Ilson (1987) permet, amb molta més eficiència que la de Rey-Debove (1967) l'establiment de distincions precises entre definicions *relacionals* i *substancials*. De fet, permet caracteritzar separadament les definicions pròpiament hiperonímiques de les que introdueixen altres tipus de relació, ja que en el primer cas **Syn₁** i **Sem₁** coincideixen en el mateix element (el descriptor), mentre que en l'altre es troben repartits en dos elements: *conjunt de X*, etc. el primer element constitueix el *descriptor sintàctic*, mentre que X constitueix el *descriptor semàntic* de la definició.

4.2 Substituïbilitat definit-definició

Tal com ja he dit a § 4.1, la simplicitat amb què es pot enunciar el programa de descripció del significat conceptual en lexicografia s'ha de contraposar a la complexitat que suposen els diferents tipus de significats (en el sentit de Leech, 1971) que s'associen a un mot. Un cop determinades amb precisió quines són les condicions

bàsiques per a la substituïbilitat definit-definició, convé plantejar-se fins a quin punt la definició lexicogràfica ha d'integrar els aspectes no pròpiament conceptuals del significat lèxic, en la línia dels aspectes complementaris que he assenyalat al començament de § 4.1. La qüestió no té una resposta senzilla, i fa posar en qüestió el principi de *substituïbilitat* com a prova de verificació fonamental de l'adequació de la definició lexicogràfica.

Les possibilitats de substitució depenen fonamentalment de la satisfacció dels diferents tipus d'*equivalències* que podem establir entre el terme definit i la definició. Agrupant els diversos fenòmens integrats en el significat lèxic podem distingir, almenys, quatre tipus d'*equivalències* entre definit i definició que van més enllà de la mera equivalència conceptual. Aquests quatre tipus d'*equivalència* són els que considero en les seccions següents d'aquest treball: l'*equivalència argumental* (§ 4.3), l'*actancial* (§ 4.4), l'*esdevenimental* (§ 4.5) i la *sincategoremàtica* (§ 4.6).

La substituïbilitat ha estat assumida com un principi d'adequació metodològica, però (almenys en la seva aplicació en lexicografia) no s'ha analitzat mai críticament: ni a partir del concurs de dades empíriques ni en relació a les aportacions de la lingüística teòrica. Determinades aportacions de la semàntica lexicalista més recent poden resultar de gran interès per a l'anàlisi d'aquesta qüestió, i poden matisar, fins a cert punt, el principi general de la substituïbilitat com a prova d'adequació. El fet que no s'hagin considerat fins ara seriosament els components semàntics que podem denominar inherents a la semàntica del mot, però que alhora poden aparèixer en determinades condicions realitzats lèxicament (com a arguments del mot en qüestió) en contextos gramaticals, és un dels principals elements d'anàlisi amb què cal matisar la prova de la substitució.

Els treballs que han fet referència a aquesta qüestió ho han fet, generalment, des d'una òptica parcial o des d'un punt de vista diferent del que adopto aquí. S'ha remarcat, sobretot a partir dels treballs de Seco (1979) i de Rafel (1989), la importància que adquireix, per a la normalització dels diccionaris convencionals, el marcatge dels

elements del *contorn* (com els anomena Seco) o dels elements *extrínsecs* (com els anomena Rafel) en la definició lexicogràfica.

Una definició conté necessàriament una sèrie d'elements que Rafel (1989) anomena *intrínsecs* i que compleixen, en el millor dels casos, la condició de ser equivalents al terme definit. Algunes definicions lexicogràfiques incorporen, també, un altre tipus d'elements que podem anomenar *extrínsecs* (veg. també Rafel, 1989) i que, sense intervenir en els processos de substitució, expliciten característiques sintàctiques o semàntiques que s'han de complir com a condicions per a la substitució. L'ús, el marcatge i la diferenciació dels elements *extrínsecs* varia d'un diccionari a un altre, però en les obres lexicogràfiques concebudes amb criteris més moderns s'observa una tendència cap a una major explicitació dels criteris d'inclusió i de tractament, i cap a un marcatge menys ambigu.

Aquestes remarques metodològiques corresponen a grans trets a intents de sistematització de la indicació de l'anomenat *complement directe potencial*,⁵⁶ i de la indicació de les restriccions seleccionals de subjecte, quan aquests elements presenten diferències sintàctiques o semàntiques respecte al descriptor de la definició.⁵⁷ En el cas

⁵⁶ Els diccionaris més tradicionals han tractat aquests tipus d'elements en les definicions verbals indicant-los entre parèntesis. Es pot dir de passada que l'ús dels parèntesis per al marcatge de diferents tipus de fenòmens és habitual, i no m'interessa ara d'establir-ne una casuística, sinó només prendre com a referència aquells casos en què els parèntesis són utilitzats per vehicular l'expressió dels elements *extrínsecs*. En algunes ocasions, els diccionaris indiquen, bé que d'una manera no estrictament formalitzada, el subjecte en els casos en què aquest presenta una restricció selectcional del descriptor. De manera més o menys implícita, doncs, la lexicografia convencional recull algunes d'aquestes informacions i les distingeix del contingut *intrínsec* de la definició.

⁵⁷ El marcatge d'aquests constituents de la definició té, en el cas del català, la seva primera manifestació amb l'aparició, l'any 1932, del DGLC. Fabra introdueix aquesta notació (manllevada del *Dictionnaire général de la langue française du commencement du XVII^e à nos jours*, de Hatzfeld & Darmesteter, circa 1900) consistent en el marcatge entre parèntesis, dins de la definició, del complement directe potencial. En ocasions el DGLC també especifica el subjecte del definit, encapçalant la definició, i separat d'aquesta per una coma. El tractament de l'extrinsicitat en el DGLC, però, presenta diferents graus de sistematicitat: mentre que el complement directe potencial és marcat més o menys sistemàticament, l'explicitació del subjecte hi és en un grau molt menor.

de Rafel (1989), aquesta sistematització arriba a un grau notable de rigor formal i analític. De tots dos treballs es desprèn la idea més o menys implícita que, per a un sistematització més aprofundida dels constituents extrínsecs de la definició, cal incidir en dos problemes específics:

a) La *intercategorialitat* del fenomen, ja que la necessitat d'utilització d'elements extrínsecs en les definicions pot ser fàcilment projectada a qualsevol tipus de categoria predicativa. De l'aprofundiment en aquesta intercategorialitat se'n podria derivar fàcilment una formulació més àmplia del paper dels elements extrínsecs en la definició lexicogràfica que comportarien un major grau de sistematització de la descripció lexicogràfica.

b) La *diversitat* dels complements susceptibles de ser tractats com a elements extrínsecs, ja que les condicions de substituïbilitat no es limiten al tractament del complement directe potencial. Cal tenir en compte, doncs, altres tipus de complements (en general complements preposicionals) que formen part de l'estructura de subcategorització de mot definit.

L'aprofundiment en aquests dos aspectes incideix directament en la complementació dels treballs que fins ara s'han ocupat de l'extrinsicitat, i es relaciona també amb certs aspectes de naturalesa semàntica que limiten en certa mesura les possibilitats de substitució. Cal analitzar amb detall algunes de les condicions semàntiques de l'equivalència a fi de mostrar, d'una banda, les limitacions que té aquesta noció com a prova d'adequació en lexicografia i, d'una altra, possibles ajustaments metodològics que poden contribuir a millorar la descripció lexicogràfica.

4.3 Equivalència argumental

Ja he mostrat a § 4.2 que, per a assegurar la plena correspondència entre definit i definició, la paràfrasi definidora ha de tenir, en moltes ocasions, una estructura

complexa on s'integren elements de dues naturaleses diferents: un discurs (intrínsec) descriu el significat en termes d'un enunciat que compleix el principi de substitució, mentre que en un altre nivell discursiu (extrínsec) s'expliciten les característiques argumentals necessàries perquè aquesta substitució es porti a terme sense problemes:

digital *v. tr.* Anotar (en una partitura) les xifres d'una digitació. [DIEC]

tatuar *v. tr.* Fer (en el cos) dibuixos indelebles inserint una matèria colorant sota la pell. [DIEC]

acomplexar *v. tr.* Fer adquirir un complex (a algú). [DIEC]

esmolar *v. tr.* Fer agut el tall o la punta (d'una eina, d'una arma). [DIEC]

guaitar *v. tr.* Vigilar amb l'esguard (un indret). [DIEC]

inhabilitar *v. tr.* Fer (algú) inhàbil. [DIEC]

A partir dels exemples precedents podem verificar les dues funcions principals que s'han associat als elements de *contorn* (com els anomena Seco, 1979) de la definició: l'element extrínsec, expressat entre parèntesis a les definicions de *digital*, *tatuar*, *acomplexar* i *esmolar*, adequa les característiques del verb que constitueix el descriptor sintàctic (i en el cas de *digital* també semàntic) de la definició a fi d'identificar-lo amb l'argument d'objecte del verb definit i, a més, explicita les restriccions seleccionals del complement d'objecte del definit. En els casos de *guaitar* i d'*inhabilitar*, en canvi, l'element extrínsec únicament restringeix les possibilitats de l'objecte del descriptor semàntic (*vigilar*, en el primer cas i *inhàbil* en el segon) de la definició a fi de fer-les coincidir amb les del terme definit.

Tot i que la majoria de treballs sobre aquesta qüestió s'han centrat en la categoria verbal (veg. més enrere § 4.2), la necessitat d'expressió dels elements extrínsecs es pot fer extensiva a la majoria de categories predicatives, no com una propietat lingüística, sinó com un mitjà per a satisfer la necessària condició d'equivalència entre definit i definició. Així, les possibilitats de generalització d'aquesta tècnica són limitades, en funció del fet que només es refereixen a cadascuna de les relacions particulars entre les entrades i els contextos de corpus que triem per a definir-les; per això la necessitat d'aquests elements extrínsecs depèn exclusivament de la relació entre les propietats de l'entrada i les propietats del descriptor i de la fórmula de la definició. Si el descriptor és

un hiperònim que permet, com en el cas precedent de *digital*, un major espectre pel que fa a les propietats semàntiques de l'objecte, caldrà restringir les seves propietats. Si la definició, en canvi, és introduïda per un descriptor que presenta especificitats sintàctiques que el diferencien del definit, caldrà adequar les propietats de l'estructura argumental del descriptor a les del terme definit.

A efectes pràctics, la lexicografia convencional ha identificat els elements extrínsecs amb l'argument corresponent a la funció sintàctica d'objecte directe del definit, tot i que es pot posar de manifest la necessitat d'expressió d'altres tipus de restriccions referides a d'altres funcions sintàctiques.

Crec que aquesta breu presentació resumeix els aspectes fonamentals d'allò que s'ha dit sobre els elements extrínsecs en les definicions lexicogràfiques. A fi d'assolir un major nivell de profunditat i de posar les bases d'una certa generalització del tractament d'aquest fenomen, convé analitzar amb més detall les diferents característiques dels arguments.

Pustejovsky (1995a) dona una classificació orientada semànticament dels diferents tipus d'arguments que podem trobar en les propietats d'inserció d'un lexema. L'autor estableix els quatre grups següents:

a) **Arguments propis** (*True arguments*): consisteixen en aquells arguments que cal realitzar sintàcticament perquè es pugui produir l'aparició d'un lexema en qualsevol enunciat. En l'oració "John arrived late", l'element lèxic *John* correspon a un *argument propi*.

b) **Arguments per defecte** (*Default arguments*): consisteixen en arguments que no necessàriament estan expressats sintàcticament, ja que participen a les expressions lògiques de l'estructura de *qualia*. La relació entre el procés *build* i el material utilitzat, com *brick*, fa que en la següent frase el segon es manifesti com a *argument per defecte* del primer: "John built the house out of bricks".

c) **Arguments ocults** (*Shadow arguments*): consisteixen en arguments que estan incorporats semànticament en el lexema, i que poden ser expressats sintàcticament només a través d'especificacions discursives: “Mary buttered her toast with an expensive butter”.

d) **Adjunts propis** (*True adjuncts*): consisteixen en arguments que modifiquen l'expressió lògica d'una frase, però que tan sols formen part de la interpretació situacional, i no estan lligats a la representació semàntica de cap lexema; això inclou els modificadors espacials o temporals: “Mary drove down to New York on Tuesday”.

En general, els treballs sobre lexicografia només s'han ocupat (i encara, com he dit, d'una manera tangencial) de la representació dels arguments del tipus *a*), que en ocasions s'integren dins de la definició com a elements extrínsecs determinats per les necessitats de restringir o ajustar les característiques semosintàctiques del descriptor a les del definit. Els tipus *b*) i *c*) no han estat objecte de cap anàlisi específica. Els modificadors del tipus *d*) cauen, en general, fora de la descripció lexicogràfica.

Convé assenyalar que els arguments exigits per una categoria (els *arguments propis*) no estan associats a cap funció sintàctica específica, és a dir, no s'identifiquen necessàriament amb els complements d'objecte (en el cas dels verbs), sinó que poden ser de tipus molt diferents (Soler, 2001), entre els quals l'expressió dels denominats (per la gramàtica tradicional) complements circumstancials no hi és absent. En el cas de verbs del tipus *incloure X en Y*, el sintagma preposicional *en Y* és exigít per l'estructura de subcategorització del verb, de manera que no es pot realitzar lèxicament aquest verb (o, més precisament, un dels sentits corresponents a aquest verb) sense l'aparició d'aquest argument.

Pel que fa a la resta dels tipus d'arguments assenyalats per Pustejovsky, cal dir que, en general, la lexicografia ha obviat les possibilitats d'aparició dels *arguments ocults*, i no ha assajat cap solució específica que permeti en aquests casos mantenir la condició de substituïbilitat:

amortallar v. tr. Embolcallar amb una mortalla. *Amortallar un difunt*. [DIEC]

ensucrar v. tr. Posar sucre (en alguna cosa), endolcir amb sucre. [DIEC]

Aquestes definicions incorporen, com es pot veure, un argument instrumental que es troba semànticament incorporat en el verb definit. Tot i que aquests arguments es troben integrats intrínsecament en la paràfrasi de definició, es poden produir contextos d'inserció en què, a través de la incorporació de determinades especificacions, aquests es trobin realitzats en enunciats lingüístics, com ara: *van amortallar el difunt amb una mortalla de lli*.

<p>li donaren. L' endemà fou portat á la fossa nú, amortallat ab un llansol y dessobre un hayart que portavan dos religióis de la Ordre de Sant Domingo; disposá que se 'l amortallés ab l' hábit de la Mare de Deu del Carme; va prohibir mólt llarga, com la d' un gegant enterch, y estava amortallada ab un hábit de frare, mólt folgat y que, no sé si maná que se li donás lo cos. Y prenent Joseph lo cos, lo amortallá ab un llansol net, y lo posá en un sepulcre seu nou, cos á Joseph. Y Joseph comprá un llansol, y baixantlo l' amortallá ab lo llansol, y lo col·locá en un sepulcre que estava demaná lo cos de Jesus, y habentlo baixat de la creu lo amortallá ab un llansol, y lo posá en un sepulcre obert en una de Jesus, y ungintlo ab aquellas especies aromáticas, l' amortalláren ab telas conforme á la costum que los jueus tenen fou la més crudel una parálisi que li aturá les cames, amortallant ab elles la ferotge energia de la Marquesa. En vostras estimadas. Dòls es prop d' ellas, al morir lo dia Emortallat ab núvols d' or y rosa, Ohir la suáu y tendra al costat del mort, vegé ab sorpresa qu' aquést no estava amortallat ab lo llensol, sino vestit de negre. Atansá l' llum, ve á la imaginació la idea del cuch de seda que fila y s' amortalla ab sa mateixa fusada, en la tardor, per sortir novella que, a l' emparedar- la, s' hi va emparedar una familia, amortallada am la seva historia..., o és un arbre que fuig de a abandonar aquell lloc. El cotxer i en Just tingueren d' amortallar amb una manta el cadáver del pobre majordom i el i flonges! La neu tenia més pietat que no pas ell! L' amortallaria amb les seves blancors puríssimes! I torná a general, les deixava respectuosament a un cantó del plat, amortallades amb una fulla d'escarola. Tampoc no vaig saber a dipòsit dels cadáveris no reclamats de l' Hospital, amortallats amb miserables bruses dins de caixes de fusta sense maná que fos entregat el cós. I rebut el cós, Josep l' amortallá amb un llençol net, i el posá en son sepulcre nou, que el qui, en plena hivernada patriòtica, volgué ésser amortallat amb la bandera catalana. Ell convertí en una simfonia no comportaren que marxás d' aquest món com un ca. L' amortallaren amb un llençol i cercaren almoines per pagar-li un discussions. Algunes hores després, madó Segrina l' havia amortallada amb sedes de colors i li havia cenyit una corona de gent jove i alegre no podia morir. En veure la meva amiga amortallada amb el vestit de la primera comunió, les mans mas, i a la nit obrí 'ls balcons de bat a bat perquè l' amortallés amb el seu raig la lluna, que ell tant estimava i que fes vivent i esperonat d' estranya gelosia, i fos ton cos amortallat amb el llençol on nia! Car, filla fosca del matí, si en les vetlles tan fredes, l' hivern el frec de les sedes amortalla amb un plor. Record i somnis a penes: en la blavor de de Jesus. Pilát maná que lei donásen, i prentintlo, l' amortaiá amb un llansól nóu i bén net, i el posá dins un morí el mateix any que havia fet la primera comunió. L' amortallaren amb el vestit blanc. ;Quins plors els de la pobra i un vell Caronte, cuirassat i gris, que se l' enduia amortallat amb boira. Ha passat setanta anys colgat de sorra i encetada aquell dia del 1970 era memòria llunyana, temps amortallat amb teranyines de boira, una crònica anònima va de tothom. El gest fou interpretat com un desig de ser amortallada amb aquella roba de religiosa plena de sets defóra, lligats els peus i les mans amb benes, i la cara amortallada amb un mocador. Jesús eis digué; Deslligueu-lo i i Pilat maná que l' hi donessin. Josep el prengué, l' amortallá amb un llençol i el posá en el sepulcre nou que ell s'</p>

Llavors cal adóbar l'amanida de la següent manera: Ensucra- la amb sucre fi tirat per damunt i rega-la amb uns posarás dins un colador gran, sucre esponjat esclafat i ensucrarás es suplillo amb aquest sucre amb un sedás molt fi.

Figura 4.3. Mostra de concordances d'amortallar+amb i d'ensucrar + amb

EL CTILC pot aportar dades en aquest sentit, i la demostració que l'especificació de l'argument ocult apareix en determinades condicions. A la figura 4.3 es poden veure els contextos del CTILC corresponents a les ocurrències d'*amortallar* seguit immediatament de la preposició *amb*, i a les del verb *ensucrar* seguit en posició pròxima de la mateixa preposició.

Els contextos recollits a la figura 4.3 mostren, justament, la propietat més característica dels arguments ocults (més endavant, en aquesta mateixa secció, m'ocupo de les similituds entre els arguments ocults i els arguments per defecte): el fet de poder manifestar-se en certes condicions, en especial quan s'introdueixen especificacions que els modifiquen o els precisen semànticament en un enunciat concret. D'alguna manera, són arguments per als quals es poden donar especificacions semàntiques que requereixen la presència física de l'argument en si. Així, el fet que en l'enunciat calgui especificar característiques associades a l'argument *amb una mortalla* que no formen part de la descripció semàntica bàsica d'*amortallar* fa necessària la manifestació externa de l'argument, normalment a causa de l'especificació del substantiu (*llençol*, etc.) que fa de *mortalla*.

De manera similar als arguments ocults, els *arguments per defecte* també solen ser incorporats com a elements intrínsecs a l'enunciat de definició. En els exemples que vénen a continuació (*engiponar* i *amonestar*), es pot veure com la presència en el text de la definició d'especificacions que representen trets semàntics inherents del definit (i que per tant es poden considerar incorporats per defecte) no impedeix la realització, en els enunciats lingüístics, d'alguns d'aquests trets com a arguments sense cap subespecificació especial. De la definició lexicogràfica dels verbs *engiponar* i *amonestar*, per exemple, es pot deduir que incorporen, respectivament, trets significatius que fan referència al mode en què es produeix l'acció verbal: *de qualsevol manera*, i *severament*.

engiponar v. tr. Arranjar de qualsevol manera, a corre-cuita. [DIEC]

amonestar v. tr. Advertir severament (algú) per tal que s'esmeni. [DIEC]

El CTILC, però, ens mostra com en ocasions la utilització d'aquests verbs comporta l'aparició explícita de components que fan referència a aquests trets, tant en el cas d'*engiponar* com en el d'*amonestar*,⁵⁸ això es pot veure a la figura 4.4.

« es passà pel coll el tortell de les faldilles, s'*engiponà* de qualsevol manera el sac de cotó musc i se n'anà a vestir la nena. » [393; 280, 13]; Caterina Albert Paradís, *Jubil·leu*, 1951
 « Intentà apartar-lo del vi. Endebades Preguntà a Laia i a Esteve. Els *amonestà* amb severitat. Ells negaven... Senyor, que es faci la voluntat teva! » [149; 114,26]; Salvador Espriu, *Laia*, 1932.
 « En Deberga esclafí a riure; però la mamà's posà molt seria y *amonestà* severament a la noya pera que's contingués com debía. » [300; 112, 19]; Narcís Oller, *Pilar Prim*.
 « la Reyna s' esvolotà també, *amonestant* severament als gremis per lo pas que acabavan de donar; los Consellers protestaren també » [1.995; 86,6]; Salvador Sanpere i Miquel, Títol : *Barcelona: son passat, present y porvenir*, 1879.
 « La Dominació que presideix el meu Cercle em va *amonestar* severament; em va condemnar a dos segles i un dia de reclusió i em va suspendre »: [18; 45,12]; Joan Oliver, *Allò que tal vegada s'esdevingué*, 1936
 « les persones de sa familia, com lo senyor oncle, per exemple, l'*amonestarian* severament á l' hora que menys ell pensés. » [3.022; 49,8]; Lluís Bertran Nadal i Canudas, *Benet Roure*, 1887.

Figura 4.4. Casos específics d'*engiponar* i d'*amonestar*.

En alguns casos l'aparició en un enunciat de l'argument per defecte pot anar acompanyada de canvis en el comportament sintàctic del verb. És el cas del verb *emparar* que el DIEC descriu, en una de les seves accepcions, com a simplement transitiu, sense indicar el fet que es pot produir el mode ergatiu del verb amb la consegüent pronominalització:
X empara Y (sota la protecció d'X) → Y s'empara (sota la protecció d'X):

emparar v. tr. Posar (algú) sota la nostra protecció. [DIEC]

Al CTILC es poden trobar amb certa facilitat mostres d'aquest sentit que posen de manifest l'argument per defecte, amb la característica que el verb presenta un patró pronominal:

Tots dos en silenci, no gosaven trencar l'encantament d'Otilia present, **emparant-se sota llur protecció**. (Francesc Trabal, *Vals*; p. 166)

Entre ells [els departaments que integren un gran municipi], el de Cultura té una missió específica a complir que transcendeix a tota la gamma de manifestacions artístiques i de belles arts que **s'emparen sota la seva tutela**. (Gonçal Lloveras, *Doctor A. Pedro Pons*, p. 60)

⁵⁸ En el cas d'aquest verb podem pensar, a més, que la realització d'aquest argument per defecte a través de l'adverbi *severament* té un cert caràcter col·locatiu que podria ser tractat com una *funció lèxica* (veg. més enllà § 3.3) d'intensificació del significat verbal.

Les definicions de Pustejovsky sobre els tipus d'arguments lògics no estan, però, exemptes de certes dificultats de classificació, si les traslladem a la realitat lingüística. Un dels aspectes més problemàtics recau en l'establiment de la diferència entre els arguments per defecte (*default arguments*) i els arguments ocults (*shadow arguments*); en molts casos, aquesta diferència no és del tot evident.⁵⁹

Pustejovsky, a més, assenyala el fet que no són exclusivament les propietats lèxiques dels mots allò que determina la selecció dels diferents tipus d'arguments lògics, i que, a més, els arguments poden, en ocasions, variar el seu tipus (*type*) en funció de les característiques semàntiques del complement que el realitza:

It is not just the lexical properties of a single item which determine the logical status of a phrase as a certain argument type. Compositional operations may create an argument or shadow an argument at a non-lexical projection, by virtue of compositionality in the phrase. In other cases, however, a logical argument is defaulted by virtue of a complement's semantics. For example, in *show a movie* the goal argument is defaulted by the semantics of the complement, and becomes an optional argument (...). Furthermore, default arguments when realized in certain ways, can mimic the behavior of shadow arguments. For example in the sentence

Mary built a wooden house with pine

the default argument has effectively been saturated indirectly by a modifier in the direct object, and the further specification *with pine* is licensed in the same manner as in shadow arguments. [Pustejovsky (1995b: 80)]

Aquestes dificultats de determinació degudes a la mobilitat existent entre un i altre tipus d'arguments també es posen de manifest en els enunciats de definició. El verb *rodolar*

⁵⁹ Pustejovsky (1995a: 65) diu sobre això: “Unlike default arguments, however, which are optionally expressed due to conditions above the level of the sentence (i. e., discourse and contextual factors), shadow arguments are expressible only under specific conditions within the sentence itself; namely, when the expressed arguments stands in a subtyping relation to the shadow argument.”

incorpora en la seva definició arguments que *a priori* podríem considerar arguments per defecte, ja que són especificacions susceptibles d'aparèixer o no en funció de l'enunciat:

rodolar *v. intr.* Caure per un pla inclinat, moure's al llarg d'una superfície, donant voltes sobre si mateix. *Els rocs despresos rodolaven fins al fons de la vall. Fer rodolar un tronc d'arbre. De l'empenta el va fer anar rodolant per terra.* [DIEC]

<p>camí més fàcil, que es ço que fa sempre l' home. Es dolç mb la superfície líquida... És clar, el negre i jo vàrem com la impressió que l' edifici trontollava. L' home i malparla. Mitja humanitat per terra, la raó i la força tornar xoroí durant un parell d' hores. En Gasparet Brau dempeus a bord i evitar que el sacseig de la nau et faci Era en aquest clot on germans i cosins solíem jugar: tornaven a esclatar els meus plans i el meu idol incendi i de destrucció, i els ornaments sagrats havien setmana que, pagues hores abans, aquell mateix mati, ha Espanya lliure rica, treta de les mans dels qui la fan per Europa: tro "spantós del llamp terrible que feu qui passa amb suau volada, cau amb peresa, assolada, i amb l' intel·lecte i amb les venes, Cor flamejant que d' abans, ara que tant comprenia? Enllardar-se, cobejar, hores. La nebotà fugia: -Està apestada! Està apestada! capital de Catalunya, ara, a l' època que estudiem, havia net com el palet allisat i arrodonit pel fregadís en anar Vejam com acabarà això?, em deia jo, a mesura que tot enbarborant els mandrons. Dos o tres pirates varen van haver de presenciar com el senyor Pelai i el bandit les comes, com més d' una vegada m' havia passat, i fent entra pels ulls de bou de la capella i un tro desmanegat blanc de la boira. Les fulles queien planejant i algunes li el vestit que mostrava les senyals d' haver tan aviat s' enfilava per l' aire grisenc amunt, com Serra Cavallera, es deixà anar fins al fons del barranc, la careta posada. D' una revolada em trec el casc, que llit amb una docilitat sobtada. Tots els seus propòsits i s' ennuvola a cada instant. Mitja hora després se sent l' embalum, en efecte, s' esfondrà de genolls: En Joan ni túniques de llustrina, les bacànals, les orgies, i com dir- ho als seus, si donar un escàndol, després d' haver que s' ho mirava no va poder crear apostes. Recordo que Passats uns moments, la mà va obrir- se i el borinot els mals sons trasbalsaren l' animeta d' en Batista. Rodolava per un pendís sense fons; obria els braços i els ulls- el cotxe i estimbar- lo una mica lluny. Se l' imaginava rogenic, amb els ulls ben negres i nets, un animaló per a estora o el raig d' un brollador, un cadell rodó per a</p>	<p>rodolar pel pendís, és treballós enfilat costa amunt... S'abocà- rodolar pel pendent... -Certament un accident així no és fàcil- rodolà pel mosaic blanc i caigué aplomat, per davant de la- rodolant pel llot. Mes, sortosament, capitulada la raó i la- rodolà pel trespol i jo mai no havia tengut una sensació tan- rodolar pel pont. El motorista ha estat despert tota la nit rodolàvem pel pendent, ens aferràvem a les arrels dels pins, rodolava pel llot. M'apartava d'ella, però hi tornava, i rodolat pel carrer, mentre les tombes eren violades a la rodolat pel patibul la testa sobirana de Lluís XVI. Allavores- rodolar pel pendís d'esdevenir una colònia lliurada a les- rodolar pel cadafalç lo cap del rey, fent croxir y tremolar rodola pel trespol. Tot se sent emperesir, dolça son mos ulls- rodoles pel glaciari d'una Europa pusil·lànim i bàrbara: Quin- rodolar per l'amarg pendent de la passió i de l'espera, ara Rodolava per les escales: -Està apestada! Quietud. Defallien rodolat per terra com la mateixa Grècia, que, havent estat, en- rodolant per un corrent d'aigua o per l'embat de les ones. rodolava per la gelera. I creu que no tenia pas gaires- rodolar per terra, ferits de pedrades al cap. Els altres- rodolaven per terra tot copejant-se bàrbarament. Amb tant- rodolar per terra les ametlles i les nous amb sonoritats mai- rodola per la teulada. El pare impàvid. L'àvia se senya. rodolaven per sobre meu. Els pardals venien a bequejar-me les- rodolat per terra. El pare aprofità el moment, per explicar- rodolava per damunt l'herba perfumada i humida. Però ai, la- rodolant per terra embolicada en el cor d'una feixina de palla rodola per terra i em poso la meva. M'atanso a l'home, prenc rodolaven per damunt de la seva voluntat estèril! VIII -Tu,- rodolar per l'empedrat un cotxe que s'acosta. El noi del rodolà per terra amb uns grans xiscles. Se'ls acudí una altra- rodolaven per sota les taules els comensals després dels rodolat per terra a la biblioteca de l'Ateneu barallant-se amb- rodolí per terra. Per evitar-ho, Hermès es veu obligat a- rodolàrem per terra, i que tan aviat veia la lluna damunt meu,- rodolà per sobre de les puntes, convertit en una piloteta Rodolava per un pendís sense fons; obria els braços i els ulls- rodolant per un pendís entre un núvol de pols i un ròssec de- rodolar per faldes i catifes, que hauria estat ben bé una- rodolar per faldes i catifes, i pensava com s'ajeuria ferida i-</p>
--	--

Figura 4.5. Mostra de concordances de *rodolar+per*.

Els contextos del CTILC que recull la figura 4.5 incorporen en llur majoria l'argument per defecte *per un pla*, etc., que forma part de la definició. En altres ocasions, però, l'aparició d'aquests arguments en el discurs és causada per les necessitats d'especificació que són característiques per a l'aparició dels arguments ocults. Aquest és el cas del verb *rosegar*:

rosegar *v. tr.* Emportar-se petites parts (d'alguna cosa dura), esmicar-la a poc a poc, donant-hi amb les dents una sèrie de mossegades. *Rosegar una crosta de pa. Les rates han rosegat tota la roba, tots els papers, que hi havia al calaix. Rosegar un gos un os.* [DIEC]

A partir de la consulta al CTILC es pot comprovar com aquest verb no sol incorporar els arguments per defecte que conté la seva definició (*amb les dents*), però almenys en un cas apareix aquest tipus d'argument com si es tractés d'un argument ocult, és a dir, amb especificacions que matisen la naturalesa de l'argument. El fet que en frases d'aquest tipus es donin especificacions superficials sobre un argument fa que, segons la interpretació de Pustejovsky, aquest argument canviï el seu tipus per adequar-lo al dels arguments ocults.

maligna legió d' insectes, els Minuts innombrables, que **roseguen**, amb dents agudes com espines, el cor dels humans. (Calvet, Agustí; *Hores viatgeres*, 1926, p.72)

A partir dels exemples adduïts per Pustejovsky per als arguments ocults (*Mary buttered her toast with an expensive butter*), cal preguntar-se, des d'un punt de vista estrictament pràctic, fins a quin punt la diferència entre els arguments ocults i els arguments per defecte no es basa en l'existència o no de relació morfològica entre un verb i el seu possible argument:

encanonar v. tr. Fer passar per un canó. || Apuntar amb una arma de foc (a algú, a una peça de caça).

En l'exemple d'*encanonar*, una de les accepcions, la primera, té com a complement un argument que en principi podem considerar ocult (*per un canó*). La segona definició incorpora, en canvi, un argument per defecte, que es pot manifestar en frases com ara *va encanonar el fugitiu amb la pistola*. Però l'argument també pot agafar el *type* d'argument ocult en determinats enunciats: *el va encanonar amb la seva pistola de juguina*.

En qualsevol cas, per a Pustejovsky, la naturalesa dels arguments no es determina (o no es determina només) com una propietat morfològica, sinó lèxica i sintàctica. Per aquest motiu, no podem considerar suficient l'existència de relació morfològica entre un lexema i l'aparició externa dels seus arguments per a determinar la diferenciació entre els arguments ocults i els arguments per defecte.

La manifestació en l'enunciat d'arguments ocults ha de complir també, una certa condició sintàctica: únicament es poden presentar com a *modificadors*, o *circumstanciadors* que introdueixen arguments que no representen funcions sintàctiques centrals de la categoria a què es refereixen. Com ja he mostrat, en general aquests arguments es manifesten introduïts a través de la preposició *amb*, i no ocupen mai la plaça d'un argument extern o d'un argument que desenvolupi la funció d'un complement d'objecte. L'exemple que ve a continuació es refereix justament a aquesta constatació:

concloure (...) Arribar a una solució definitiva (d'alguna cosa) com a conseqüència de les premisses posades. *D'aquestes dues proposicions atorgades es conclou que va ser responsabilitat seva. De tot això, no en podem concloure res.*

En aquest cas, “a una solució definitiva” equival a “a una conclusió”. Però no es pot admetre una frase com *van concloure la conclusió d'aquell problema*, sinó més aviat *van concloure aquell problema amb una conclusió absolutament brillant*.

Ja he mostrat a través dels exemples adduïts que els diccionaris convencionals no assignen cap tractament específic als arguments ocults ni als arguments per defecte, que solen ser incorporats com a elements intrínsecs a la definició lexicogràfica. Des d'un punt de vista tècnic o metodològic, aquesta manca d'explicitud entra en conflicte amb les propietats de substitució de l'enunciat definidor. Com ja he dit anteriorment, si la paràfrasi d'*amortallar* és “embolcallar (algú) amb una mortalla” la substitució esdevé impossible en el cas que l'argument ocult es manifesti en un context determinat, com ara *van amortallar el cadàver amb una mortalla de lli*, o bé *van amortallar el cadàver amb un llençol*. El fet que en el primer cas l'argument es manifesti com a ocult i en el segon com a argument per defecte és un problema interpretatiu que, des del punt de vista lexicogràfic, resulta irrellevant.

L'assaig de donar una solució tècnica a aquestes qüestions requereix un breu examen de les limitacions que presenta aquest sistema en la lexicografia convencional. En general, el tipus de notació amb què els diccionaris determinen l'aparició d'elements extrínsecs (l'ús dels parèntesis o notacions similars) correspon a mecanismes molt poc adequats

per a representar la complexitat lingüística que en moltes ocasions presenten aquests elements. En algun dels diccionaris concebuts a partir de criteris més moderns, l'expressió d'aquests mateixos elements es fa de forma una mica més explícita. El DSLE, per exemple, introdueix l'ús de diferents tipus de parèntesis per al marcatge d'aquests elements:

en todas las definiciones aparece marcado el tipo de sujeto con un paréntesis en ángulo <>. Los complementos se marcan con los corchetes []. El complemento directo de persona lleva la preposición *a* dentro del corchete. Los otros complementos llevan las preposiciones fuera. [DSLE, “Introducción”: XI]

Semblantment, el DEA també estableix un sistema de representació dels elements extrínsecs més elaborat que el dels diccionaris precedents, i hi destaca el marcatge de determinades funcions sintàctiques:

El método que hemos seguido no prescinde de la mención de tales elementos [sujeto y complementos] en la fórmula definidora; pero los presenta como lo que son, *elementos habituales del contorno*, en los enunciados del habla, de la palabra definida, y no como parte sustancial del contenido de esa palabra. Para ello los escribimos entre corchetes, indicando, cuando es preciso, la función que a cada uno corresponde en los enunciados vivos (sujeto, complemento directo, complemento con la preposición *en*, etc.). [DEA, “Guía del lector”: XXII]

L'estreta relació que mantenen els elements extrínsecs amb l'estructura argumental del definit no ha estat representada fins al moment, però, en cap dels diccionaris convencionals de forma totalment explícita. Aquesta observació és especialment vàlida per a la lexicografia catalana. Aquest tractament parcial del fenomen implica un cert grau d'ambigüitat en l'expressió dels elements extrínsecs, i limita les possibilitats d'un tractament rigorós de l'expressió del significat, ja que exclou tàcitament (per l'ambigüitat intrínseca del sistema) la possibilitat d'incorporar el nombre suficient i necessari d'elements extrínsecs en el cas de verbs que presenten més d'un argument

intern. En funció d'aquestes consideracions, podem considerar inexplícites i, fins i tot, inadequades des del punt de vista de la seva substituïbilitat, definicions com:

incloure (...) v. tr. Posar (una cosa) dins una altra. [DIEC]

En el cas precedent, el complement de la paràfrasi de definició *dins una altra* forma part intrínseca de la definició. Si el considerem des del punt de vista de la substituïbilitat, però, cal tenir en compte que es tracta d'un element que justifica la seva presència en el fet que és exigít pel verb definit. Es produeix, així, una falta de correspondència entre el patró sintàctic del mot definit i l'estructura de la definició. Bastarà adduir uns quants exemples del CTILC, que es poden veure a la figura 4.6, per comprovar que és així.

sobre l' ús de les llengües vernacles en l' ensenyament	inclou dintre dels "factors socio-culturals" que de vegades s'
el pagament, un mèrit que, molt evidentment, l' escriptor	inclou dintre de la seva estricta visió del món, essencialment
de coordinació, d' unificació fisiològica, que Kraus	inclou dins la denominació general de "sistema vegetatiu", la
haver inclòs l' iglesia de Sant Pere; però en Lampérez l'	inclou dintre'l tipo general, creyent equivocadament que les
llibre The autonomic nervous system (1921). Langley	inclou dintre del sistema autònom el "simpàtic" i el vague i
fulles i de les tiges de tabac. Si hé aquest producte s'	inclou dins el grup d'insecticides de contacte, s'ha comprovat
el premi d' assaig religiós Francesc Eiximenis que s'	inclou dins els premis Ciutat de Girona, amb un projecte titulat
amb una extraordinària habilitat. Aquesta exposició s'	inclou dins el programa "Cultures indígenes d'Amèrica" que ha
crystal·lina que té l' estructura d' una lactona. S'	inclou dins del grup dels colorants derivats de les ftaleïnes.
groc. S' usa en solució aquosa al 0,1% (p- v). S'	inclou dins del grup de colorants azo: @ En medi àcid té color
havien escrit Nancy Mckeith i altres dones. El llibre s'	inclou dintre la col·lecció "Manuales de Salud", col·lecció
productes que, d' altra banda, moltes classificacions	inclouen dins el grup dels industrials. Respecte als sectors
al concepte original de Freud. Molts d' altres, però, les	inclouen dins la transferència. Sullivan ha intentat d'evitar la
incloure dintre l' esquizofrènia. Willmanns i Mayer-Gross	inclouen dintre l'esquizofrènia les formes depressives de les
del canal. ¡I com és veritat que tots els progressos	inclouen dins seu un món de contrarietats! Qui no havia de
Asplenion glandulosi de terres baixes, i totes dues s'	inclouen dins l'ordre de l'Asplenietalia glandulosi. Això vol
assalariats (per compte d' altri); així també s'	inclouen dins aquest camp les pensions de viduïtat originades
per aprovació de programes d' actuació urbanística. No s'	inclouen dins del concepte de béns urbans els que abans es
en el coneixement científic. Una gran part d' aquests s'	inclouen dintre el que es denomina enginyeria genètica. De fet
n' hi ha algunes que en els darrers temps es tendeixen a	incloure dintre l'esquizofrènia. Willmanns i Mayer-Gross
molt sensible per la banda oest de la vila, tendint a	incloure dintre d'ella, no ja la plaça del segle XIII, adjacent
del seu geni propi, Poincaré no s' és pas acontentat amb	incloure dins la seva obra proplament tècnica, el conjunt de
per superar les dificultats que anirem trobant. Cal	incloure dintre d'aquestes línies l'homenatge sincer que el club

Figura 4.6. Mostra de concordances d' *incloure+dins/dintre*.

La inexplicitud de la lexicografia convencional respecte a aquesta qüestió particular es posa de manifest, també, en verbs que el diccionari considera intransitius, però que incorporen complements d'objecte introduïts per preposició. Aquests complements no es recullen en la part extrínseca de la definició, sinó que solen il·lustrar-se a través dels exemples lexicogràfics:

intervenir (...) v. intr. Prendre part; ficar-se en un afer per arranjar-lo, modificar-lo, etc. *Intervenir en una negociació. Intervenir en un procés. El guàrdia intervingué en la baralla.* [DIEC]

4 – Definició lexicogràfica i equivalència

EL CTILC confirma novament el fet que en molts casos la part corresponent al constituent “*en un afer*” de la definició lexicogràfica apareix realitzat en el verb *intervenir*, com es pot veure a la figura 4.7.

tant, sotmeses al règim comú o general. Quan les Agències intervinguin en d'altres operacions de transport que no fassin la mateixa. De manera que al noy que li agradi intervenir en las feynas de construcció pot ferse estucador, siga evident que la fauna, la vegetació i l' agricultura intervenen en la diferenciació comarcal, resten, malgrat tot, a ambients especials, o s' hi han adaptat. En indrets així intervenen en l'ecologia general de la zona. Hi ha ocells amb que hi havia o bé les que pogués augmentar. L' Ajuntament intervenia en la direcció de la Societat nomenant tres regidors si pot estendre 's la virtut educadora del Dret, al intervenir en tantes relacions qu' afecten intimament la entendre lo nostre meritissim company Sr. Maspons al intervir en aquest debat, es clar que perdriam llastimosament sotmès a una humitat constant. La vegetació. Aleshores intervè en el paisatge un altre element, conseqüència del clima industrials catalans no s' hi presentaren, bé que alguns intervinguessin en pla d'executors de pavellons o stands que els els diferents factors que d' una manera o altra intervenen en la regulació del tonus tensional, podem establir banda els elements modificadors que en una forma o altre intervenen en la regulació de la pressió arterial, és evident ni 'l bras popular tolerava que 'ls uns ni 'ls altres intervinguessin en sos assumptos. Fundamentalment era una que actuï en nom propi o per compte d' altri intervingui en una prestació de serveis, s'entendrà que ha rebut Perque vostè es l' amo de la casa y no està bè que l' amo intervingui en certs detallets... perque vostè si bè es fill de de sexualitat. Però no és necessari que l' amor intervingui en les relacions sexuals. No són termes idèntics ni per exemple a Brutus va matar Cèsar. Aquesta anàlisi intervè en la teoria del raonament exposada en una secció

Figura 4.7. Mostra de concordances d' *intervenir+en*.

Aquesta inadequació en l'enunciat de definició es pot resoldre a partir d'un sistema de notació que representi clarament la correspondència entre cada element extrínsec i l'estructura argumental del definit. Aquest sistema pot tenir diverses formulacions (Picoche, 1998) i no n'examinaré cap de concreta, però correspon a grans trets al sistema amb què els DEC (veg. més enrere § 3.3) introdueixen el patró sintàctic en les accepcions dels seus articles. Aquest sistema pren com a element de descripció no pas el mot considerat aïlladament, sinó tot el seu patró d'inserció lèxica. Així, una definició per a *incloure* no es dona sinó a través d'algun dels seus patrons d'inserció:

X inclou Y en Z → X posa Y dins Z

Aquest sistema es caracteritza per dues propietats principals:

1. Expressa el patró d'inserció lèxica de cada argument del definit, i caracteritza cada argument amb una variable.
2. Permet que la definició repregui les variables que representen els arguments del definit i les integri en la seva estructura de constituents extrínsecs.

A partir d'una formulació com la que acabem d'examinar podem, també, subcategoritzar sintàcticament el definit si aprofitem les variables X, Y, Z, i les substituïm per la categoria sintàctica que les representa. Així, la definició d'aquest sentit del verb *incloure*, encapçalada per un element que conté la descripció del patró sintàctic, té la forma següent:

$[N_1 \mathbf{V} N_2 \text{ en } N_3] \rightarrow \text{Posar [una cosa]}_2 \text{ dins d' [una altra]}_3$

Com es pot veure en l'exemple precedent, l'expressió dels elements extrínsecs s'associa a l'element del patró sintàctic de l'accepció a què es refereix, de manera que cada element extrínsec va coindexat amb el seu corresponent del seu patró sintàctic. Aquest tipus d'expressió de l'extrinsicitat en la definició lexicogràfica compleix amb un grau d'adequació molt més satisfactori el principi de substituïbilitat de l'enunciat de definició i, alhora, és susceptible de permetre tractaments específics per als arguments ocults i per defecte.⁶⁰

El tractament d'aquests arguments de manera indistingida de la resta no resol, però, el fet que aquests arguments poden realitzar-se o no en les expressions en què intervé el definit, per la qual cosa la relació d'equivalència entre definit i definició pot no verificar-se en determinats contextos.

Una possible solució a aquesta qüestió passa per la representació notacional dels diferents tipus d'extrinsicitat, en funció del seu diferent comportament en relació al test de la substitució: d'una banda, un primer tipus que representa els arguments que es trobaran sempre realitzats en qualsevol enunciat lingüístic en què intervingui el definit i, d'una altra banda, els arguments que poden aparèixer facultativament en el mateix enunciat. Aquest segon tipus d'elements extrínsecs pot ser, al seu torn, dividit en dos subtipus: els que condicionen les seves possibilitats d'aparició a l'expressió de

⁶⁰ Un sistema molt semblant d'expressió de l'extrinsicitat s'aplica en la redacció del DDLC.

característiques semàntiques específiques (cas dels arguments ocults) i els que poden aparèixer sense necessitat de cap mena d'especificació (cas dels arguments per defecte).

L'explicitació d'aquestes característiques, que de fet corresponen en sentit estricte a condicions sobre la naturalesa dels arguments i no pas de la definició, s'ha de projectar sobre el patró sintàctic. La presència de parèntesis en l'expressió dels elements afectats indica aquí el seu caràcter opcional, que en uns casos se subindexa amb la nota *espec* (arguments ocults) i en una altres amb la nota *int* (per defecte). Així, tindriem:

amortallar: [N₁ V N₂ (amb N₃)_{espec}] → [algú]₁ embolcallar [algú]₂ [amb una mortalla]₃

emparar(-se): [N₁ Vpron (sota N₂)_{int}] → [algú]₁ posar-se [sota la protecció d'algú altre]₂

L'adopció de sistemes més rigorosos d'expressió de l'extrinsicitat dotats d'un major grau d'adequació al test de la substituïbilitat permet, doncs, la integració de mecanismes per a una formulació més adequada de les definicions lexicogràfiques. Un element fonamental en l'expressió d'aquests mecanismes és, com he mostrat en aquesta mateixa secció, la representació no ambigua de la correspondència existent entre els elements extrínsecs i el patró sintàctic del definit.

4.4 Equivalència actancial

La necessitat d'equivalència definit-definició i la correcta explicitació del significat lèxic implica en moltes ocasions que la definició ha d'expressar quines són les relacions que es produeixen entre els components que representen els arguments del definit. Són especialment coneguts els casos dels substantius de tipus relacional, com els termes de parentiu, que en són un cas emblemàtic. La correcta definició d'un d'aquests termes implica l'explicitació de les *relacions* referencials que es produeixen entre els

arguments. És a dir, només es pot donar la informació semàntica suficient i necessària per a l'explicitació d'una estructura d'inserció com *X és cosí d'Y* si assenyalem el fet que entre X i Y ha d'existir un determinat tipus de relació, que en aquest cas és que un dels progenitors d'Y és germà (o germana) d'un dels progenitors d'X. Una definició lexicogràfica del mot *cosí* ha d'integrar aquesta informació i es pot considerar incompleta si no l'explicita amb claredat.

La lexicografia convencional es mostra vacil·lant en aquest aspecte. El DIEC, per exemple, no mostra cap estratègia explícita ni uniforme per al tractament d'aquests casos, tot i que en algunes definicions (*sogra, cònjuge*) s'introdueixen determinades informacions en el sentit de precisar les relacions.

cosí *m.* Fill d'un oncle o d'una tia d'algú. [DIEC]

oncle *m.* Germà del pare o de la mare. [DIEC]

sogra *f.* Mare del marit o la muller (respecte del cònjuge). [DIEC]

nora *f.* Muller del fill. [DIEC]

gendre *m.* Marit de la filla.

marit *m.* Home casat respecte a sa muller. [DIEC]

espòs esposa *m. i f.* Cònjuge. [DIEC]

cònjuge *m. i f.* Persona en relació amb el seu marit o la seva muller. [DIEC]

tia *f.* Germana del pare o de la mare. [DIEC]

pare *m.* Home o mascle que ha engendrat un o més fills. [DIEC]

mare *f.* Dona o femella que ha tingut un o més fills, esp. respecte a aquests. [DIEC]

Cap d'aquests mots no pot ser definit sense tenir en compte la *relació* existent entre els seus arguments en tant que persones que participen en la *situació* determinada per cadascun dels substantius, de manera que la comprensió d'una formulació del tipus *X és cosí d'Y* exigeix necessàriament que es produeixin determinades circumstàncies que van més enllà del que habitualment es considera l'especificació semàntica d'X i d'Y, és a dir [+ HUMÀ].

La falta d'un mètode descriptivament adequat per al tractament lexicogràfic d'aquests mots s'evidencia en el fet que les definicions no passen, en general, la prova de la substitució, ni segueixen cap model de definició sistemàtic. La inadequació de les

definicions del DIEC es produeix per inexplicitud de l'element *d'algú*, o pel fet que aquest element desapareix. La inadequació es deu, doncs, al fet que la descripció semàntica del mot no integra l'expressió de les relacions entre els constituents de la definició.

En aquesta mena de mots es posa especialment de manifest l'existència de lligams estrets entre tres aspectes de la definició lexicogràfica:

- La necessitat de distingir entre elements intrínsecs i extrínsecs.
- La necessitat de reflectir el patró de comportament sintàctic del mot a què es refereix.
- La necessitat d'explicitar certes coreferencialitats o el valor de certs *actants* semàntics: $X \text{ és cosí de } Y = X \text{ és fill de l'oncle o de la tia de } Y$.

Sovint, en la lingüística teòrica, les nocions d'*actant* i d'*argument* s'han tingut com a sinònimes o, almenys, molt pròximes pel que fa a la seva naturalesa. En general, l'ús de l'una o de l'altra expressió en la bibliografia es deu a les preferències de cada corrent de la lingüística teòrica. Mentre que els corrents generativistes i les gramàtiques d'unificació prefereixen la denominació d'*argument* per referir-se als elements governats sintàcticament per un lexema, altres corrents (els pròxims a la semàntica conceptual, com la MTT) solen preferir la denominació d'*actant*. En la secció anterior (§ 4.3) he reservat el terme *argument* per als elements que formen part de l'estructura de subcategorització sintàctica. En aquesta secció utilitzo una noció d'*actant* que es refereix a un concepte diferenciable del de l'*argument* sintàctic, a partir d'una distinció introduïda per Mel'čuk (1984, 1993).

Mel'čuk distingeix entre *actant semàntic* i *actant sintàctic*. Els actants semàntics corresponen a les variables que apareixen com a arguments d'un lexema (X i Y amb els participants de la situació denotada per una accepció d'un mot (per exemple *X perseguir Y*, en el cas de *perseguir*). Els actants sintàctics, per altra banda, es refereixen a les actualitzacions d'aquests actants semàntics en els enunciats lingüístics, és a dir els

sintagmes nominals que expressen els actants semàntics del definit. El conjunt d'actants configura la *valència* d'un lexema.

L'aproximació teòrica de la MTT a la sintaxi, doncs, es basa en els següents components (Nakhimovsky, 1990):

- una situació i els seus participants (actants)
- un mot i els seus actants semàntics que formen la *valència semàntica*
- un mot i els seus actants sintàctics que formen la *valència sintàctica*

Per a Mel'čuk és essencial la distinció entre la valència sintàctica i la valència semàntica. La primera està completament determinada pel significat del mot, això és, per la seva definició; les variables de la definició són els actants semàntics. La segona està determinada per les dependències “fortes” (subjecte i complements) que es combinen sintagmàticament amb el mot.

Prenem com a exemple el verb *condemnar*; aquest verb sol presentar tres arguments —o actants sintàctics— quan apareix en frases del tipus *A condemna B a D* (per exemple: *el jutge condemnà l'acusat a tres anys de presó*); des del punt de vista semàntic, però, l'anàlisi que en fa MTT és diferent: només es pot parlar de condemna si la persona condemnada ha comès una acció punible, que fins i tot pot aparèixer realitzada sintàcticament: *El jutge condemnà l'acusat per agressió a tres anys de presó*; aquest element, que s'inclouria en una descripció semàntica de *condemnar* del tipus:

A causa a B un dany D amb el propòsit de fer que B (o altres persones) no tornin a fer C

La figura de l'actant semàntic apareix integrat en la metallengua de definició dels DEC com a vehiculador de la informació de subcategorització:

X teaches Y to Z = X, having knowledge of, or skills in, Y, causes Z intentionally and methodically to learn Y.⁶¹

La noció d'actant a què em refereixo aquí es correspon, doncs, amb els *actants semàntics* de la MTT. Es tracta dels elements de l'estructura argumental considerats en relació a llurs referents semàntics i al paper que desenvolupen en la *situació* denotada pel nucli sintàctic. Així, utilitzo aquest terme per posar de manifest la necessitat d'establir coreferències o distincions entre els elements de la definició i el patró sintàctic del definit, i també per la necessitat de desambiguar referencialment els elements que formen part de l'enunciat de definició.

Aquesta necessitat d'establiment de relacions de coreferència entre els elements del definit i de la definició es veu també complementada per una altra necessitat: establir dins de l'enunciat els mecanismes de desambiguació referencial suficients que permetin una interpretació correcta de l'enunciat de definició, que determinin les equivalències i les diferències entre els elements, a fi de reflectir les relacions que mantenen els elements de l'estructura semàntica interna del mot.

La lexicografia convencional presenta alguns indicis de consciència d'aquesta problemàtica, que es pot veure reflectida en la utilització que fan alguns diccionaris dels pronoms personals. L'expressió lingüística de la coreferencialitat (i de la resolució de l'anàfora) recau fonamentalment en aquesta classe de mots. Com que el discurs de definició lexicogràfica és bàsicament la llengua natural, es poden rastrejar en els hàbits d'utilització dels pronoms i de les diferents *persones* gramaticals, quines són les qüestions fonamentals que presenta aquest problema.

⁶¹ Segons Mel'čuk, el *definendum* és *X teaches Y to Z*, de manera que l'objecte de descripció no és el mot aïllat (com en la lexicografia convencional) sinó una estructura que incorpora certes propietats d'inserció lèxica del mot. Això implica un major grau de rigor descriptiu pel que fa a les coreferències que cal establir entre el definit i la paràfrasi definidora per a la interpretació de la definició.

La definició lexicogràfica constitueix un tipus de discurs bàsicament descriptiu que troba la seva fórmula *natural* d'expressió en la tercera persona gramatical. Quan un diccionari diu que un mot *significa* una determinada expressió tracta, en general, amb entitats que no impliquen el *jo* i el *tu* gramaticals. Un diccionari descriu el tipus de relació existent entre mots i significats en termes de *tal mot vol dir tal expressió*, o bé *dir això equival a dir això altre*.

Aquesta afirmació de tipus general explica la causa que el discurs lexicogràfic sempre s'ha tingut per un tipus d'enunciació notablement diferenciat de la utilització del llenguatge en altres funcions. La tercera persona gramatical hi té una predominança gairebé exclusiva.

La publicació del diccionari COBUILD va suposar la incorporació a la lexicografia pràctica moderna d'un nou sistema de construcció de definicions basat en la idea d'acostar-se al màxim possible a la llengua ordinària (Hanks, 1987), integrant dins de la definició les característiques que solen consignar-se com a extrínseques (informacions que habitualment van entre parèntesis o que formen part d'altres elements estructurals del diccionari), i prescindint de les expressions formulàries típiques de les definicions lexicogràfiques (com ara *relatiu o pertanyent a*, *acció i efecte de*, etc.).

El COBUILD formula les definicions sobre un context que integra el mot d'entrada, i que, alhora que el descriu semànticament, informa sobre les seves característiques gramaticals bàsiques; així, per exemple, els noms comptables són introduïts sempre per l'article indeterminat (*A brick of ice cream is...*), els no comptables no incorporen cap article (*Calligraphy is...*), mentre que els que duen sempre article determinat s'introdueixen amb el mateix article (*The future is...*).

Els canvis introduïts en la presentació de les definicions del COBUILD modifiquen d'una manera força evident l'aspecte extern del discurs lexicogràfic; però tal vegada la seva aportació més remarcable sigui en la definició dels verbs. Tant les restriccions selectives de complement com de subjecte s'integren de manera natural en la definició. Així, el COBUILD introdueix les definicions de verbs de subjecte personal amb

fórmules com ara *If you...*, o *When you...* L'ús de la segona persona en aquests casos, fins i tot en els verbs que poden realitzar-se amb un subjecte més general (*menjar*, *caminar*, *dormir*, etc.), és general al llarg de tot el diccionari.⁶²

Malgrat que aquest és un sistema aparentment poc formalitzat, permet de fer front a les necessitats de referenciació interna de les definicions. A part de determinar la categorització general dels arguments del definit, aquest sistema permet la presentació no ambigua de cadascun dels actants que intervenen en la situació. La introducció dels pronoms personals de les tres persones gramaticals és l'estratègia bàsica utilitzada per aquest diccionari, que usa la segona persona (amb preferència sobre la tercera) per a totes les referències a elements de tipus [+HUMÀ] en les definicions de qualsevol categoria gramatical:

Your **cousin** is the child of your uncle or aunt. [COBUILD]

La segona persona és especialment útil també per a la distinció actancial en aquelles definicions que incorporen alhora una tercera persona de tipus [+HUMÀ]:

Shame is an uncomfortable feeling of guilt and failure that you have because you have not behaved in an acceptable way or because someone close to you has not behaved in an acceptable way. [COBUILD]

A **friend** is someone who you know well and like, but who is not related to you. [COBUILD]

Aquest sistema, malgrat el seu caràcter fàcilment interpretable i intuïtiu, és en molts casos sobregenerador pel que fa a l'ús motivat de la segona persona gramatical. En els exemples que segueixen, no hi ha cap necessitat objectiva de recórrer a aquest ús:

A **hair** is one of the long, fine, thread-like things that grows in large numbers on the top part of your head and on other parts of your body. [COBUILD]

⁶² Les úniques excepcions són els casos que denoten accions punibles o immorals; en aquests casos hom introdueix directament l'infinitiu: *to strangle someone means...* Ocasionalment, també s'usa la tercera persona: *If someone swindles a person or organization...*

If you are **thirsty**, you feel that you want or need to drink something.⁶³
[COBUILD]

Això contrasta fortament amb el que, fins a la publicació d'aquest diccionari, havia estat la norma general d'utilització dels pronoms personals en el discurs lexicogràfic. Els motius d'aquesta tria cal buscar-los, en part, en la major familiaritat que suposa l'ús de la segona persona, i en el reconeixement per part del lexicògraf de l'antropocentrisme que domina els prototips lingüístics. A jutjar per diccionaris de tipus similar (com el CIDE) que s'han publicat posteriorment, aquesta innovació ha tingut un cert èxit, ja que també han incorporat la utilització dels pronoms de segona persona.

En el CIDE hi ha una certa tendència a recórrer a la segona persona com a marca de subjectivitat, sobretot per als mots que es relacionen amb estats mentals o físics de l'individu. Aquest ús no s'aplica sistemàticament a aquest conjunt de mots com a classe, però és certament freqüent:

mood (...) *n* [U] the way you feel at a particular time. [CIDE]

sick (...) *adj* feeling ill as if you are going to vomit. [CIDE]

La segona persona també s'usa amb freqüència en la definició de verbs que signifiquen actes de parla:

forgive (...) to stop blaming or being angry with (someone) for something they have done, or to ask someone not to be angry with you. [CIDE]

ask (...) to request or invite (someone) to go somewhere with you or to come to your home. [CIDE]

⁶³ Compareu amb les definicions corresponents del LDOCE:

hair (...) a fine threadlike growth from the skin of a person or animal.

cousin (...) the child of one's uncle or aunt.

thirsty (...) feeling thirst.

Finalment, es poden trobar molts casos de segones persones com a substituïts de subjectes de tipus impersonal en determinades clàusules de la definició, sense que això estigui lligat a la naturalesa gramatical del definit:

sight (...) *n* [C] part of a gun or other device through which you look to help you aim at something. [CIDE]

En la lexicografia catalana la funció dels pronoms personals és més limitada. L'ús dels pronoms personals forts és pràcticament inexistent, ja sigui per vehicular informació semosintàctica sobre algun dels arguments o per a l'establiment de lligams interns a la definició. Però si examinem amb un cert deteniment algunes definicions lexicogràfiques del DGLC es poden trobar certs casos d'utilització de les diferents persones gramaticals que tenen a veure amb les necessitats d'establiment de coreferències internes a la definició. Aquests casos corresponen a quatre tipus bàsics de motivacions:⁶⁴

⁶⁴ Deixo de banda alguns casos d'inclusió supèrflua d'aquest recurs. En determinats casos la utilització de fórmules com ara *ens indica*, *ens mostra* no sembla encara gaire justificada metodològicament (el subratllat és meu):

areòmetre *m.* Instrument per a determinar pesos específics, esp. de líquids, consistent en un flotador proveït d'una escala graduada que, pel que s'enfonsa en un líquid, *ens indica*, directament o fent ús d'una taula, el pes específic d'aquest líquid; o en un flotador que, per mitjà de pesos addicionals, hom fa que s'enfonsi sempre fins al mateix nivell, amb el qual instrument, podent-se determinar el que pesa un cos sòlid i el que, immersit en l'aigua, perd del seu pes, *ens és donat* de calcular el seu pes específic. [DGLC]

clau *f.* (..) | En mús., signe col·locat al començament d'un fragment que *ens indica* quina ratlla del pentagrama correspon a una nota determinada, i per consegüent les ratlles i espais en què cal escriure les altres notes de l'escala musical. [DGLC]

Un cas específic clarament evitable el constitueixen les definicions de noms referits a animals i plantes, on és fàcil de trobar expressions de possessius de primera persona en relació al fet que l'animal o planta tingui el seu hàbitat dins del territori corresponent al domini lingüístic català; aquesta informació és clarament enciclopèdica:

cranquet *m.* Nom amb què es designen algunes espècies de crancs *del nostre* litoral, esp. els que viuen com a comensals en la túnica de certs mol·luscs (gèn. *Pinnotheres*). [DGLC]

1. En determinats casos la primera persona s'utilitza per a la descripció de mots de significats inherentment subjectius, lligats en general a estats d'ànim o a sentiments o opinions individuals que no són generalitzables pel fet que no són propietats objectives de les coses. En general, es tracta de substantius (deverbals o no) que subcategoritzen complements (*el desdeny d'*X* per *Y*, etc.*), un dels quals (*X*) fa de subjecte i l'altre (*Y*) d'objecte. Per naturalesa, *X* és sempre de tipus humà, mentre les possibilitats de *Y* són més àmplies.

La semàntica de les definicions d'aquest tipus de mots en el DGLC, mostra sovint la relació de causalitat que existent entre l'element que representa *Y* i el descriptor de la definició (que representa el mot d'entrada):

desdeny *m.* Sentiment que ens fa desdenyar algú o alguna cosa.⁶⁵ [DGLC]

desgrat *m.* Desplaer que ens causa una cosa que no és del nostre grat.⁶⁶ [DGLC]

anagall *m.* Petita planta herbàcia anual, l'*Anagallis arvensis*, de la família de les primulàcies, freqüent en els nostres horts i conreus; té dues varietats, l'una de flors rogenques i l'altra de flors intensament blaves. [DGLC]

En alguns altres casos l'ús de la primera persona (defensable metodològicament o no) és clarament restrictiu, ja que s'aplica a complements que no només poden ser de tipus [+HUMÀ], sinó més aviat [+ANIMAT]:

cel *m.* Part de l'espai que ens apareix com una volta sobre nostre circumscrita per l'horitzó. [DGLC]

llum *f.* Forma d'energia que per la seva acció sobre els ulls ens fa veure els cossos. [DGLC]

patir *v. tr. i intr.* Tr. Rebre l'acció (d'alguna cosa que ens causa un dany, dolor, ofensa, etc.). [DGLC]

⁶⁵ El DIEC manté la mateixa definició, mentre que el DLC elimina les referències a la primera persona en l'article *menyspreu*, que és el mot que s'utilitza en la definició sinonímica de *desdeny*:

desdeny *m* Menyspreu. [DLC]

menyspreu *m* Sentiment pel qual hom considera com a indigna d'estima, d'ésser tinguda en compte, una persona o una cosa. [DLC]

⁶⁶ El DIEC manté la mateixa definició, mentre que el DLC elimina les referències a la primera persona:

desgrat *m* Desplaer que causa a algú allò que no és del seu grat. [DLC]

despit *m.* Irritació que ens causa el desdeny que algú ens mostra, la preferència que mostra a un altre.⁶⁷ [DGLC]

fàstic *m.* Sentiment desagradable que ens causa una cosa que ens repugna.⁶⁸ [DGLC]

gelosia *f.* (...) enveja que ens causa el qui gaudeix d'alguna cosa que voldríem per a nosaltres. [DGLC]

injúria *f.* Dany inevitable que ens fa alguna cosa. [DGLC]

maliciada *f.* Forta malícia que ens fa una cosa.⁶⁹ [DGLC]

remordiment *m.* Pena que ens causa de sentir-nos culpables d'una mala acció.⁷⁰ [DGLC]

2. La subjectivitat que introdueix la primera persona també es manté en algunes interjeccions, etc. (definides amb metallengua de signe) relacionades amb manifestacions de valors primàries i no generalitzables:⁷¹

⁶⁷ El DIEC manté la mateixa definició, mentre que el DLC elimina les referències a la primera persona; en aquest cas, però, l'eliminació representa una certa pèrdua (o almenys un canvi) d'informació, de manera que el *despit* ja no queda caracteritzat com la "irritació que causa a la persona X el desdeny que la persona Y mostra per X"; a més, el manteniment de la part de la definició "la preferència que mostra a algú altre" perd funcionalitat si no s'ha especificat el complement directe de *mostrar*:

despit *m* Irritació que causa el desdeny que algú mostra, la preferència que mostra a algú altre. [DLC]

⁶⁸ El DIEC manté la mateixa definició, mentre que el DLC elimina les referències a qualsevol persona gramatical, amb la qual cosa la definició de *fàstic* esdevé una definició equiparable a la d'un terme *objectiu*:

fàstic *m* Sentiment desagradable que causa una cosa que repugna. [DLC]

⁶⁹ El DIEC manté la mateixa definició, mentre que el DLC elimina les referències a la primera persona i, alhora, les referències a la causa externa, amb la consegüent pèrdua d'especificitat de la definició, que passa a ser:

maliciada *m* Malícia intensa. [DLC]

⁷⁰ El DIEC manté la mateixa definició, mentre que el DLC elimina les referències a la primera persona, i incorpora l'expressió 'd'una acció tinguda per dolenta' com a element subjectivitzador, sense assenyalar el fet que cal que l'acció sigui *tinguda per dolenta* pel mateix que l'ha realitzada:

remordiment *m* Pena, neguit, que hom experimenta en sentir-se culpable d'una acció tinguda per dolenta. [DLC]

ecs, interjecció que expressa el fàstic, la repugnància, que ens causa algú o alguna cosa.⁷² [DGLC]

uix, interj. que expressa el fàstic, la repugnància, que ens fa algú o alguna cosa.⁷³ [DGLC]

3. En les definicions de mots que fan necessària l'explicitació d'un punt de vista, una situació, o una naturalesa que es pot identificar amb el de l'emissor, és freqüent la utilització de la primera persona:

endarrera *adv.* En direcció a la banda que és darrera nostre; més o menys lluny en aquesta direcció.⁷⁴ [DGLC]

⁷¹ També hi ha ús no sistemàtic de primeres persones en casos d'interjeccions que corresponen a expressions que manifesten reaccions a experiències que suposadament són únicament individuals:

ui, *interj.* que proferim quan ens fem mal, quan volem manifestar que una cosa ens sorprèn granment. [DGLC]

ui *interj.* Expressió usada per a manifestar que ens hem fet mal, que una cosa ens sorprèn granment. [DIEC]

La poca sistematicitat d'aquest ús es pot il·lustrar a partir dels casos següents:

uf, *interj.* que denota l'opressió, l'aclapament; també, el fàstic, el menyspreu. [DGLC]

uf *interj.* Expressió usada per a manifestar opressió, aclapament; també fàstic, menyspreu. [DIEC]

⁷² El DIEC manté la primera persona; el DLC, no:

ecs *interj.* Expressió usada per a manifestar el fàstic, la repugnància, que ens causa algú o alguna cosa. [DIEC]

ecs! *interj.* Partícula que expressa fàstic, repugnància. [DLC]

⁷³ El DIEC manté la primera persona; el DLC, no:

uix² *interj.* Expressió usada per a manifestar el fàstic, la repugnància, que fa algú o alguna cosa. [DIEC]

uix! *interj.* Interjecció que expressa fàstic, la repugnància, que fa a algú una persona o una cosa. [DLC]

⁷⁴ Tant el DIEC com el DLC eliminen l'ús de la primera persona:

endarrere *adv.* En direcció a la banda que és darrere algú; més o menys lluny en aquesta direcció. [DIEC] **endarrera** *adv.* En direcció a la banda que és darrera d'algú. (...) Més o menys lluny en aquesta direcció. [DLC]

estranger -a *adj.* Que és d'un altre país; que és altre que el nostre país.⁷⁵ [DGLC]

4. En altres casos, la presència de la primera persona es justifica per la necessitat de distingir diferents components de la definició que, reduïts conjuntament a la tercera persona, produirien una definició difícil d'entendre o ambigua:

escurabutxaques *m. i f.* El qui o allò que ens fa despendre tots els diners de què disposem.⁷⁶ [DGLC]

do *m.* Acció de donar allò que és nostre a algú; la cosa així acordada.⁷⁷ [DGLC]

dret *m.* Facultat d'exigir allò que ens és degut, de fer allò que la llei no defèn, de tenir, exigir, usar, etc., allò que la llei o l'autoritat estableix a favor nostre o ens és permès per qui pot.⁷⁸ [DGLC]

Aquest ús de la primera persona és el que amb més facilitat pot trobar-se en les definicions verbals:

⁷⁵ En aquest mot es troba un dels pocs casos d'ús de la primera persona que he constatat en el DLC:

estranger -a *adj.* Que és d'un altre país. (...) Que és altre que el nostre país. [DLC]

estranger -a *adj.* Que és d'un altre país; que és altre que el nostre país. [DIEC]

⁷⁶ El DIEC i el DLC, respectivament, eliminen de la definició aquestes referències a la primera persona:

escurabutxaques *m. i f.* Persona o cosa que fa despendre tots els diners de què hom disposa. [DIEC]

escurabutxaques *m. i f.* (...) 2 *a* Persona que entabana algú per fer-li despendre els diners en profit propi. *b p ext* El joc és el seu escurabutxaques. [DLC]

⁷⁷ El DIEC manté la definició, mentre que el DLC n'elimina les referències pronominals:

do *m.* Acció de donar allò que és nostre a algú; la cosa així acordada. [DIEC]

do *m.* Acte de donar. [DLC]

⁷⁸ El DIEC manté la definició, mentre que el DLC l'adapta a la tercera persona:

dret *m.* Facultat d'exigir allò que ens és degut, de fer allò que la llei no defèn, de tenir, exigir, usar, etc., allò que la llei o l'autoritat estableix a favor nostre o ens és permès per qui pot. [DIEC]

dret *m.* Facultat d'exigir hom allò que li és degut, de fer allò que la llei no defèn, de tenir, exigir, usar, etc., allò que la llei o l'autoritat estableix a favor seu o li és permès per qui pot. [DLC]

contragirar *v. tr.* Girar, contra qui ens ha girat una lletra, una altra lletra (per una quantitat equivalent o menor).⁷⁹ [DGLC]

despatxar *v. tr.* Desfer-se (d'algú) donant-li comiat quan no necessitem o no volem més els seus serveis.⁸⁰ [DGLC]

demanar *v. tr.* Expressar a algú que desitgem obtenir d'ell (alguna cosa).⁸¹ [DGLC]

desemparar *v. tr.* Deixar sense empara, sense el nostre ajut o assistència (algú que ho necessita); (...); abandonar (una cosa) renunciant als drets que hi tenim.⁸² [DGLC]

En un cert nombre de definicions verbals, l'ús de la primera persona es produeix en els elements extrínsecs, i sol correspondre's a una identificació entre els complements de

⁷⁹ El DIEC manté l'ús de la primera persona, mentre que el DLC l'evita. Notem de passada, però, que tots dos suprimeixen els parèntesis en la definició, amb la qual cosa s'altera el contingut informatiu de la definició, ja que no és possible de determinar quin tipus de complement directe incorpora el verb *contragirar*:

contragirar *v. tr.* Girar, contra qui ens ha girat una lletra, una altra lletra per una quantitat equivalent o menor. [DIEC]

contragirar *v. tr.* Girar algú, contra el qui li ha girat una lletra, una altra lletra per una quantitat equivalent o menor. [DLC]

⁸⁰ L'evitació de l'ús de la primera persona es produeix en tots dos diccionaris, però en el DLC s'hi afegeix un tractament desafortunat dels elements extrínsecs:

despatxar *v. tr.* Desfer-se (d'algú) donant-li comiat perquè no es necessiten o no es volen més els seus serveis. [DIEC]

despatxar *v. tr.* Donar comiat a algú, desfer-se'n, engegar-lo. [DLC]

⁸¹ Tant en el DIEC com en el DLC s'evita l'ús de la primera persona:

demanar *v. tr.* Expressar a algú el desig d'obtenir d'ell (alguna cosa). [DIEC]

demanar *v. tr.* Expressar a algú el desig d'obtenir d'ell (alguna cosa). [DLC]

⁸² El DIEC elimina el pronom de la primera part de la definició, però manté el de la segona; el DLC adapta el pronom de la segona definició:

desemparar *v. tr.* Deixar sense empara, sense ajut o assistència (algú que ho necessita). (...) Abandonar (una cosa) renunciant als drets que hi tenim. [DIEC]

desemparar *v. tr.* Deixar sense empara, sense ajut o assistència. (...) Abandonar una cosa algú renunciant als drets que hi té. [DLC]

subcategorització i les diferents persones gramaticals, amb la finalitat d'aclarir fenòmens de naturalesa actancial:

dictar *v. tr.* (...) || Prescriure, imposar, formular a algú (allò que exigim d'ell).⁸³
[DGLC]

No es pot dir que hi hagi una utilització sistemàtica de les diferents persones gramaticals per a la descripció d'aquestes qüestions. Però sí que es pot destacar el fet que en alguns casos l'eliminació d'aquest recurs és innocu, mentre que en d'altres implica una pèrdua (o, almenys, un canvi) en la descripció del mot. Examinem els següents dos exemples:

reconvenir *v. tr.* Fer un càrrec o demanda (a algú que ens fa un càrrec, que adreça una acció contra nostre) per destruir o restringir els efectes d'aquest càrrec o acció.⁸⁴ [DGLC]

És a dir: *X* reconvé *Y* quan fa un càrrec o demanda contra *Y* i es dona el cas que *Y* ha fet abans un càrrec o demanda contra *X*.

tornar *v. intr.* Tr. (...) lliurar a algú (allò que ell ens havia lliurat, deixat, posat sota la nostra custòdia, etc.). [DGLC]

És a dir: *X* torna *Z* a *Y* quan lliura *Z* a *Y* i es dona el cas que *Y* ha posat abans *Z* sota la custòdia d'*X*.

⁸³ El DIEC manté la mateixa definició, mentre que el DLC n'altera força l'estructura, convertint l'element extrínsec en una llista d'elements impropis:

dictar *v. tr.* Prescriure, imposar, formular a algú (allò que exigim d'ell). [DIEC]

dictar *v. tr.* Prescriure, imposar, formular, lleis, sentències, condicions, etc. [DLC]

⁸⁴ Tant el DIEC com el DLC canvien les definicions eliminant-hi els pronoms de primera persona:

reconvenir *v. tr.* En dret processal, la persona demandada, exercir, en contestar, una acció (contra qui ha promogut la demanda) per destruir o restringir-ne els efectes. [DIEC]

reconvenir *v. tr. dr* Reclamar una persona contra una altra, que l'ha demandada o acusada, i convertir-se, d'acusat o demandat, en acusador o demandant. [DLC]

Tant el DIEC com el DLC eliminen l'ús de la primera persona en el cas de *tornar*, amb la qual cosa s'altera la definició del DGLC, ja que es perd la coreferència actancial entre el realitzador de l'acció de *tornar* i el receptor de l'acció de *deixar*:⁸⁵

tornar v. *intr.* (...) lliurar (alguna cosa deixada) a qui l'havia deixat. [DIEC]

tornar v (...) Lliurar a algú allò que ell havia lliurat, deixat, etc. [DLC]

Tots els casos d'utilització de la primera persona que he analitzat fins ara tenen en comú el fet que la primera persona gramatical introdueix en la definició un element de tipus [+HUMÀ] que sovint concorre amb altres elements, intrínsecs o extrínsecs, enunciats en tercera persona, que no incorporen necessàriament la característica [+HUMÀ], i que solen introduir-se amb pronoms de qualsevol mena (possessius, personals, indefinits, etc.) o bé amb expressions nominals (*alguna cosa*, etc.). Alguns d'aquests elements identifiquen arguments que són comuns entre el definit i la fórmula de la definició; són els actants de la situació denotada pel mot definit.

Aquests actants corresponen, en els exemples de *reconvenir* i de *tornar*, als elements que he reformulat amb l'ús de les variables *X*, *Y*, *Z*, etc. Notem que és condició *sine qua non* perquè un complement rebi aquest tractament el fet que es pugui identificar amb un argument del mot definit. En el cas de *reconvenir* no són actants complements del tipus *un càrrec o demanda*, ja que es refereixen exclusivament al descriptor de la definició i no s'identifiquen amb cap complement del verb *reconvenir*.

La presència d'actants en la definició, doncs, sol ser necessària en el cas dels verbs, però també en determinats tipus de noms capaços de contreure relacions lèxiques de tipus sintagmàtic amb elements subcategoritzats (habitualment introduïts per determinades preposicions), o bé en adjectius que presenten determinades restriccions selectives que cal descriure.

⁸⁵ Una altra qüestió és si aquesta identitat és exigible en la descripció semàntica d'aquest verb. Probablement ho és, a la vista de la incertesa de casos com ara: "En Pere va deixar el llibre a en Pau i en Joan l'hi ha tornat".

Ja he mostrat que el DIEC sol mantenir, respecte al DGLC, les definicions amb referències pronominals de primera persona. El DIEC, però, elimina totes les marques de *subjectivitat* en el cas dels elements extrínsecs de les definicions verbals.⁸⁶ El DLC, per la seva banda, elimina totes les referències a pronoms de primera persona.

Pel tipus de metallengua que solen utilitzar els diccionaris catalans, la diversificació de persones gramaticals ha de ser vista com un factor distorsionador de la regularitat expositiva del discurs lexicogràfic. De tota manera, cal convenir en el fet que la introducció de la primera persona del plural que fa el DGLC en la majoria dels casos tractats no és gratuïta, sinó que està motivada per la claredat expositiva. Com he dit, en lexicografia la tercera persona és la manera natural d'expressió de les definicions, però també he mostrat com, en el cas que s'acumulin en una mateixa definició diversos pronoms personals corresponents a diferents actants, es pot fer difícil (o massa artificios) de distingir aquestes persones si només ens atenim a l'ús dels pronoms de tercera persona.

Tal com he fet notar més enrere (veg. § 4.3), una condició fonamental per a l'expressió no ambigua de la definició lexicogràfica és la resolució del problema de l'associació dels components de la definició a l'estructura sintàctica del mot descrit. En bona mesura, l'adopció d'un sistema de notació que permeti de reflectir aquestes associacions esdevé fonamental també en el moment de la formulació de paràfrasis actancialment equivalents amb el definit. A partir de l'adopció d'aquest sistema, i sense alterar fonamentalment les definicions del DIEC que he citat anteriorment, obtindríem el següent resultat:

⁸⁶ En el cas que la marca de primera persona no figuri explícitament com un element extrínsec, l'eliminació no sembla ser tan sistemàtica. Es produeix en casos com *desfardar*, però no en casos com *despullar*:

desfardar v. tr. (...) || Pron. Treure's, discordar-se, afluir-se, allò que ens embarassa, ens acalora, etc. [DGLC]

despullar v. tr. (...) || Extreure (d'un conjunt de documents, d'un llibre, d'un text, etc.) els documents que ens convé d'examinar, els passatges, les notícies, que ens convenen.

- cosí** *m.* [N_1 és **N** de N_2] → Fill d'un oncle o d'una tia [d'algú]₂.
oncle *m.* [N_1 és **N** de N_2] → Germà del pare o de la mare [d'algú]₂.
sogra *f.* [N_1 és **N** de N_2] → Mare del marit o la muller [d'algú]₂.

Aquest sistema, que descriu el sentit dels mots referits mitjançant el patró d'inserció sintàctica (que ens indica que aquests substantius apareixen en construccions atributives amb *ser*), explicita mitjançant la subindexació la coreferència entre N_2 i l'element extrínsec [d'algú]₂, que és bàsica per a la correcta interpretació del significat.

Una mica més complexos són els casos de substantius que indiquen subjectivitat o estat d'ànim. De tota manera, el fet de recórrer al suport de les estructures argumentals com a element introductor dels actants semàntics (com a elements extrínsecs de la definició) és també un element determinant per a la formulació no ambigua de les definicions d'aquests mots. No és la meua intenció aquí de revisar l'adequació conceptual del contingut de les definicions, per la qual cosa faig una adaptació de cadascuna amb el que considero els mínims canvis indispensables d'estructura:

- desdeny** *m.* [**N** de N_1 per N_2] → sentiment que causa que [algú]₁ desdenyi [algú o alguna cosa]₂.
desgrat *m.* [**N** de N_1 per N_2] → desplaer que causa [a algú]₁ [alguna cosa que no és del seu grat]₂.
despit *m.* [**N** de N_1 per N_2] → irritació que causa [a algú]₁ [el desdeny que algú li mostra]₂.
fàstic *m.* [**N** de N_1 per N_2] → sentiment desagradable que causa [a algú]₁ [alguna cosa que li repugna]₂.
gelosia *m.* [**N** de N_1 per N_2] → enveja que causa [a algú]₁ [el qui gaudeix d'una cosa que es vol per a un mateix]₂.

Tal com es veu en algunes d'aquestes definicions, en ocasions es produeix la represa pronominal de l'actant corresponent a un element extrínsec dins d'un altre: *del seu grat*, *que li mostra algú*, etc. Aquestes coreferències no expressen cap ambigüitat; de tota manera, es podrien evitar mitjançant la inclusió de la fórmula *algú* amb el subíndex corresponent, dins de l'element extrínsec on es produeix la coreferència:

- desgrat** *m.* [**N** de N_1 per N_2] → desplaer que causa [a algú]₁ [alguna cosa que no és del grat d'algú]₂.

despit *m.* [**N** de N₁ per N₂] → irritació que causa [a algú]₁ [el desdeny que algú mostra a algú]₂

fàstic *m.* [**N** de N₁ per N₂] → sentiment desagradable que causa [a algú]₁ [alguna cosa que repugna a algú]₂

gelosia *m.* [**N** de N₁ per N₂] → enveja que causa [a algú]₁ [algú que gaudeix d'una cosa que algú₁ vol]₂.

remordiment *m.* [**N** de N₁ per N₂] → pena que causa [a algú]₁ [una mala acció de què algú₁ se sent culpable]₂.

En els casos precedents es pot vincular amb el patró sintàctic de cadascun dels substantius definits amb la situació que representen pel que fa a la distribució d'actants semàntics. Això fa que aquest tipus de solucions, a part d'artificioses des del punt de vista dels hàbits de la lexicografia convencional, siguin clarament dependents del patró sintàctic, ja que en darrer terme la possibilitat de declaració d'un actant com a element extrínsec depèn fortament de la presència en el patró del constituent sintàctic corresponent.

En la revisió dels casos d'utilització de la primera persona en els diccionaris convencionals he mostrat, a més, que hi ha un cert tipus de mots amb una descripció argumental que difícilment podria permetre l'aparició d'elements necessaris per a la desambiguació de la definició. En aquests casos els actants semàntics s'han d'incorporar a la part intrínseca de la definició lexicogràfica, pel fet que no constitueixen un argument del definit i, per tant, són elements propis pel que fa a la prova de la substitució. Examinem els casos vistos anteriorment d'*endarrera* (*adv.*) i *estranger* (*adj.*).

endarrera *adv.* En direcció a la banda que és darrera nostre; més o menys lluny en aquesta direcció. [DIEC]

estranger -a *adj.* Que és d'un altre país; que és altre que el nostre país. [DIEC]

Un possible tractament d'aquests casos (que fins i tot es podria fer extensible a la solució anterior a fi de fer-la menys formalitzada per a la interpretació humana) consisteix en la utilització d'elements que no impedeixin les possibilitats de substitució de la paràfrasi i que alhora indiquin amb claredat quin és l'element que s'estableix com

a punt de referència. Una possible solució a aquesta qüestió passa per la pressuposició de l'emissor de l'enunciat com a punt de referència, la qual cosa permet la inclusió de fórmules que l'identifiquin com l'actant semàntic que s'estableix com a punt de referència:

endarrera *adv.* [V ADV] → en direcció a la banda que és darrere d'allò de què es parla.

estranger *adj.* [N és A] → d'un país diferent al de qui parla o al d'algú de qui es parla.

Els casos d'utilització de la primera persona tractats en el grup 4 no presenten una problemàtica específica des d'aquest punt de vista. En el cas d'*escurabutxaques* i de *do* l'estructura argumental no subcategoritza elements ambigus dins de la definició. La solució a aplicar en el cas de *dret* no difereix de la que hem vist anteriorment.

Pel que fa a les definicions verbals la solució proposada per als substantius de naturalesa subjectiva (amb represa pronominal o sense dins dels elements extrínsecs) resol amb un grau d'adequació prou satisfactori els diferents tipus de requeriments de coreferenciació:

contragirar *v. tr.* Girar, contra qui ens ha girat una lletra, una altra lletra (per una quantitat equivalent o menor). [DGLC]

contragirar *v. tr.* [N₁ V N₂ a N₃] → [Algú]₁ girar una lletra [per una quantitat equivalent o menor]₂ contra [algú que ha girat prèviament una altra lletra a algú]₃.

despatxar *v. tr.* Desfer-se (d'algú) donant-li comiat quan no necessitem o no volem més els seus serveis. [DGLC]

despatxar *v. tr.* [N₁ V N₂] → [Algú]₁ desfer-se d'[algú que no és necessari per a algú]₂ donant-li comiat.

dictar *v. tr.* (...) || Prescriure, imposar, formular a algú (allò que exigim d'ell). [DGLC]

dictar *v. tr.* [N₁ V N₂ a N₃] → [Algú]₁ formular [a algú]₃ [alguna cosa que algú₁ li exigeix]₂.

Val la pena destacar, finalment, que un sistema d'aquest tipus no introdueix una major complexitat interpretativa en la definició respecte a la que ja és habitualment present en la lexicografia convencional. L'avantatge d'aquest tipus de notació és que introdueix

sobre els enunciats convencionals uns canvis mínims que són bàsics per a l'adequació descriptiva, i per aquest motiu donen lloc a definicions lexicogràfiques que, llegides seqüencialment (això és, sense valorar la distinció entre elements intrínsecs i elements extrínsecs), sistematitzen les definicions convencionals i hi aporten una major claredat descriptiva.

4.5 Equivalència esdevenimental

Una de les qüestions de què més s'ha ocupat la semàntica de les últimes dècades —sobretot a partir dels estudis de Vendler (1967)— és l'anàlisi i la determinació d'allò que es coneix com *aspecte*. L'aspecte estudia els tipus de predicats lingüístics en relació a les situacions que denoten, i que es determinen en les seves característiques esdevenimentals. Aquest marc d'estudi, que parteix ja de les observacions aristotèliques i que ha estat un centre d'interès remarcable en la filosofia del llenguatge, té avui una importància considerable en la lingüística teòrica. Els darrers corrents lexicalistes, a més, han fet d'aquesta qüestió un tema fonamental, cosa que ha afavorit el progrés teòric en aquest sentit.

Des del punt de vista de la consideració lexicogràfica del problema de l'aspecte, podríem enunciar la qüestió de la següent manera: en principi, un mot *A* que denota una situació esdevenimental *S* ha de tenir com a paràfrasi de definició l'expressió *P* que denoti la mateixa situació esdevenimental *S*.

Aquest principi formula la necessitat (no exempta d'alguns problemes teòrics i pràctics, com mostraré en aquesta mateixa secció) de l'equivalència esdevenimental entre definició i definit. L'aspecte ha estat un dels elements en què s'han basat les teories d'anàlisi composicional del significat lèxic; per tant, les informacions aportades per les categories aspectuals poden tenir, almenys a priori, incidència en l'expressió de l'enunciat de definició, ja que aquest és, en darrer terme, la descomposició del significat

lèxic d'un mot. Un dels requisits de les possibilitats d'intercanvi entre definit i definició és, doncs, l'equivalència també en aquest terreny.

Cal, però, que abans d'abordar la transcendència d'aquesta qüestió en la descripció lexicogràfica, examini sumàriament la noció d'aspecte i faci algunes consideracions associades a la seva naturalesa. Marín (1999: 91) defineix l'aspecte com “la informació temporal associada inherentment al significat dels predicats o (...) l'estructura temporal interna en la qual se situa l'acció denotada per la predicació”.⁸⁷

Miguel (1999) fa una presentació resumida de les diferents informacions que vehicula el denominat *aspecte lèxic*:

El aspecto informa sobre la manera en que un evento se desarrolla u ocurre: implicando un cambio (por ejemplo, en el caso de *madurar*) o la ausencia de cambio (por ejemplo en el caso de *estar verde*); alcanzando un límite (por ejemplo *llegar*) o careciendo de él (*viajar*); de forma única (por ejemplo, *disparar*) o repetida (*ametrallar*); de forma permanente (*ser español*), habitual (*cortejar*) o intermitente (*parpadear*).

El aspecto informa también sobre la extensión temporal del evento: un período no acotado de tiempo (com en el caso de *ser inteligente*), un intervalo acotado (en el caso de *madurar*) o un instante (en el caso de *explotar*); sobre cual es la fase principal del evento descrito: el inicio (como en *floreecer*), la fase media (como en *envejecer*) o la fase final (como en *nacer*); asimismo, el aspecto puede informar sobre la intensidad con que el evento tiene lugar: por ejemplo, *peinar* es un evento de intensidad neutra con respecto al intensivo *repeinar* y al atenuativo *atusar*. [Miguel (1999: 2979)]

⁸⁷ Alguns autors utilitzen o esmenten el terme *Aktionsart* per referir-se a aquesta noció. El fet que determinades llengües distingeixin, com assenyala Marín (1999: 92), entre les característiques temporals inherents al predicat (*aktionsart*) i l'expressió morfològica d'aquest significat (*aspecte*) no ha estat acollit en general com una distinció terminològica general. En aquest treball utilitzaré el terme *aspecte*, doncs, per referir-me a la caracterització de la denotació de les situacions, mentre que per referir-me a la denotació introduïda per l'ús de variants morfològiques determinades (en el cas dels verbs) em referiré (seguint Miguel, 1999) a *aspecte flexional*.

Malgrat que els estudis sobre l'anàlisi aspectual s'han referit de forma preferent a la categoria verbal, cal remarcar la *intercategorialitat* que representen les categories aspectuals. Determinats adjectius (*intel·ligent, savi, holandès*) determinen o seleccionen la naturalesa esdevenimental dels predicats en què poden aparèixer; de forma similar, certs noms que subcategoritzen estructures argumentals són també susceptibles de constituir-se en nuclis d'àmbits de predicació i, per tant, de ser analitzats en termes del seu aspecte lèxic: *la redacció d'un diccionari* denota una situació processiva, mentre que *l'esternut d'en Pere* denota una situació puntual.

Una altra qüestió important en la teoria de l'aspecte, que s'ha remarcat a bastament, és la seva *composicionalitat* en tant que categoria lingüística. Tot i que en la descripció de l'aspecte es pressuposa un cert caràcter inherent de les categories aspectuals als lexemes, cal aclarir que aquestes categories no solen estar associades de manera estàtica a una entrada lèxica, sinó que poden ser modificades per la naturalesa dels elements que formen part del predicat.

Pres aïlladament, el predicat *fumar* denota una situació no delimitada, en el sentit que no és necessari que hagi acabat perquè hagi tingut lloc, ja que alhora que algú fuma ja ha estat fumant. Si modifiquem els arguments del verb i hi afegim un complement directe (*fumar una cigarreta*), la situació denotada esdevé diferent, finalitzarà quan el subjecte hagi acabat de fumar la cigarreta, i no s'esdevé que mentre algú està fumant una cigarreta, ja hagi fumat una cigarreta.⁸⁸

Lea anàlisis lingüístiques sobre l'aspecte assenyalen que el valor aspectual d'un predicat també pot veure's modificat per l'*aspecte flexional*. Així, un verb que denota estat com

⁸⁸ Miguel (1999) fa referència a aquesta qüestió: "Puede decirse, entonces, que en este caso la raíz verbal no es la responsable única de la información aspectual referida a la ausencia o presencia de límite interno en el evento. Este no es un hecho trivial. Por el contrario, los estudios sobre el aspecto léxico de los verbos tienen que enfrentarse de forma constante a la paradoja (al menos terminológica) de que lo que se conoce como aspecto léxico no tenga que ver –al menos, no sólo– con el verbo como unidad léxica." [Miguel (1999: 2985)]

ara *saber* (*saber francès, saber multiplicar*) pot adquirir, en temps de perfet, les característiques d'un verb de tipus incoatiu: *aleshores va saber que l'enganyaven* denota un canvi 'passar a saber' que no admet la categoria de verbs estatus. La lexicografia sol tractar aquests casos com a accepcions diferents d'una mateixa entrada.

El fet que l'aspecte sigui un fenomen de naturalesa composicional ha fet que l'aproximació teòrica al problema de l'anàlisi aspectual hagi desenvolupat sistemes de descripció basats en combinatòries de components que s'acosten als primitius semàntics. Tal com diuen Chierchia & McConnell-Ginet (1990):

Decompositional accounts of lexical meaning have been proposed to elucidate semantic relations among lexical items, apparent semantic universals, aspectual differences, and constraints on possible interactions of conceptually complex words with modifying expressions such as adverbials. [Chierchia & McConnell-Ginet (1990: 360)]

Aquest composicionalisme arrenca de l'anàlisi de Dowty (1979), que delimita tres operadors aspectuals: DO, BECOME, CAUSE, i suggereix que els verbs poden ser definits a partir de predicats estatus formats en termes d'aquests operadors. Pustejovsky (1992) assenyala el reduccionisme que representa el tractament d'aquest tipus de fenòmens com a constructes de diferents primitius, i afegeix que:

any method assuming a fixed number of primitives, however, runs into some well-known problems with being able to capture the full expressiveness of natural language. [Pustejovsky (1992: 54)]

Aquesta variabilitat que presenten les categories aspectuals en funció del tipus de predicat i de les característiques flexionals dels lexemes condueix a una qüestió fonamental: és l'aspecte una categoria lèxica? O, dit d'una altra manera, fins a quin punt el lèxic és el nivell d'anàlisi adequat per a les categories aspectuals? Marín (1999: 113) assenyala que "hi ha arguments prou sòlids per creure que l'anàlisi aspectual no se situa al nivell lèxic"; de tota manera, cal no oblidar que els treballs sobre els *esdeveniments* (*events*) en general s'han basat en el marc de les teories composicionalistes del

significat lèxic, de manera que s'assumeix implícitament que les propietats esdevenimentals estan associades a l'entrada lèxica. Pustejovsky (1992) introdueix una interpretació de l'estructura esdevenimental que permeten l'explicitació dels canvis que es produeixen en la sintaxi dels *event types* de cada lexema:

Event structure is just one level of the semantic specification for a lexical item, along with subcategorization, argument structure, and qualia structure. Because an event structure is recursively defined in the syntax, "event type" is also a property of phrases and sentences. [Pustejovsky (1992: 55)]

Segons Pustejovsky (1992: 55), els tres components bàsics de l'estructura esdevenimental són: *a)* l'*event type* primari del lexema, *b)* les regles de la composició d'*events*, i *c)* les *mapping rules* de l'estructura lèxica.

Així, doncs, parteix de la base que, malgrat que el nivell lèxic no és l'únic nivell determinant per a l'anàlisi esdevenimental, hi ha característiques bàsiques de la denotació aspectual que es determinen en el pla del lèxic. Així un verb, per exemple, està associat a un dels tres tipus bàsics d'esdeveniment (estats, processos, transicions), però la denotació final dels predicats en què intervé no està només determinada lèxicament, sinó també gramaticalment; de manera que un verb com *tancar*, que bàsicament és una transició, pot aparèixer en predicats de naturalesa diferent:

La porta és tancada [ESTAT]

La porta s'ha tancat [PROCÉS]

En Pere ha tancat la porta [TRANSICIÓ]

Els primers treballs sobre els tipus aspectuals recullen classificacions com la que es pot trobar a Vendler (1967). Vendler estableix, referint-se als verbs, quatre classes aspectuals basant-se en les distincions dels paràmetres de *processivitat* i de *telicitat*, que es poden resumir en el quadre de la figura 4.8.

	Processivitat	Telicitat
Estat	–	–
Activitat	+	–
Realització	+	+
Assoliment	–	+

Figura 4.8. Classes aspectuals de Vendler (1967).

La processivitat fa referència al tipus denotatiu basat en les fases de què consta una operació qualsevol; les operacions processives consten de diverses fases mentre que les no processives consten d'una sola fase. La distinció entre situacions tèliques i atèliques està basada en la denotació del punt final; un predicat és tèlic quan implica una culminació, mentre que és atèlic si no porta implícit cap punt final intrínsec.

Chierchia & McConnell-Ginet (1990) caracteritzen genèricament cadascun d'aquests tipus de la següent manera:

States are like snapshots of the world at a given instant. They lack a natural culmination or end point, and their subject is perceived not as an agent (as doing something) but as an experiencer (as experiencing something). (...) Activities share with states the property of lacking a natural culmination. Yet they are agentive in that they typically involve a subject doing something. (...) Telic eventualities have a natural end point or culmination. [Chierchia & McConnell-Ginet (1990: 353)]

Cadascuna d'aquestes categories aspectuals pot ser identificada amb tipus de predicats lingüístics. *Romandre* s'identifica amb un estat, *caminar* amb una activitat (o procés); pel que fa als esdeveniments tèliques, se sol distingir entre *realitzacions* (accions amb un acabament implícit, com ara *construir una casa*) i *assoliments* (situacions puntuals, com ara *descobrir*).

Dowty (1979) estableix de manera similar tres classes aspectuals per a classificar els diferents predicats: estats, activitats i esdeveniments tèlics, i desenvolupa una bateria de qüestions per a la classificació aspectual dels verbs. Marín (1999) fa una adaptació al català d'aquest test, que es pot veure a la figura 4.9.

		Estats	Activitats	Realitzacions	Assoliments
1	Admet el progressiu i ser predicat de <i>persuadir</i>	No	Sí	Sí	?
2	Interpretació habitual en present	No	Sí	Sí	Sí
3	Durant X temps	Sí	Sí	Sí	No
4	En X temps	No	No	Sí	Sí
5	Φ durant una hora implica Φ a qualsevol moment de l'hora	Sí	Sí	No	--
6	Està Φ -ant implica ha Φ -at	--	Sí	No	--
7	Complement de <i>deixar</i>	Sí	Sí	Sí	No
8	Complement d' <i>acabar</i>	No	No	Sí	No
9	Ambigüïtat amb <i>quasi</i>	No	No	Sí	No
10	Φ en una hora implica estar Φ -ant durant aquesta hora	--	--	Sí	No
11	Apareix amb adverbis del tipus <i>acuradament</i>	No	Sí	Sí	No

Figura 4.9. Adaptació al català del test de Dowty (1979) extreta de Marín (1999).

Cal assenyalar, però, que aquests tipus de classificacions aspectuals són, des del punt de vista lexicogràfic, més o menys problemàtiques, sobretot perquè hi ha factors que condicionen la suposada existència de les *característiques aspectuals inherents*. La suposició d'aquesta existència equival a postular que les categories aspectuals poden identificar-se en els lexemes de forma descontextualitzada; d'alguna manera, aquesta suposició representa un cert contrasentit amb la idea que les denotacions aspectuals es verifiquen sempre sobre predicats i no pas estrictament sobre els lexemes presos de manera aïllada.

Assumint de manera acrítica l'existència d'aquestes categories inherents podem pensar, doncs, que les característiques aspectuals dels verbs (o d'altres classes gramaticals) poden tenir incidència en (o, fins i tot, determinar) algunes seleccions de complements o

algunes possibilitats d'inserció sintàctica. De tota manera, sembla clar que, en qualsevol cas, el tipus aspectual d'un verb no determina sempre els tipus aspectuals dels enunciats en què apareix. Un verb com *anar* (un procés), per exemple, pot aparèixer en una frase com *anar a la masia*, que pertany a una classe esdevenimental diferent de la que correspon a *anar* en virtut d'una anàlisi descontextualitzada. El fet que la lexicografia, almenys en la forma amb què es presenta en aquest treball, no es pugui desvincular de l'anàlisi contextualitzada dels lexemes posa en dubte les possibilitats reals d'integrar (almenys en la paràfrasi de definició) informacions relatives a l'aspecte.

A fi de determinar amb un grau major de claredat l'àmbit que és susceptible de tractament lexicogràfic, es pot partir de l'esquema aportat per Miguel (1999),⁸⁹ que determina les diferents manifestacions de l'aspectualitat. Aquest esquema correspon al quadre que es recull a la figura 4.10.

Aspectualitat					
Verbal				Oracional	
Oposició de formes d'un mateix verb (imperfet/perfet simple)	Afixos derivatius (<i>re-</i>)	Oposició de classes aspectuals de verbs (<i>viatjar/arribar</i>)	Certes combinacions de verbs (modes d'acció analítics: perífrasis verbals)	Marques lèxiques i funcionals (adverbis, negació)	Característiques gramaticals dels participants en l'event (funció semàntica i sintàctica, nombre, determinació i quantificació)
Aspecte flexional	Aspecte lèxic		Aspecte lexicosemàntic		

Figura 4.10. Tipus d'aspectualitat segons Miguel (1999).

⁸⁹ No he consultat la font en què es basa Miguel per a la composició d'aquest esquema: MASLOV, Y. S. (1978), "An Outline of Contrastive Aspectology", in Maslov, Y. S. (ed.) (1985), *Constructive Studies in Verbal Aspect*, Heidelberg: Gross, 1-44.

D'acord amb aquesta classificació, l'àmbit de l'aspecte lèxic es refereix només a dos subtipus de l'aspecte verbal: el dels afixos de derivació que poden modificar o matisar les característiques aspectuals dels verbs i les classes aspectuals. Resten fora de l'àmbit lèxic les modificacions introduïdes per altres verbs (per exemple, els verbs modals), la modificació adverbial i les característiques dels arguments.

Tot i que no m'he ocupat específicament de la incidència en l'aspecte lèxic dels afixos derivatius, sembla evident que aquests solen tenir una repercussió clara en la definició lexicogràfica. Prefixos com *re-*, *contra-*, etc. sovint impliquen la inclusió en la definició de fórmules (*tornar a...*, etc.) que traslladen a l'enunciat de definició la modificació esdevenimental en tant que modificació semàntica. La modalitat esdevenimental, en tant que característica lèxica, pot mantenir certes relacions, manifestades de manera no tant evident, amb determinats processos derivatius, de manera que sembla que en alguns casos la naturalesa del verb pot veure's alterada per la presència de determinats afixos i processos. Examinem, per exemple, el cas d'alguns verbs relacionats amb la base lèxica *dormir*; en concret, ens referirem als recollits al DIEC, és a dir: *dormir*, *dormisquejar*, *dormitar*, *adormir(-se)*, *adormissar-se* i *condormir-se*. El DIEC dona com a descripció d'aquests verbs la informació que resumeixo en el quadre de la figura 4.11.

Atenent a les característiques aspectuals d'aquests verbs podem distingir, a grans trets, dos grups amb diferents propietats de predicació:

- Les activitats, o processos sense una culminació: *dormir*, *dormitar*, *dormisquejar*.
- Els assoliments, que denoten situacions processives amb un punt de culminació: *adormir*, *adormir-se*.

El cas d'*adormissar-se* i de *condormir-se* mereix un comentari a part, ja que, si n'analitzem la definició, es tracta de verbs que es refereixen a *adormir* com a activitat, més que no pas com a assoliment. *Mig adormir-se* o *començar a adormir-se* es poden

identificar amb activitats que no tenen culminació, però que (almenys si ens atenim a la definició lexicogràfica) es refereixen a situacions que són culminatives (com *adormir*).

Entrada	Categoria	Definició	Exemples
dormir	<i>v. intr.</i>	Una persona o un animal, estar en aquell estat de repòs, anomenat son, caracteritzat per la suspensió de certes funcions de l'activitat vital.	<i>Ara dorm: no el despertis. Anar-se'n a dormir. Posar-se a dormir. Aixecar-se de dormir. Aquesta nit he dormit sis hores. He dormit tota la nit d'un son, en un son. No he pogut dormir en tota la nit. Dormir com les llebres</i>
dormisquejar	<i>v. intr.</i>	Estar mig adormit.	
dormitar	<i>v. intr.</i>	Estar mig adormit.	
endormiscar-se	<i>v. pron.</i>	Mig adormir-se.	
adormir	<i>v. tr.</i>	Fer dormir.	<i>L'han adormit bressant-lo.</i>
	<i>pron.</i>	Començar a dormir.	<i>No m'he adormit fins a mitja nit.</i>
adormissar-se	<i>v. pron.</i>	Mig adormir-se, ensopir-se, abaltir-se.	
condormir-se	<i>v. pron.</i>	Començar a adormir-se, adormissar-se.	

Figura 4.11. Informacions lexicogràfiques del verb *dormir* i derivats.

Si s'apliquen els tests de Dowty (en la formulació de Marín, 1999) a aquesta sèrie de verbs, s'obté un quadre com el de la figura 4.12.

Analitzant componencialment cadascuna de les definicions lexicogràfiques que corresponen a la sèrie de mots relacionats lèxicament amb *dormir* es pot veure amb certa facilitat que el discurs de la lexicografia convencional no sol descriure els predicats esdevenimentals en funció de la seva naturalesa situacional. *Dormisquejar* i *dormitar* incorporen *estar* com a descriptor *esdevenimental* (a la manera dels descriptors que caracteritzen sintàcticament el definit, que he tractat a § 4.1; veg. també Ilson, 1987), la qual cosa els situaria entre els estats, des del punt de vista del seu aspecte lèxic. El fet, però, que siguin estats que es refereixen a processos de tipus culminatiu fa que no siguin interpretables en termes d'estats purs, sinó més aviat en termes de processos. La prova d'això és que no es comporten com a estats d'acord amb els tests de

Dowty, i que d'acord amb aquestes proves són equiparables a *dormir*. També es poden veure en aquesta sèrie de verbs altres tipus de descriptors esdevenimentals com ara *començar a*, que representa genèricament la idea del caràcter *incoatiu* de certs verbs respecte de la situació que denoten. Altres elements de la definició com ara *mig* seguit d'infinitiu, també matisen les característiques esdevenimentals del predicat.

		<i>Dormir</i> <i>Dormitar</i> <i>Dormisquejar</i>	<i>Adormir</i>	<i>Adormir-se</i>	<i>Endormiscar-se</i>	<i>Adormissar-se</i>	<i>Condormir-se</i>
1	Admet el progressiu i ser predicat de <i>persuadir</i>	Sí	Sí	Sí	Sí	Sí	Sí
2	Interpretació habitual en present	Sí	Sí	Sí	Sí	Sí	Sí
3	Durant x temps	Sí	No	No	Sí	Sí	?
4	En x temps	No	Sí	Sí	No	No	?
5	Φ durant una hora implica Φ a qualsevol moment de l'hora	Sí	No	No	Sí	Sí	?
6	Està Φ -ant implica ha Φ -at	Sí	No	No	No	No	?
7	Complement de <i>deixar</i>	Sí	No	No	Sí	Sí	Sí
8	Complement d' <i>acabar</i>	No	Sí	Sí	No	No	No
9	Ambigüitat amb <i>quasi</i>	No	No	No	No	No	No
10	Φ en una hora implica estar Φ -ant durant aquesta hora	--	No	No	--	--	--
11	Apareix amb adverbis del tipus <i>acuradament</i>	No	No	No	No	No	No

Figura 4.12. Anàlisi aspectual del verb *dormir* i derivats.

Des del punt de vista estrictament formal, en els casos de *dormitar* i d'*endormiscar-se*, tal vegada fóra preferible la selecció d'un descriptor no estatiu com ara *esdevenir*, però no és clar que des del punt de vista de la interpretació final de la descripció semàntica d'aquest verb aquest tipus de canvi fos defensable en termes absoluts.

Les limitacions que imposa el mètode lexicogràfic per la naturalesa del metallenguatge de descripció que utilitza fan difícil la formulació de propostes de tractament genèriques per a assegurar l'equivalència aspectual entre definit i definició. En molts casos, el fet

que el descriptor de la definició reculli només una part de la naturalesa semàntica del definit (deixant de banda el seu aspecte lèxic) implica que definit i definició no tinguin el mateix valor aspectual, com per exemple:

fumar v. tr. (...) Aspirar el fum (d'un cigar, d'una pipa, etc.). *Fumar un cigar, una pipa. Els xinesos fumen opi.* [DIEC]

Un predicat com ara *fumar opi* és no culminatiu, i representa un procés que pot ser identificat, a grans trets, amb el d'*aspirar el fum (de l'opi)*. El mateix verb, però, en el predicat *fumar un cigar*, intervé en un assoliment: X no ha fumat un cigar fins que no ha acabat de fumar el cigar. Si projectem aquesta distinció sobre el predicat *fumar un cigar* i sobre la definició *aspirar el fum (d'un cigar)* obtindrem esquemes de comportament aspectual diferents per a cadascun dels predicats. Una prova intuïtiva d'això és que X pot aspirar el fum de tres cigars i no haver-ne fumat cap. Tal vegada aquesta definició podria millorar les seves capacitats de descripció amb un descriptor més pròxim al tipus esdevenimental del verb a què es refereix:

fumar (...) Consumir (un cigar, una pipa) aspirant-ne el fum.

La formulació d'una proposta fonamentada en aquest terreny implica la necessitat de fer un estudi aprofundit sobre la casuística i sobre les possibilitats reals de delimitar el que he anomenat descriptors esdevenimentals, a fi de veure quina ha de ser la seva interacció amb els descriptors semàntics d'una definició.

Tot i reconeixent la dificultat de plantejar esquemes de solució en aquest sentit, sí que fóra desitjable de remarcar el fet que una lexicografia normalitzada ha d'evitar la tendència a acumular diferents paràfrasis en una mateixa accepció, tal com assenyalen Rey-Debove (1971: 208) i Martin (1977b: 62). Un cert nombre de definicions múltiples (que podem caracteritzar com una de les desviacions metodològiques més corrents de la lexicografia convencional) poden deure's justament a la dispersió aspectual que poden presentar els lexemes:

caminar *v. intr.* Traslladar-se a peu d'un lloc a un altre; avançar a passes, no aixecant completament un peu que l'altre no toqui a terra; els quadrúpedes, marxar tenint sempre almenys dos peus a terra; un animal, posar en execució la seva manera normal de locomoció; alguna cosa, fer el seu camí. *Caminar a bon pas. No van trobar cap cotxe i van haver-hi d'anar caminant. Avui haurem de caminar vuit hores. I el noieta, camina que caminaràs, no va arribar a la cabana fins al vespre. Caminar de puntetes. Els cangurs caminen a salts.* [DIEC]

viatjar *v. intr.* Fer viatge, anar de viatge. *M'agrada molt de viatjar. Viatjar en tren, en auto. Viatjar d'incògnit.* [DIEC]

encendre [quant a la flexió, com atendre] *v. tr.* Fer que (una cosa) cremi, comunicar el foc (a una cosa), esp. amb la intenció de servir-se'n. *Encendre el carbó, la llenya, el foc. Encendre el llum, l'espelma. Encendre el cigar, la pipa.* [DIEC]

El tractament de l'equivalència esdevenimental des del punt de vista lexicogràfic ha de fer front a diferents tipus de qüestions, entre les quals ocupa un lloc destacat la variabilitat aspectual dels lexemes en funció dels predicats de què poden formar part. Ja he mostrat que el mode aspectual d'un lexema no és estàtic, i que pot variar en funció de l'estructura argumental i de la naturalesa dels seus arguments. Aquest fet implica, novament, que l'equivalència depèn en bona mesura de la relació existent entre patrons sintàctics i definició.

L'estat actual de sistematització dels coneixements sobre l'estructura aspectual dels lexemes no està encara en condicions de permetre un tractament d'aquests fenòmens de forma integrada amb el discurs de la definició lexicogràfica, i les úniques propostes de tractament van en la línia de la utilització de *marques aspectuals* externes a la definició (Sol, 1992b). En part això es deu al fet que no s'ha anat més enllà dels models teòrics i que encara no s'ha pogut constatar que es pugui disposar d'un sistema de primitius integrable en la descripció lexicogràfica. La formulació de definicions adequades depèn fonamentalment de les capacitats de detectar i d'aïllar fórmules de definició que permetin el tractament dels tipus esdevenimentals com a component semàntic del lexema.

De tota manera, convé destacar que, si hom assumeix que hi ha raons per determinar el caràcter lèxic de certs atributs aspectuals, és possible, tal com he mostrat, la descripció

lexicogràfica d'aquests atributs a partir d'enunciats lingüístics que recullin descriptors de tipus esdevenimental, referits a les característiques que podem atribuir al lexema.

4.6 Equivalència sincategoremàtica

Els corrents més recents de la semàntica lèxica han posat una atenció especial en l'establiment de regularitats referides a les propietats de subcategorització sintàctica del mot. La predictibilitat adquireix, així, un cert protagonisme en el terreny de la semàntica, que tradicionalment havia concentrat bona part dels seus esforços en l'estudi dels comportaments idiosincràtics. Els formalismes d'unificació de les teories lexicalistes i la semàntica generativa s'han ocupat especialment de l'establiment de mecanismes que permeten la projecció regular de determinats fenòmens de naturalesa lèxica, com ara les variacions metonímiques del significat o el que podríem anomenar les modificacions selectives d'atributs semàntics (Soler, 2001).

Aquesta tendència ha afavorit, complementàriament, la reflexió sobre els aspectes estructurals dels diccionaris, i sobre fins a quin punt les recerques teòriques i les aportacions de la lexicografia computacional poden determinar estratègies per a la descripció de determinats fenòmens del comportament del lèxic.

Una de les construccions que han estat estudiades per la semàntica lèxica amb un interès preferent és un subtipus dels verbs que admeten la construcció de les *anàfores de complement nul* (Brucart, 1999). Aquesta construcció es produeix en verbs que accepten de controlar una subordinada d'infinitiu amb un complement d'objecte i permeten l'elisió de tota la subordinada,⁹⁰ com es pot veure en l'exemple de (5.1).

⁹⁰ Genèricament, aquesta construcció afecta els verbs modals (*voler, poder, etc.*), els verbs denominats aspectuals (*començar a, acabar de, etc.*), els verbs que (seguint Bosque, 1999) expressen predisposició, actitud o propòsit (*acceptar, dubtar de, aprendre a, etc.*), els verbs que expressen incidència sobre l'actitud

(5.1) Li van proposar viatjar per tot Europa, però va refusar \emptyset

En alguns casos, però, aquesta subordinada d'infinitiu, com en els casos de (5.2), pot ser elidida i el complement d'objecte pot passar a ser complement del verb controlador; això dóna lloc a construccions del tipus (5.3), que poden tenir el significat de les expressions *a* o *b* de (5.2).

(5.2a) En Pere va començar a llegir el llibre

(5.2b) En Pere va començar a escriure el llibre

(5.3) En Pere va començar el llibre

El problema que plantegen aquest tipus de construccions té una doble dimensió: sintàctica i semàntica. Des del punt de vista sintàctic, *començar* demana com a complement una oració infinitiva, no un substantiu. Des del punt de vista semàntic, *començar* exigeix un argument amb una denotació esdevenimental (*començar la lectura d'un llibre*), no pas un nom amb una denotació d'objecte. A més, aquest tipus de construcció, que implica la conservació del valor semàntic de les frases alternants, no l'accepten altres verbs controladors d'infinitiu, com mostra l'exemple de (5.4).

(5.4) En Pere va acceptar de llegir l'informe \neq En Pere va acceptar l'informe

En principi, es pot suposar que el tipus de construccions esmentades més amunt són possibles no només en funció de les característiques dels verbs que les permeten, sinó també en funció de la naturalesa semàntica (o, més precisament, esdevenimental) del substantiu que es projecta com a complement del verb principal, i que en bona mesura la interpretació prototípica de les construccions sense infinitiu depèn de l'estructura

d'altri (*autoritzar a, ensenyar a, incitar a, etc.*), i finalment alguns verbs pronominals que admeten l'elisió del seu complement de règim (*abstenir-se de, disgustar-se per, decidir-se a, etc.*)

semàntica del substantiu. Aquest fenomen s'ha designat amb el terme de *sincategorematicitat* per Bosque (1999, 2000):

El término 'sincategorematicidad' se ha usado con varios sentidos, pero es muy habitual referirse a las 'palabras sincategoremáticas' para aludir a las que denotan propiedades que requieren hacer mención a eventos o a acciones que no se mencionan expresamente en su propia definición ni prevé –en principio– la naturaleza léxica de tales sustantivos. Naturalmente, el problema es el de precisar cómo obtenemos exactamente la información ausente en esos casos. [Bosque (1999: 264)]

En estreta connexió amb el caràcter sincategoremàtic de determinats substantius, també hi ha, tal com assenyala Bosque (1999), el fenomen de determinades modificacions adjectivals, que no es projecten sobre la totalitat dels trets semàntics del substantiu, sinó que només es refereixen a una part del seu significat conceptual, tal com es mostra en els exemples de (5.5).

(5.5a) Un problema difícil de resoldre

(5.5b) *Un problema difícil*

Així, les construccions de 5.5 es poden considerar sinònimes, ja que l'adjectiu modifica l'atribut semàntic de la *qualitat de resoluble* del problema. Això permet l'elisió de la construcció infinitiva. En canvi, altres construccions, com les de 5.6a i 5.7a no permeten el mateix tipus de modificació (5.6b i 5.7b):

(5.6a) Un fet difícil d'explicar

(5.6b) **Un fet difícil*

(5.7a) Distàncies difícils de calcular

(5.7b) **Distàncies difícils*

Tal com assenyala Bosque (1999), el tractament lexicogràfic d'aquest tipus de fenòmens no sol reflectir la possibilitat d'aquestes alternances, ni en la descripció dels noms ni en

la dels lexemes corresponents a les categories adjectival o verbal. Una solució lexicogràfica a aquestes qüestions no és evident ni fàcil de donar, atesa la naturalesa *absent* de la informació que resulta necessària per a la interpretació dels casos precedents.

El fet que aquestes modificacions tinguin una naturalesa selectiva sobre el nucli (és a dir, no es modifica la totalitat del nucli sinó només algun dels seus atributs semàntics) ha fet que s'hagin assajat propostes d'explicació com la de les *zones actives* de Langacker (1984), que es basa en el principi que un nom pot intervenir en determinats predicats no denotant-se com una totalitat sinó denotant només alguna de les seves parts. Kleiber (1999) fa una adaptació crítica de la proposta de les *zones actives* al cas de *començar un llibre* i conclou que la proposta de Langacker no és descriptivament satisfactòria, ja que en darrer terme implica el tractament del cas com a polisèmia verbal entre les dues estructures possibles del verb començar (*començar a llegir un llibre / començar un llibre*).⁹¹ L'anàlisi de Langacker es concentra en el procés verbal molt més que no pas en allò que permet l'alternança de predicats, que són més aviat les característiques específiques que presenta el substantiu.

Un concepte que s'ha introduït amb un cert èxit per al tractament d'aquestes estructures és el de la subespecificació. La subespecificació en la representació del lèxic (Buitelaar, 1998) es basa en la idea que la descripció d'un lexema no ha de contenir de manera desenvolupada i estàtica totes les seves característiques semàntiques, sinó que cal descriure de forma general els mecanismes a partir dels quals pot adquirir un determinat

⁹¹ Segons Kleiber (1999); “Langacker joue la carte de la polysémie verbale et défend l’hypothèse que le verbe *commencer* en construction directe avec un SN objet (...) n’a plus le même sens que dans la construction avec infinitif. (...) Dans les cas de la construction avec infinitif, c’est le procès denoté par l’infinitif qui se trouve choisi comme landmark [= argument objet, en la terminologia de Langacker], alors que dans la construction avec le SN comme objet direct c’est le landmark du procès qui est landmark dans le premier cas qui se trouve choisi comme landmark du predicat *commencer*.” [Kleiber (1999: 160)]

sentit en un determinat context. La idea de subespecificació parteix de la idea bàsica de l'existència de la polisèmia sistemàtica:⁹²

Systematically polysemous nouns (...) occur with different, but related interpretations, throughout a coherent discourse. The noun 'school', for instance, can refer in the same discourse to *group of people, learning process, an institution* and *a building*. This noun will thus be underspecified for any of these related senses, unless in a local part of the discourse (a coherent discourse segment, a sentence or even a part of a sentence – a collocation) one sense is more prominent than any of the others. Still, even the other senses will be present, although on the 'background'. [(Buitelaar, 1998: 5)]

La subespecificació dels lexemes com a estratègia per a la descripció dels fenòmens lligats a la semàntica lèxica està present en els formalismes d'unificació, que s'hi basen per a explicar certs fenòmens de naturalesa semosintàctica —com l'ergativitat, per exemple, en què el canvi argumental és d'una predictibilitat evident—. En aquesta línia Pustejovsky desenvolupa el concepte de *type coercion*,⁹³ que determina, a partir de les estructures de *qualia*, la naturalesa dels canvis semàntics de què poden ser objecte els lexemes:

In the verb phrase *begin a novel*, the verb *begin* expects a phrase whose semantic type is an event. Because the NP *the book* does not satisfy this type, the verb coerces the NP into an event denotation, one which is available from the head's own qualia structure. Thus, formally, each qualia aspect is a partial function from noun denotations into one of their subconstituents (...). The verb *begin* therefore, can be said to semantically select for an argument of type **event**, instead of requiring three syntactic subcategorization frames. [Pustejovsky (1995b: 86)]

⁹² Aquesta idea es pot veure il·lustrada amb un exemple com *Un llibre d'astronomia amb unes cobertes vermelles*. Els dos modificadors de *llibre* fan referència a dues lectures semànticament diferents del mot *llibre*, el llibre com a objecte cultural té com a tema l'astronomia, i en tant que objecte físic podem dir que té les cobertes vermelles. Per tant, la descripció semàntica de *llibre* està en aquest sentit subespecificada fins que no la sotmetem a una determinada anàlisi contextual.

⁹³ Pustejovsky (1995a) defineix la *type coercion* com "A semantic operation that converts an argument to the type which is expected by a function, where it would otherwise result in a type error". [Pustejovsky (1995a: 59)]

La *type coercion*, doncs, permet interpretar, en funció de les informacions aportades per l'estructura de *qualia*, alguns tipus de predicació que no concorden exactament amb el tipus denotatiu sobre el qual s'han de realitzar. Així, per exemple, si *començar a* espera com a complement un predicat esdevenimental en forma de subordinada d'infinitiu, l'estructura de *qualia* de *llibre* (que incorpora la característica tèlica *per ser llegit*) modifica el tipus bàsic d'aquest substantiu (que és el d'objecte) i el converteix en un predicat esdevenimental que pot ser complement de *començar*. Les estructures de *qualia* proporcionen la informació necessària per a la modificació del tipus de denotació en funció de les característiques tèliques de *llibre*, que denoten el procés que exigeix el verb *començar*.

Aplicada al cas de la modificació adjectival, es pot entendre que és l'estructura de *qualia* dels substantius la que determina sobre quin element de la seva estructura semàntica es poden projectar les modificacions. Es pot il·lustrar aquest comportament a través de l'article del DIEC corresponent a l'adjectiu *fort*:

fort -a *adj.* Que té una gran potència per a obrar o resistir físicament, moralment; potent a produir un efecte, a accomplir un resultat; capaç de fer un gran efecte, una pregona impressió; que ateny un alt grau d'intensitat, de concentració; que té poderosos mitjans. *Els bous són forts per al treball. Un home fort com un roure. Tenir els braços forts. Tenir una constitució forta. Ha crescut sa i fort. Una voluntat forta. Una roca forta. Una plaça forta. Tenir la veu forta. Ésser fort en matemàtiques. Un argument fort. Un fort motiu de greuge. Un àcid fort. Un cafè fort. Una llum forta. Una olor, una pudor, forta. Un cop fort. Un vent fort. Una passió forta. Vi, licor, fort, que té una gran proporció d'alcohol. Un banc fort. Un exèrcit fort. Un fort adversari.* [DIEC]

La multiplicitat de definicions en una sola accepció d'aquest article permet suposar, d'entrada, la dificultat que presenta el tractament lexicogràfic d'aquest tipus de polisèmia. El conjunt d'exemples d'aquest article, complementàriament, presenten diferents tipus de modificacions que l'adjectiu *fort* pot realitzar en funció de les característiques semàntiques del substantiu a què s'aplica. En lexicografia són força corrents aquests tipus de solucions (que no deixen de ser una mena de tractament

inexplícit del fenomen) a fi de no prodigar en excés el nombre d'accepcions per a aquest tipus de modificacions semàntiques que poden ser percebudes com a regulars o sistemàtiques. En termes de la teoria de Pustejovsky, diríem que aquesta dispersió d'exemples es deu al fet que aquest adjectiu pot modificar selectivament diferents papers que formen part de l'estructura de *qualia* dels substantius a què es refereixen.

Finalment, i en la línia de dotar de solucions interpretatives aquest tipus de fenòmens, Bosque (1999) assenyala que les FL de la MTT (veg. més enrere § 3.3), i més específicament les funcions REAL i OPER, que tenen a veure amb la designació de verbs d'activació ($Real_1(\text{problema}) = \text{resoldre}$; $Real_2(\text{consell}) = \text{seguir}$; $Oper_1(\text{pregunta}) = \text{plantejar}$; $Oper_2(\text{consell}) = \text{rebre}$) poden ser una via de tractament per a aquestes pressuposicions que es presenten sobre els verbs sintàcticament absents, i en especial les relacionades amb la modificació adjectival. Així, *un problema fàcil* s'interpreta com *un problema fàcil de resoldre*.

No és la meua intenció d'aprofundir en aquesta observació, però val la pena de remarcar de passada el fet que el sistema de les FL no passa de ser, aplicat a aquest cas, un formalisme de representació i que resten pendents de resolució les múltiples qüestions que es relacionen amb les modificacions selectives, que no es projecten de manera indistinta sobre els valors d'aquestes FL (una pregunta, prototípicament, no és difícil de plantejar, sinó de respondre). Bosque (1999: 267) conclou que la relació entre *resoldre* i *problema* és de naturalesa lèxica i que “la forma en que se establezca esta relación en el léxico no es tan importante como la decisión sobre si debe establecerse o no”.

En escometre la qüestió de quin tractament lexicogràfic pot tenir el fenomen de la sincategorematicitat dels lexemes cal que, d'entrada, es delimiti la naturalesa del problema a nivell lèxic. L'absència d'informació i les seves diferents estratègies de recuperació són un fenomen corrent en els enunciats lingüístics, però no totes les estratègies de recuperació tenen el nivell lèxic com a base. Bosque (1999) al·ludeix a aquest fet:

Por defecto, parece claro que la recuperación anafórica es la primera que el hablante intenta, puesto que es la que la sintaxis proporciona. La recuperación léxica es la que se obtiene de las informaciones, formalizadas mediante *Qualia*, funciones léxicas u otros procedimientos que proporcione un léxico enriquecido. La recuperación discursiva es la que tiene lugar cuando fallan las otras dos, es decir, la que estamos obligados a obtener de los factores pragmáticos que podamos extraer del contexto previo. [Bosque (1999: 269)]

Assumint el fet que, almenys en part, la qüestió de la informació absent té un nivell de tractament lèxic que ha estat objecte d'estudi per part de certs models teòrics, cal plantejar-se des del punt de vista de la lexicografia si els diccionaris poden recollir en les seves descripcions aquest tipus de fenòmens. Cal valorar, doncs, fins a quin punt la lexicografia pot integrar en la seva descripció els elements necessaris perquè les seves parafrasis de definició conservin la relació d'equivalència sincategoremàtica amb el definit. Una primera qüestió derivada d'aquesta assumpció és la determinació de quin és el marc de tractament adequat o, d'una manera més concreta, quina classe lèxica cal prendre com a referència per a la descripció de les qüestions analitzades en aquesta part del treball.

Ja he mostrat que el concepte de sincategorematicitat introduït per Bosque (1999) es defineix en relació als substantius, i que Pustejovsky (1995a) vincula també aquest fenomen al canvi de tipus que poden experimentar, en funció de l'estructura de *qualia*, els substantius que per aquest procediment s'adapten a les característiques de l'argument esperat pel verb. Ja he mostrat els punts de contacte existents entre l'estructura de *qualia* i l'enunciat de definició. Malgrat l'existència d'aquestes similituds, però, no es pot considerar genèricament que la definició dels diccionaris convencionals aportin les informacions necessàries per a l'explicitació d'aquests comportaments i per a la descripció de les regularitats que representen.

D'altra banda, les regularitats polisèmiques associades a aquests fenòmens no afecten únicament (o no afecten principalment) els substantius sinó també els verbs del tipus *començar (a)*, i els adjectius que, com *fort*, poden modificar selectivament (Soler, 2001) determinats substantius. Tenint en compte aquests fets, assajaré de donar una proposta

de solució descriptiva, dins de l'estratègia general de tractament de l'extrinsicitat en la definició, a partir dels enunciats de definició dels verbs aspectuals i dels adjectius que són susceptibles de modificar selectivament els substantius amb els quals es combinen.

Tot i que la *type coercion* es realitza sobre el substantiu que fa d'objecte en les construccions en verbs aspectuals, es pot determinar que el canvi semàntic que es produeix en el verb aspectual consisteix en l'alternança entre *començar a fer X* / *començar l'acció prototípica associada a X*. La constatació d'aquesta alternança es manifesta en el diccionari com dos sentits separats i sense cap relació estructural, si ens atenim a la descripció lexicogràfica de *començar* al DIEC:

començar v. tr. Fer la primera part o començament (d'una cosa). *Començar un edifici. Començar un poema.*

(...)

7 començar a fer una cosa [o **començar de fer una cosa**] Fer el primer acte o la primera part de l'acció de fer aquesta cosa. *Començà d'anar d'una banda a l'altra. La neu comença a caure. Comença a ploure, a nevar.*

Una solució d'aquest tipus pot estar avalada pel fet que una i altra manifestacions del verb tenen de fet estructures argumentals diferents (*començar a V_{inf}* / *començar N*). Però això no integra en la descripció que aquest és un cas de polisèmia sistemàtica i que, fins a cert punt, és desitjable que un diccionari sigui sensible a les regularitats semàntiques que es produeixen en el lèxic.

Sense alterar fonamentalment la paràfrasi de definició, es poden assajar solucions alternatives que permetin posar de manifest l'alternança que es produeix en la selecció dels arguments d'aquest verb. La necessitat d'una solució genèrica es fa evident a partir de l'existència de casos de sentits diversos com els de 5.8.

(5.8a) El paleta va començar la casa / el paleta va començar a construir la casa

(5.8b) El pintor va començar el quadre / el pintor va començar a pintar el quadre

(5.8c) El pilot va començar la cursa / el pilot va començar a córrer la cursa

També és evident que la solució descriptiva a aquesta alternança passa pel tractament com a elements extrínsecs dels diferents components de l'estructura argumental, i per l'admissió de la possibilitat de plantejar alternances en l'expressió de l'extrinsicitat:

començar – [N₁ V a Vinf₂]; [N₁ V N₂]; Fer la part inicial [de fer alguna cosa | de l'acció de fer alguna cosa]₂

En el cas dels adjectius de modificació selectiva, el tractament també passa per admetre la possibilitat de la presentació de l'element extrínsec com a alternant. La lexicografia convencional no sol tractar de cap manera especial aquestes alternances:

rodó -ona *adj.* De figura circular; que té la secció transversal circular; esfèric. [DIEC]

robust -a *adj.* Que té una constitució forta, vigorosa. *Un home robust.* | *Una complexió robusta.* || FIG. *Un enginy robust.* *Un estil robust.* *Una fe robusta.*

Com es pot veure en l'exemple precedent, el component semàntic del nom modificat per l'adjectiu forma part intrínseca de la definició. Això implica que la substitució no és possible quan aquest element es manifesta en l'enunciat lingüístic (*un objecte de forma rodona, un home de constitució robusta*) i que per tant la presència d'aquests elements en la definició té un caràcter *impropi* (veg. Rafel, 1989).

Aquest fet es podria resoldre de la següent manera:

rodó - [N₁ és A] [objecte | forma de l'objecte]₁ que té la secció transversal circular.

robust - [N₁ és A] [algú | constitució física d'algú]₁ vigorós.

Solucions d'aquest tipus permeten l'explicitació de les característiques dominants dels substantius que són modificades pels adjectius i, alhora, permeten, si considerem el diccionari com a estructura, explicitar les característiques associades als noms modificats respectant els principis fonamentals del mètode de descripció lexicogràfica i sense incórrer en la necessitat d'especificar informació impròpia en les definicions. A més, la possibilitat de fer tractaments sistemàtics sobre el contingut dels diccionaris permetria, en un diccionari estructurat d'acord amb aquests principis, l'anàlisi de les

diferents fórmules d'expressió a fi de detectar-ne les principals invariants: [*objecte, forma d'un objecte*] per als adjectius de forma, [*objecte, color d'un objecte*] per als adjectius de color, etc.; d'aquesta manera, l'obtenció de dades més precises sobre la naturalesa i l'abast dels fenòmens lligats a la sincategorematicitat aportaria noves dades de gran interès per a l'estudi del fenomen.

5 Condicions sobre la descripció lexicogràfica

5.1 Desviacions estructurals del vocabulari definidor: la circularitat

La metallengua amb què s'expressen les definicions d'un diccionari monolingüe està basada en la llengua natural que descriu. La correspondència que s'estableix entre les unitats lèxiques que usa la metallengua i les entrades de diccionari condiona una de les desviacions estructurals més importants que presenten els diccionaris convencionals: l'anomenada *circularitat*. La circularitat provoca que la projecció de les relacions semàntiques que es poden establir a partir de la vinculació de les definicions amb els definits no sigui analitzable en termes de jerarquies o de taxonomies lèxiques (Cruse, 1986). Des d'un punt de vista purament formal, es pot afirmar que la circularitat provoca en els diccionaris una pèrdua de capacitat descriptiva i informativa, que en general l'usuari soluciona amb el seu propi coneixement de la llengua. L'impacte *real* de la circularitat en l'alteració de les capacitats descriptives, però, no és sempre el mateix, sinó que varia en funció de la importància dels components de la definició afectats.

La circularitat està associada al mètode de descripció lexicogràfica. Els diccionaris convencionals no poden evitar el fenomen de la circularitat, ja que aquesta es produeix necessàriament per raó del seu sistema de descripció del significat, que condiona de manera inevitable l'existència d'algun grau de recurrència entre els lexemes i llurs definicions. És justament el caràcter inevitable de la circularitat allò que sol fer que els

seus efectes negatius siguin relativitzats per algun dels autors que han tractat aquest tema, com Béjoint (2000):

Lexicographers may have been excessively rigorous about non-circularity. (...) After all, a dictionary definition has to use words, and cannot be understood if the user does not know what the words mean; there is a fundamental circularity in the genre of the monolingual general-purpose dictionary. [Béjoint (2000: 203-204)]

A fi de determinar i mesurar l'extensió de les relacions circulars en un conjunt de mots, Marcus (1970) introdueix el concepte d'*índex de circularitat*. Segons Marcus, es pot associar a un mot a un índex de circularitat n , en què n és l'enter més petit tal que hi ha una sèrie x_0, x_1, \dots, x_n que té $n + 1$ termes amb les següents propietats:

- (i) $x_0 = a$
- (ii) x_{i+1} s'utilitza en la definició de i
- (iii) per a qualsevol i , $0 \leq i \leq n - 1$, hi ha un enter j tal que x_j s'utilitza en la definició de x_n

Aquesta definició formula matemàticament les relacions circulars del diccionari, que habitualment es troben en els tractats de lexicografia descrites com: $A \rightarrow B, B \rightarrow A$ (índex de circularitat = 1); $A \rightarrow B, B \rightarrow C, C \rightarrow A$ (índex de circularitat = 2); $A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow A$ (índex de circularitat = 3); etc.

Vegem un cas de circularitat a través de l'exemple següent:

facultat *f.* [...] || Poder, dret, de fer alguna cosa. [DIEC]
poder *m.* Facultat que posa en estat de fer alguna cosa; (...) [DIEC]

Posem que $x_0 = poder$ i $x_1 = facultat$; es compleix que x_1 ha estat utilitzat a la definició de x_0 i per tant l'índex de circularitat del mot *poder* (i, consegüentment, també de *facultat*) és igual a 1.

Una *facultat* es caracteritza, doncs, com un *poder* i un *poder* es caracteritza com una *facultat*. En la projecció de les relacions jeràrquiques del diccionari no podríem assignar a cap d'aquests dos elements un superior jeràrquic fora d'ells mateixos. Aquesta

situació forçaria la conclusió que aquests elements, i, per consegüent, tots els elements inclosos per aquests són indeterminables per unitats de rang més alt.

Convé aclarir, però, d'entrada, que tant la formulació de Marcus (1970) com les que sol recollir la lexicografia teòrica no són del tot precises si les apliquem als diccionaris i a la seva estructura. Les formulacions del tipus $A \rightarrow B$, $B \rightarrow A$ per a descriure la circularitat obvien el fet que, en un diccionari, els elements que apareixen, respectivament, a l'esquerra i a la dreta de la implicació no tenen la mateixa naturalesa. En una proposició del tipus $A \rightarrow B$, hom considera que A és una entrada de diccionari i B un element del vocabulari de definició. Quan s'inverteixen els termes de la proposició, doncs, s'altera la seva naturalesa en l'estructura del diccionari: en la proposició $B \rightarrow A$, B és una entrada i A és un element de la metallengua.

En contra del que pot semblar, aquesta distinció no és trivial: la circularitat en un diccionari no es produeix entre entrades i vocabulari de definició, sinó exclusivament entre elements de la metallengua de definició. No és l'aparició d'un mot B a la definició del mot A, i viceversa, allò que provoca l'existència de circularitat, sinó l'aparició del sentit 'b' del mot A a la definició del sentit 'a' del mot B.

En l'exemple de *poder* \leftrightarrow *facultat*, la circularitat no es produeix pel simple fet que aparegui *poder* a la definició de *facultat*, sinó pel fet que el significat 'facultat que posa en estat...' del mot *poder* surt a la definició del significat 'poder de fer alguna cosa' del mot *facultat*.

La circularitat, doncs, es produeix entre aquests dos sentits concrets, que convencionalment podríem anomenar **poder²** i **facultat⁴**, respectivament, i que de fet són unitats de la metallengua de definició i identifiquen cadascun una accepció concreta de l'entrada a què corresponen. Allò que provoca en aquest cas l'existència de circularitat és, doncs, l'existència de les següents relacions:

poder² \rightarrow facultat⁴

facultat⁴ \rightarrow poder²

Mentre que no podem parlar pròpiament de circularitat en el cas següent, ja que el sentits de *facultat* que apareixen a banda i banda de la implicació són diferents:

poder² → facultat⁵

facultat⁴ → poder²

Aquestes precisions sobre el concepte de circularitat han estat sempre obviades, segurament pel fet que l'absència de desambiguació semàntica en la metallengua de definició dels diccionaris convencionals afavoreix la identificació entre entrades de diccionari i elements del vocabulari definidor (anomenant-los tots dos, genèricament, mots). Aquesta formulació de circularitat resultaria inviable en un diccionari que tingués la seva metallengua de definició completament desambiguada, ja que les diferències entre unitats d'una i altra mena esdevindrien molt més evidents.

A més d'imprecisa en relació a la naturalesa dels diccionaris, la definició de Marcus (1970) es pot considerar sobregeneradora des del punt de vista de l'impacte que tenen els diferents casos de circularitat en l'estructura i en la coherència descriptiva dels diccionaris. Alhora, la definició de Marcus no recull determinats casos de recurrència de la base lèxica en el cas de les definicions formulàries, que ha estat vist (Landau, 1984) com un tipus particular de circularitat. Com a mostra de totes dues limitacions hom pot adduir els exemples següents:

metall m. Qualsevol dels elements químics situats a les primeres columnes de la classificació periòdica o en els períodes de transició, com l'alumini, el ferro o l'or, caracteritzats per llur brillantor (dita metàl·lica), pels alts valors de llur ductilitat, mal·leabilitat i conductivitat de la calor i de l'electricitat, per llur capacitat de formar òxids bàsics com l'oxigen i de substituir els ions hidrogen d'un àcid per a formar sals. [DIEC]

oxigen m. Element químic no metàl·lic, gas incolor, inodor, insípid, que es troba en estat lliure en l'atmosfera, de la qual constitueix el 21% en volum i és un dels components de l'aigua, dels òxids i dels oxoàcids i l'element més abundant de l'escorça terrestre (símbol, O; nombre atòmic, 8; pes atòmic, 16,00). [DIEC]

metàl·lic -a adj. De metall, pertanyent a un metall o als metalls, de la natura d'un metall. [DIEC]

L'aplicació de la definició de Marcus als casos precedents dóna com a resultat, si ens atenim a la determinació de l'índex de circularitat de *metall*, que en aquest cas el seu valor és 2. El substantiu *oxigen* surt a la definició de *metall*; la definició d'*oxigen* conté *metàl·lic*, i finalment la definició de *metàl·lic* torna a contenir *metall*. D'aquesta anàlisi se'n desprenen dues constatacions força clares des del punt de vista de la interpretació lexicogràfica del cas:

a) El grau d'importància d'*oxigen* en la caracterització semàntica de *metall* és molt baix, però la formulació de Marcus no introdueix cap element que discrimini els diferents components de la definició en relació al seu pes descriptiu. El resultat és que el recorregut per a la constatació d'aquesta circularitat s'inicia a partir d'un element que podem considerar no essencial per a la caracterització semàntica de *metall*.

b) La formulació de Marcus arriba a l'establiment de la circularitat de *metall*, però no té en compte la relació existent entre *metall* i *metàl·lic*. El fet que el segon estigui format sobre la base lèxica del primer, i que això es tracti en termes de definició formulària que inclou *metall* implica un tipus de recurrència molt pròxima a la circularitat (Landau, 1984), però que la formulació de Marcus no recull, pel fet de tractar-se de lexemes diferents.

A partir del concepte d'*índex de circularitat* associem als mots que formen part del vocabulari de definició d'un diccionari un valor numèric que indica la distància que presenta un recorregut circular d'aquest mot en un diccionari. Aquest recorregut, doncs, pot ser més o menys llarg en relació al nombre d'entrades que s'hi impliquen. Diversos autors han assenyalat que la circularitat només implica problemes de descripció si es produeix entre un nombre reduït d'articles (això és, si el mot d'una sèrie circular presenta un índex de circularitat baix). Com a conseqüència, també és freqüent

d’afirmar que aquest mateix fenomen és perfectament tolerable a condició que el nombre d’articles implicats sigui prou gran.⁹⁴

El criteri de l’extensió dels recorreguts circulars no sembla el més adequat per establir distincions que ajudin a caracteritzar els diferents tipus de circularitat en relació al seu grau d’importància com a factor de desviació de la descripció lexicogràfica.

Una primera constatació per a l’establiment de classes d’importància en les desviacions de circularitat sorgeix si considerem aquest problema en relació a les categories gramaticals que afecta. A partir d’aquest criteri es poden establir algunes distincions intuïtives que resulten útils per a discriminar-ne els diversos graus d’importància. Si partim de la distinció entre *mots gramaticals* (o *buits*) i *mots lèxics* (o *plens*), podem constatar que en els primers resulta freqüent la presència de circularitat, que en alguns casos pot arribar a ser d’índex 0.⁹⁵ El fet, però, que aquest tipus de mots siguin caracteritzats, en general, a partir de descripcions gramaticals no substituïbles fa que, malgrat que la definició de Marcus se’ls pot aplicar, convingui distingir-los a fi de no

⁹⁴ Dubois & Dubois (1971: 46) expressen aquesta idea: “Si la circularité est trop courte, elle soulève des critiques, mais elle est dans la nature même de cet énoncé”. Martin (1974: 62) assenyala que en lexicografia “tout l’art consiste à rendre le cercle le plus large possible” i que mentre que la circularitat del tipus $V_1 \rightarrow V_2 \rightarrow V_1$ “est répréhensible”, s’accepta $V_1 \rightarrow V_2 \rightarrow \dots V_n \rightarrow V_1$ com a “imposée par la nature même de l’entreprise lexicographique”. Werner (1982b: 280) hi coincideix: “El caso extremo de una definición circular se da cuando un semema de una unidad léxica *A* es caracterizado directamente por un semema de una unidad léxica *B*, mientras que el semema utilizado para la caracterización de esta unidad léxica *B*, a su vez, es caracterizado sustancialmente de manera directa por el semema de la unidad léxica *A*, caracterizado ya por él”. Béjoint (2000: 203) es manifesta en un sentit similar: “Circularity is inevitable (...). The only thing that lexicographer can do is avoid the simplest types $A = A$, or $A = B$ and $B = A$. The longer type $A = B$, $B = N$, and $N = A$ is unavoidable and probably does not cause any practical problems either.”

⁹⁵ Es pot veure un exemple d’això en l’article del DIEC corresponent a la preposició *de*:

de [*d’* davant de vocal o *h*] *prep.* Expressa relacions que denoten: 1 El punt *de* partida *d’*un moviment, *d’*un espai *de* temps, *d’*una transició, *d’*un mesurament, l’origen; allò *de* què és separat, privat, alliberat, etc. quelcom.

tenir-los en compte per a la valoració dels casos problemàtics de la circularitat en relació a l'estructura dels diccionaris.

Soler (1992a, 1992b) és un primer intent de caracterització dels tipus més problemàtics de circularitat, i s'hi defineix la *circularitat estructural* com la que afecta cadenes de descriptors. Des del punt de vista de l'establiment de jerarquies lèxiques a partir de definicions lexicogràfiques la circularitat es percep com una desviació estructural important en les definicions que Ilson (1987) anomena *analítiques* (també s'han anomenat *hiperonímiques*, *definicions per inclusió*,⁹⁶ etc.) quan el recorregut circular afecta els elements que corresponen a la caracterització sintàctica i semàntica del

⁹⁶ No m'ocuparé ara de detallar ni de caracteritzar la idea de definició per inclusió, que d'altra banda es pot entendre com a paradigmàtica en la definició lexicogràfica; es poden trobar detalls d'això a Rey-Debove (1971). Assenyalo només algunes particularitats de les definicions nominals no (o no directament) hiperonímiques en els termes en què ho fa Martin (1977b). Entre aquestes hi ha:

- a) les definicions sinonímiques (sintètiques): **droguista** *m.* droguer, droguera. [DIEC]
- b) les definicions metonímiques: **àbac** *f.* Part superior del capitell d'una columna (...). [DIEC]
- c) les definicions antonímiques: **innutrició** *f.* Manca de nutrició. [DIEC]
- d) les definicions per analogia: **maça** *m.* (...) || Mena d'aixa per a escorxar. [DIEC]

Tampoc no es poden considerar definicions per inclusió les que integren elements més pròpiament del metallenguatge de signe, per dir-ho en termes de Rey-Debove:

teta *f.* Terme infantívol per a designar la germana gran. [DIEC]

roburita *f.* Nom d'un explosiu. [DIEC]

Martin (1977b) proposa per a l'establiment de les seves jerarquies de descriptors un sistema de reducció de tots aquests tipus de definicions al model inclusiu, a partir dels següents criteris:

- a) Inscriure els sinònims en la cadena d'inclusions.
- b) No tenir en compte el quantificador de les definicions metonímiques.
- c) No tenir en compte la negació de les definicions antonímiques.
- d) Ignorar els elements d'analogia en les definicions.

Pel que fa a les definicions amb elements de metallengua de signe, Martin proposa de no tenir en compte aquests elements. Val a dir que més que no pas la taxonomització de diccionaris, l'objectiu final de Martin és la detecció dels elements de màxima recurrència dins del diccionari, com a mètode per arribar a un conjunt d'elements primitius de la descripció lexicogràfica.

definit. De l'anàlisi de casos que presenten aquest tipus de circularitat de descriptors es poden detectar, a grans trets, tres tendències que es poden caracteritzar com segueix:

1. La circularitat afecta diferentment les unitats del vocabulari de definició, en funció d'allò que Martin (1974) anomena els *nivells d'inclusió*⁹⁷ del diccionari. Les unitats dels nivells superiors (o numèricament inferiors, veg. nota 97) estan més exposades a l'aparició de circularitat que no pas les dels nivells inferiors, que presenten més especificitat semàntica.
2. La circularitat afecta diferentment els mots en funció de llur denotació; mentre que els mots amb una denotació concreta presenten circularitat amb més dificultat, en els mots de significat menys específic o més *abstracte*, la presència de circularitat és força freqüent.
3. La circularitat es presenta amb més freqüència en els casos de definició múltiple, i sobretot en els diccionaris que utilitzen habitualment aquesta pràctica.

Es pot il·lustrar la tendència número 1 amb el següent exemple:

amonal *m.* Qualsevol explosiu constituït per barreges de nitrat amònic, alumini polvoritzat, carboni i trinitrotoluen. [DIEC]

azotina *f.* Explosiu compost de nitrat sòdic, carbó, sofre i petroli. [DIEC]

bel·lita *f.* Explosiu compost de cinc parts de nitrat d'amoni i una de dinitrobenzè o trinitrobenzè més una curta quantitat de nitrat potàssic. [DIEC]

dinamita *f.* Explosiu que consisteix en una mescla de nitroglicerina i una substància porosa inerta, d'un gran poder rompedor. [DIEC]

haloxilina *f.* Explosiu fet amb carbó, nitre, ferrocianur potàssic i una matèria inerta. [DIEC]

⁹⁷ Martin (1974) explica aquesta mena de "saturació semàntica" amb els seus "nivells d'inclusió". El nivell 0 és el del vocabulari general de la llengua, el nivell 1 correspon al conjunt de mots genèrics utilitzats en les definicions de tots els mots del nivell 0, el nivell 2 correspon al conjunt de descriptors utilitzats en les definicions de tots els mots del nivell 1, i així successivament. Cal aclarir que el concepte de "nivell" de metallengua que utilitza Martin no és coincident amb el del meu treball

Els mots definits resten caracteritzats per un únic element genèric, *explosiu*, de manera que els podem assignar un mateix hiperònim i, per tant, una relació de cohiponímia entre ells. Notem que els set mots presos com a exemple presenten en general la característica de ser mots que no formen part del vocabulari de definició.⁹⁸ La probabilitat de trobar relacions de tipus circular disminueix a mesura que aprofundim en el nivell d'anàlisi, i les projeccions basades en la hiperonímia solen presentar organitzacions pròpiament jeràrquiques. En canvi, en els nivells superiors (*explosiu* està definit amb els descriptors 'agent' i 'cos', que al seu torn es defineixen com 'allò que té el poder de produir un efecte' i 'porció limitada de matèria') els recorreguts circulars entre els seus elements són molt més freqüents. Aquesta és una tendència ben caracteritzada pels treballs d'extracció automatitzada d'informació lèxica dels diccionaris que ha dut a terme la lexicografia computacional (veg. per exemple Martí & al., 1991; Fontenelle, 1992; Ide & Véronis, 1994).

La tendència que he assenyalat amb el número 2 es pot observar a través dels mots que no formen part de sistemes de classificació cultural ben definits. És, generalment, el cas dels substantius anomenats "abstractes", que solen presentar moltes recurrències sobre els mateixos descriptors. Examinem els exemples següents:

- aptitud** *f.* Disposició o capacitat natural o adquirida per a fer alguna cosa. [DIEC]
- disposició** *f.* (...) | Manera d'ésser que fa apte per a fer alguna cosa, aptitud. [DIEC]
- capacitat** *f.* (...) || Qualitat del qui està en estat de fer alguna cosa. [DIEC]
- qualitat** *f.* Allò que, posseït per alguna cosa, fa que aquesta sigui tal com és; qualsevol dels atributs o de les propietats que distingeixen les coses (...). [DIEC]
- propietat** *f.* Allò que és propi d'alguna cosa; qualitat peculiar o característica; atribut essencial. [DIEC]
- atribut** *m.* Qualitat o manera d'ésser considerada com a pròpia d'algú o d'alguna cosa. [DIEC]
- manera** *f.* Forma particular de fer alguna cosa, de procedir, d'ésser. [DIEC]
- forma** *f.* (...) || Manera d'existència, d'acció o de manifestació d'una mateixa cosa o substància. [DIEC]

⁹⁸ El mots *amonal*, *azotina*, *bel-lita*, i *haloxilina* no pertanyen a la metallengua de definició. El mot *dinamita*, en canvi, apareix en algunes definicions, però no s'utilitza com a descriptor.

En aquestes definicions es produeix un grau elevat de circularitat estructural. Aquest tipus de fenomen dóna lloc a una representació gràfica completament diferent a la descrita anteriorment, tal com es veu a la figura 5.1.

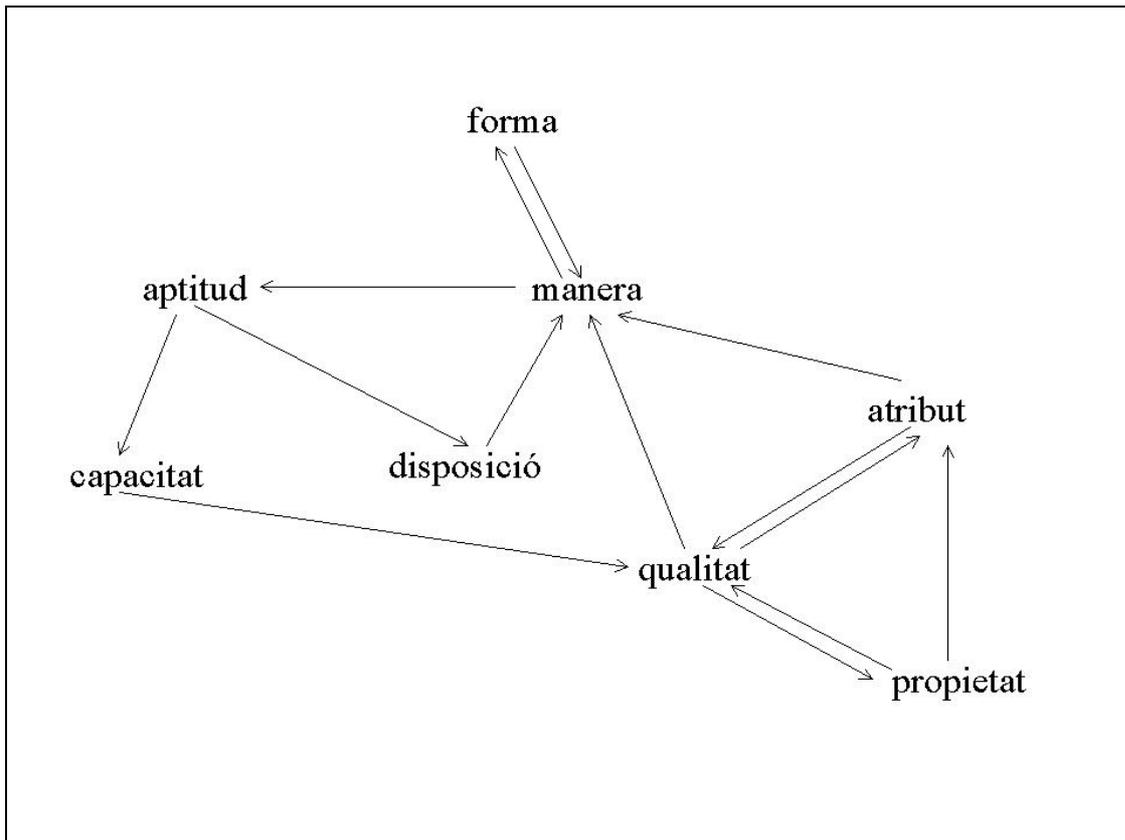


Figura 5.1. Representació de la circularitat en els substantius abstractes.

He restringit el nombre de casos involucrats en aquests diferents recorreguts circulars a fi d'atènyer-me al màxim possible a les definicions (o a les parts de definició) de tipus analític, i a més no he tingut en compte certs elements formularis com ara *conjunt de*, etc. com a descriptors. Cal tenir en compte, però, que si s'inclouessin en l'exploració els casos d'allò que Dubois & Dubois (1971: 59) anomenen “nomenclature sémantique

générique” que “est formée de termes comme *espèce, sorte, forme, qualité, action, resultat, se dit de, effet de, etc.*”,⁹⁹ les relacions circulars verificables hi serien en un nombre molt més alt (Soler, 1992b). Aquest tipus d'elements definidors són utilitzats amb certa profusió en la lexicografia convencional, en general sense que cap diccionari els identifiqui com a elements d'un valor metalingüístic especial.¹⁰⁰

Finalment, la tendència número 3 pot il·lustrar-se a partir de l'observació particular d'un fenomen corrent en l'exemple anterior. La vinculació múltiple de mots com ara *atribut* o *propietat* amb més d'un descriptor implica en general que la possibilitat d'identificar recorreguts circulars en sèries de mots sigui més gran.

La circularitat estructural (Soler 1992a, 1992b) és un concepte que cal sotmetre a revisió. En els termes en què l'he plantejada fins ara, la circularitat estructural presenta dues dificultats, l'una conseqüència de l'altra: els recorreguts circulars es manifesten

⁹⁹ Val la pena assenyalar que les relacions circulars entre aquest conjunt d'elements són força freqüents:

fet *m.* Acció de fer alguna cosa; allò que hom ha fet. [DIEC]

acció *f.* Manifestació d'una força, manera d'obrar d'una causa, acompliment d'una funció. [DIEC]

efecte *m.* Resultat de l'acció d'una causa. [DIEC]

resultat *m.* Allò que resulta. [DIEC]

estat *m.* Manera d'ésser, de trobar-se, en un moment donat, algú o alguna cosa. [DIEC]

¹⁰⁰ Com es pot veure a partir dels exemples que s'addueixen, en general aquest tipus de marcadors són elements que permeten el traspàs a una altra categoria gramatical (normalment verb o adjectiu, en el cas de les definicions nominals), del contingut semàntic de la definició. A més d'aquesta observació general, però, es pot veure l'existència de com a mínim dos tipus diferents d'elements: d'una banda, els que fan palès el caràcter “componencial” de l'estructura o les característiques morfosemàntiques del mot, com ara:

maldiença *f.* Acció de maldir; qualitat de maldient. [DIEC]

dramatització *f.* Acció de dramatitzar; l'efecte. [DIEC]

esclavatge *m.* Estat d'esclau. [DIEC]

i, d'altra banda, els que inclouen especificacions en molts casos accessòries i que no es troben a la part “genèrica” de la definició:

organogen -ògena *adj.* Dit de la roca que té un origen orgànic, i més pròpiament la formada pel creixement dels propis organismes. [DIEC]

maça *f.* (...) Mena d'aixa per a escorxar. [DIEC]

exclusivament en mots d'una mateixa categoria sintàctica (noms, verbs, adjectius, etc.) i, com a conseqüència d'això, les estructures que es manifesten com a desviacions descriptives formen part de conjunts aïllats de la resta del diccionari. De més a més, el concepte de subordinació i la definició per inclusió o *hiperonímica* no és aplicable a totes les categories gramaticals, la qual cosa exclou d'aquest tipus d'anàlisi una part de la macroestructura del diccionari.¹⁰¹

Una definició lexicogràfica pot analitzar-se com un conjunt finit de constituents semàntics que mantenen amb el definit diversos tipus de relacions. Fins ara m'he ocupat fonamentalment de l'anomenada relació hiperonímica, que podríem transcriure (per als substantius) amb la fórmula *A és_un_tipus_de B*, on *B* és l'hiperònim d'*A* i *A* és l'hipònim de *B*; aquesta és la relació típica que sol mantenir el definit amb el *genus* o element caracteritzador de la definició analítica. Les relacions que presenten els diversos tipus de constituents de la definició són, però, de naturalesa molt més diversa i, considerades des del punt de vista estructural, poden determinar tipus de circularitat que introdueixen factors de desviació en una xarxa de relacions molt més complexa (Copestake, 1992).

No és la meua intenció d'entrar a fons en aquesta qüestió, ja que des del punt de vista formal no es diferencia fonamentalment del que hem vist fins ara, sinó només cridar l'atenció sobre la importància que pot tenir una minimització conscient de la circularitat no només en el descriptor i en les estructures jerarquitzables del diccionari, sinó també en la xarxa semàntica del diccionari considerada com un tot. El control de descriptors no basta per a assegurar la coherència descriptiva dels diccionaris, sinó que cal partir de

¹⁰¹ Rey-Debove representa la visió més restrictiva. Seguint la idea d'inclusió establerta per la lògica, restringeix la inclusió als substantius i parla en to crític de l'extensió d'aquest mètode a altres categories: "Les définitions de dictionnaires (...) montrent que le système de l'inclusion est utilisé au moins pour le nom, le verbe, l'adjectif qualificatif et l'adverbe. Mais pour le nom, ce système est fondamental, alors que pour ces autres mots il n'est qu'accessoire et en quelque sorte dérivé du model nominal." [Rey-Debove (1971: 23)]

la base que allò que hem de controlar és el conjunt de la metallengua de descripció i les implicacions que, en funció d'aquesta metallengua, s'estableixen entre les diferents definicions d'un diccionari.

En analitzar els diversos models de lèxic m'he referit, en relació al lèxic generatiu (veg. més enrere § 3.5), a la interpretació de les estructures lèxiques com un graf de relacions multidimensional. La vinculació entre la caracterització del significat que fa la definició lexicogràfica i els models teòrics com ara l'estructura de *qualia* pot prendre's com a referència d'allò que intento de mostrar amb aquest replantejament de la circularitat estructural.

Examinem la següent sèrie de definicions:

calor f. Manifestació de l'energia interna d'un sistema que es mesura per mitjà de la temperatura. (...) || Sensació que experimenta el cos animal quan rep calor de l'exterior o quan augmenta la seva temperatura interior a causa d'un estat febril (oposat a *fred*). [DIEC]

temperatura f. Magnitud termodinàmica que indica el grau de calor d'un cos mesurant-ne l'energia tèrmica en relació amb la d'un altre. [DIEC]

tèrmic -a adj. Relatiu o pertanyent a la calor o a l'energia calorífica. [DIEC]

calorífic -a adj. Que produeix calor, que escalfa; pertanyent a la calor. [DIEC]

termodinàmic -a adj. Produït per una força deguda a l'acció de la calor. [DIEC]

febril adj. Pertanyent a la febre; que indica la febre. [DIEC]

febre f. Elevació de la temperatura del cos que manifesta la reacció de l'organisme a algun procés patològic. [DIEC]

Els dos sentits adduïts de *calor* caracteritzen aquest substantiu com una *manifestació de l'energia* i com una *sensació*, respectivament. Les classes determinades per aquests elements de caracterització principal estan subespecificades, en relació a la mesura de la temperatura en el primer cas, i en relació a *calor*¹ i a l'augment de la temperatura interior causada per la febre. El constituent *temperatura* té, tant en *calor*¹ com en *calor*², una importància notable en la caracterització semàntica en el cas de les dues definicions de calor. *Temperatura* és caracteritzada en primer terme com una *magnitud termodinàmica*, especificat en el sentit que indica el grau de *calor* a partir de la mesura

de l'energia *tèrmica*, etc. El resultat que una anàlisi feta en aquests termes donaria es pot veure representat a la figura 5.2, que mostra les recurrències internes que presenta la sèrie de definicions adduïdes com a exemple. En aquest cas, però, les relacions que s'estableixen entre els mots, tot i que es representen a partir del mateix tipus de fletxes, indiquen tipus diversos de caracteritzacions del definit, de manera similar a com l'estructura de *qualia* d'un lexema especifica de manera diferenciada els diversos constituents semàntics.

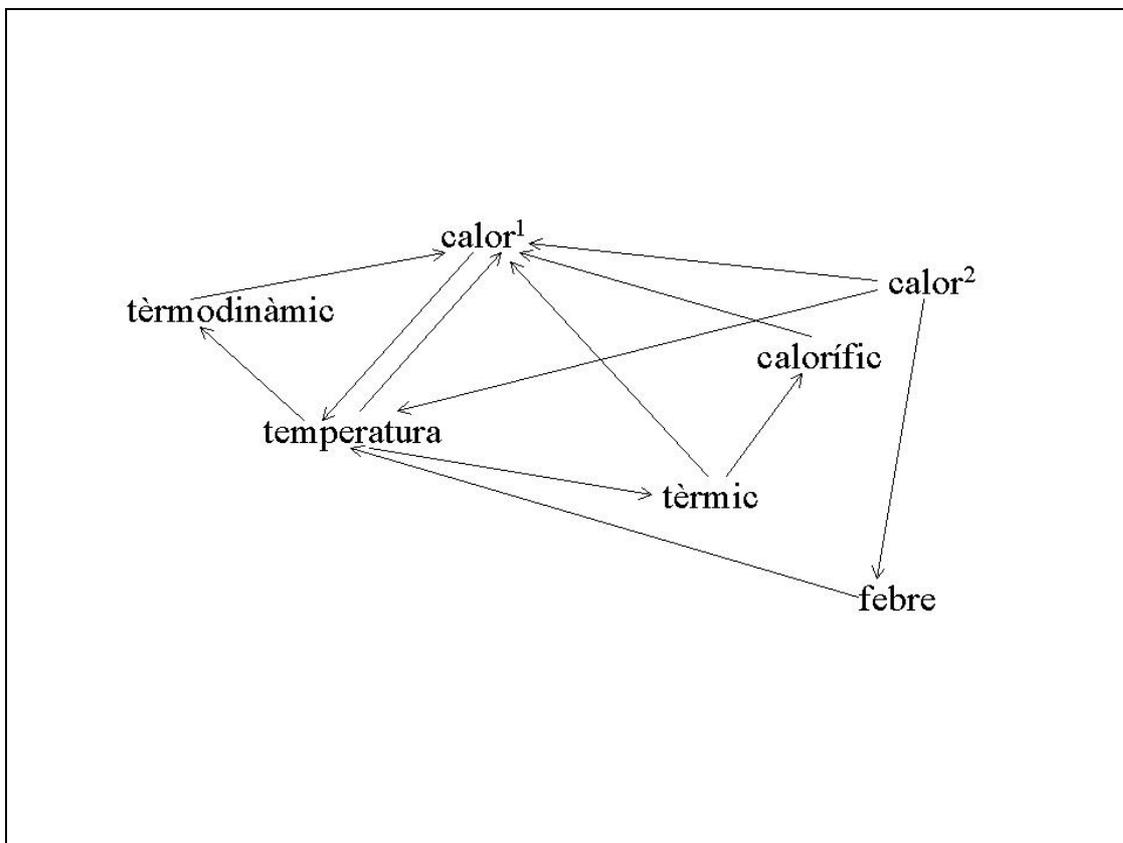


Figura 5.2. Representació de la circularitat en la sèrie relacionada amb *calor*.

El resultat de l'anàlisi que es mostra a la figura 5.2 posa de relleu l'existència de nombrosos recorreguts circulars en la sèrie de mots considerada, que afecten els dos sentits bàsics de *calor* de què he partit. Convé destacar el fet que, si aquesta anàlisi s'hagués fet considerant exclusivament els *descriptors* principals de cada definició, la

circularitat hauria estat nul·la. El tipus de circularitat que es produeix en aquests casos està, d'altra banda, fortament determinat per factors de *redundància* (que analitzo a § 5.2), però cal destacar el fet que aquesta circularitat es produeix entre elements que, malgrat que formen part dels elements diferenciadors del significat lèxic (això és, de les *differentia* en termes de definició analítica), desenvolupen un paper descriptiu important en la caracterització lexicogràfica dels mots de la sèrie.

Si partim de la base que els constituents semàntics d'una definició poden tenir graus d'importància diversos en la caracterització lexicogràfica d'un mot, es pot assajar d'ampliar el concepte de *circularitat estructural* a fi d'integrar-hi casos com el precedent. Amb aquesta integració poden superar-se les limitacions de descripció que representa circumscriure la circularitat a conjunts formats per una categoria gramatical homogènia i, per tant, s'estableixen els mecanismes adequats per a la consideració del diccionari com una sola estructura de múltiples dimensions, cosa que s'adequa molt més a la seva naturalesa pel que fa a les seves funcions de caracterització semàntica dels lexemes.

Considerada des d'aquest punt de vista més general, la circularitat estructural presenta aspectes de gran importància lexicogràfica i que no recull la caracterització formal de Marcus (1970); des del punt de vista formal el següent cas presenta circularitat:

recipient *adj.* (..) || m. Vas, esp. l'adaptat a una retorta per a rebre i condensar el producte de la destil·lació; vas per a rebre i contenir gasos. [DIEC]

gas *m.* Estat de la matèria en què les molècules que el componen resten poc lligades entre elles per les forces de cohesió i que sempre omple totalment i uniformement el recipient que el conté. [DIEC]

La segona de les definicions de *recipient* conté una especificació (que podríem denominar *tèlica*, en la terminologia de Pustejovsky) sobre la funció del definit: rebre i contenir gasos. El mot *recipient*, al seu torn, forma part de la caracterització semàntica de *gas*, però apareix formant part d'un constituent secundari, i que no estableix cap

relació semàntica directa amb el definit. Es tracta, doncs, d'un cas de circularitat no estructural i que no implica els problemes descriptius que he examinat anteriorment.

Aquest plantejament d'una visió menys restrictiva de la circularitat estructural satisfà, en certa manera, l'observació de Béjoint (2000) en el sentit que:

How should one judge a definition of *ear* (the organ of hearing) that uses the verb *hear*? Or of *mouth* that contains *eat*? (...) Such definitions are never condemned, but they also exhibit a sort of semantic circularity. [Béjoint (2000: 204)]

Cal considerar l'observació de Béjoint en relació a les recurrències entre sèries de definicions, molt pròximes a les que he presentat en el cas de *calor*. L'establiment de mecanismes per a la minimització de la circularitat, doncs, ha de tenir en compte la facilitat amb què es produeixen aquest tipus d'estructures als diccionaris convencionals. A partir de les consideracions precedents sobre la circularitat estructural, podem considerar que el fet que la circularitat es produeixi de manera necessària en la descripció lexicogràfica no justifica que no s'entengui com un problema de descripció. Complementàriament, la sola presència de circularitat en les definicions lexicogràfiques no invalida l'enunciat de definició com un sistema descriptivament adequat per a l'explicitació del significat lèxic, ja que, com he mostrat, en diversos casos la presència de circularitat no implica *necessàriament* cap desviació estructural notable del diccionari ni cap pèrdua de les seves capacitats descriptives.

5.2 Desviacions metodològiques del vocabulari definidor: la redundància

La utilització del llenguatge com a eina de descripció en els diccionaris tendeix també de manera habitual a la redundància; però a diferència de la circularitat, la redundància és evitable perquè no està associada necessàriament al mètode de descripció. Rafel

(1988) estudià de manera aprofundida el tractament lexicogràfic de les sèries sinonímiques i establí una proposta normalitzada per al seu tractament. L'estudi de Rafel revela que el tractament que sol fer la lexicografia convencional d'aquest tipus d'estructures implica sovint problemes de circularitat i de redundància. Però així com la circularitat està vinculada a l'estructura dels diccionaris i a l'ús del llenguatge com a eina de descripció, la redundància, en el sentit de la informació innecessària per a la caracterització lingüística del definit, representa de fet un problema que té una solució metodològica en la pròpia tècnica lexicogràfica.

La redundància lingüística és un fenomen complex que presenta multitud d'aspectes i que, amb graus diversos de presència, és una característica inherent a les llengües naturals, que la presenten en la majoria dels seus components d'una manera o d'una altra. Per exemple, la concordança de gènere i nombre entre l'article i el substantiu que determina, o el fenomen conegut com a *dobla negació* poden ser considerats casos de redundància gramatical del català i, per tant, un dels tipus de redundància lingüística.

Se sol dir, a més, que la redundància desenvolupa una funció bàsica en la comprensió, per part del receptor, de la informació vehiculada per l'emissor en un acte de comunicació produït amb enunciats lingüístics.

L'existència de redundància en els diccionaris és un problema que pot ser abordat des de perspectives molt diverses i que, a més, presenta una casuística rica i complexa que depassa les pretensions de l'anàlisi que vull portar a terme per a les finalitats d'aquest treball. En concret, em limitaré a examinar alguns aspectes de la redundància que presenten un impacte directe sobre la formulació de definicions lexicogràfiques.

En una definició lexicogràfica es troben elements redundants si algun dels seus components incorpora informació que pertany a la seva pròpia descripció en el mateix diccionari. Em refereixo, doncs, a la redundància provocada per l'aparició d'elements que, tot i que no són de naturalesa enciclopèdica, són superflus i, per tant, innecessaris per a la descripció. Rafel (1989) situa aquests elements dins dels elements *impropis* de

la definició. Atesa la naturalesa lingüística de la informació aportada per aquests elements, podem anomenar-los *redundants*.

Podem veure pràcticament a què em refereixo prenent com a exemple la següent sèrie de definicions del DIEC:

baixar v. *intr.* Anar de dalt a baix, d'un punt a un altre situat en un nivell més baix, davallar. *Baixar del quart pis al carrer. L'estel baixa: no té vent. El torrent baixa entre roques.* (...) || *tr.* Portar de dalt a baix, d'un punt a un altre situat en un nivell inferior. *Baixar la galleda al pou. Baixar les cadires al jardí.* [DIEC]

davallar v. *intr.* Anar, traslladar-se, moure's, d'un punt a un altre de més baix. *Davallarem de la muntanya a la plana. Un torrent que davalla de Puigcercós.* || *tr.* (...) Portar de dalt a baix. *Davallar el sagrat cos de la Creu.* [DIEC]

descendir v. *intr.* Baixar, davallar. *Descendir a l'infern.* [DIEC]

Una anàlisi superficial d'aquesta sèrie de definicions pot portar a consideracions de complexitat variable, algunes de les quals tenen molt a veure amb els problemes d'anàlisi de les sèries sinonímiques examinats per Rafel (1988).¹⁰² Aquí m'interessa, però, remarcar un altre fet que es produeix sobre la naturalesa enunciativa de les definicions i que està relacionat amb el grau de recurrència que presenten en les diverses definicions alguns dels components semàntics de la sèrie.

Les definicions de *baixar* i *davallar* incorporen, amb diferències no rellevants, l'element "d'un punt a un altre de més baix". Tots dos articles, a més, contenen una accepció intransitiva (introduïda per "anar") i una accepció transitiva (introduïda per "portar"). En aquests dos casos es fa evident la redundància que provoca la repetició de components de la definició que ja són presents en la descripció d'un o altre verbs, i la introducció de definicions repetides o equivalents en una forma que en principi està reservada a les definicions diferenciades.

¹⁰² *Davallar*, per exemple, apareix com a sinònim de *baixar* en la definició d'aquest, mentre que la que li correspon es pot considerar difícilment distingible de la definició de *baixar*.

Si ens fixem, però, en el tercer element de la sèrie d'articles (*descendir*) podem il·lustrar un tipus de redundància menys evident però igualment estès en la lexicografia convencional. *Descendir*, que d'acord amb el DIEC és sinònim de la sèrie {*baixar*, *davallar*, *descendir*} s'usa com a descriptor en la següent definició del DIEC:

abaixar v. tr. Fer descendir a un nivell més baix. *Abaixar el cap. Abaixar l'esquena. Abaixar els braços. Abaixar veles. Abaixar-se l'ala del capell. Abaixar-se les mànigues. Abaixar-se el coll de l'abric.* [DIEC]

En aquesta definició s'ha utilitzat “fer” com a marcador causatiu per fer correspondre la definició a la del verb transitiu *abaixar*. La presència del component “a un nivell més baix” resulta inequívocament redundant, ja que aquest component semàntic ja està incorporat a la descripció de *descendir*, que forma part de la mateixa sèrie sinònima que *abaixar* i *davallar*.

En tant que desviació metodològica, la redundància no té per què associar-se exclusivament a sèries sinònimes, ni tan sols a fenòmens de naturalesa semosintàctica com el que acabem d'examinar.

La redundància pot adquirir en el diccionari una casuística molt diversa. Gillette & Wit (1998) distingeixen dos grans tipus de redundància lingüística: la gramatical i la contextual; i encara, dins de la segona (que és la es presenta de forma més directa en els diccionaris) estableixen quatre categories diferents que representen tipus particulars de redundància:

a) Repetició sinònima: l'expressió conté dos o més components sinònims.¹⁰³

b) Repetició per aïllament: l'expressió conté dues o més subexpressions, de les quals una conté implícitament un o més trets o característiques de l'altra.¹⁰⁴

¹⁰³ “(...) the expression contains two (or more) identical or synonymous words or sub-expressions”. [Gillette & Wit (1998: §2.3)]

¹⁰⁴ “An isolating redundant expressions contains at least two subexpressions, of which one implicitly contains one or more features or characteristics of the other.” [Gillette & Wit (1998: §2.3)]

c) Repetició contrastiva: dos o més mots o expressions que constitueixen un contrast s'inclouen com una opcionalitat en una mateixa expressió.¹⁰⁵

d) Repetició distintiva: forma de repetició consistent en l'acumulació de característiques distintives d'un objecte.¹⁰⁶

Els quatre tipus de redundància descrits es donen amb freqüència en els diccionaris convencionals, ja sigui com a resultat “natural” del sistema de descripció que utilitzen o bé com a producte de certes desviacions metodològiques. Hi ha fenòmens de desviació prou coneguts, com la repetició de definicions (idèntiques o semànticament equivalents), que poden considerar-se casos particulars de redundància per repetició sinonímica; en aquest sentit, podem preguntar-nos quina justificació pot tenir la definició que es dóna a *giravoltar*, si la comparem amb la definició de *girar*:

girar v. *intr.* Un cos, moure's com una roda al voltant del seu eix; moure's circularment o espiralment al voltant d'un eix. [DIEC]

giravoltar v. *intr.* Moure's circularment o espiralment al voltant d'un eix. [DIEC]

Els tipus c) i d) no són tampoc estranys a la naturalesa de les definicions dels diccionaris convencionals, fórmules de definició com ara *amb X o sense*, etc., que es troben amb certa freqüència com a caracteritzadores del significat corresponen al tipus c). L'acumulació de característiques distintives¹⁰⁷ més enllà de la caracterització

¹⁰⁵ “(...) occurs when two (or more) words or expressions that semantically: constitute a contrast are repeated or in some other way redundantly coded.” [Gillette & Wit (1998: §2.3)]

¹⁰⁶ “(...) is a form of repetition of information in a context of differentiating one object from another” [Gillette & Wit (1998: §2.3)]

¹⁰⁷ Se'n pot veure un exemple en aquestes tres definicions:

mel f. Substància viscosa, dolça, elaborada per les abelles en una distensió de l'esòfag a partir del nèctar de les flors i que dipositen en les cavitats hexagonals de la bresca. [DIEC]

abella f. Insecte de l'ordre dels himenòpters (*Apis mellifica*) que viu en eixams i produeix mel i cera. [DIEC]

necessària i suficient del significat lèxic correspon al tipus *d*), que també és freqüent en la lexicografia convencional.

No em centraré, però, en aquests casos de redundància, sinó en el tipus *b*), que en el diccionari pot originar allò que podríem anomenar *redundància de components*. Seguint amb els exemples anteriors de *girar* i *giravoltar*, el component definidor *al voltant d'un eix*, apareix en totes dues definicions. Comparem aquestes definicions amb les següents:

roda *f.* Peça circular rígida de poc gruix relativament al seu radi, capaç de girar al voltant d'un eix, fix o giratori, que passa pel seu centre. [DIEC]

eix *m.* Peça cilíndrica, espiga, etc., al voltant de la qual giravolta una roda o que giravolta ensems amb la roda. [DIEC]

Tant la definició de *roda* com la d'*eix* permeten veure sengles casos de circularitat amb *girar* i *giravoltar* respectivament. La presència de circularitat pot implicar complementàriament l'existència d'aquest tipus de redundància de components.¹⁰⁸ La redundància, però, ve donada en tots dos casos per l'aparició de la fórmula “al voltant d'un eix”, que ja forma part de la descripció semàntica de *girar* i de *giravoltar*. Una de les possibles explicacions per a l'aparició d'aquests elements podria ser la necessitat de desambiguació dels verbs esmentats. Si examinem amb més detall la definició de *girar*, podem concloure que conté dues denotacions clarament diferenciables, i que es fonamenten respectivament en el fet que el cos que gira ho faci al voltant del seu eix

cera *f.* Substància groguenca, fàcilment fusible, combustible, que segreguen les abelles i amb la qual fan la bresca dels ruscós. [DIEC]

¹⁰⁸ De fet, Dubois & Dubois han volgut fer veure la circularitat com una conseqüència lògica del caràcter redundant dels diccionaris: “La redondance existe (...) d'un article à un autre d'un dictionnaire. Les définitions contiennent un certain nombre de termes qui constituent eux aussi des entrées du dictionnaire, car elles sont composées de mots de la langue. Et ces mots doivent être à leur tour définis. Par ailleurs, ils entrent dans la définition d'autres termes. Les définitions *se recouvrent* donc partiellement; le lexicographe est, en effet, conduit à utiliser le même stock de termes puisque ces derniers doivent être tous contenus dans la nomenclature du dictionnaire. Les définitions sont donc naturellement *circulaires* puisque, composées de quelques termes dont il faut connaître la définition, elles se recourent et que tous ces termes sont eux-mêmes définis par des paraphrases synonymiques.” [Dubois & Dubois (1971: 45)]

intern (del seu centre) o al voltant d'un eix extern, situat fora de la seva forma física. *Giravoltar*, en canvi, conté només la segona d'aquestes dues denotacions.

La possible justificació de la resolució de l'ambigüitat queda sense raó de ser pel fet que la utilització que el DIEC fa, en aquests casos, de *girar* i de *giravoltar* no es correspon amb la diferenciació que el mateix diccionari estableix, ni en el cas de la definició de *roda* ni en el cas de la definició d'*eix*.

La redundància pot tenir, sobre cada diccionari, efectes diferents en funció del tipus de metallengua de definició que utilitzi. En general, un diccionari amb una metallengua de definició no controlada a partir de cap criteri de restricció presentarà un nivell de redundància més elevat que un diccionari fet amb un vocabulari de definició restringit. Per contra, un diccionari de vocabulari restringit presentarà, per a una mateixa sèrie de mots analitzada, trajectòries circulars més estretes, ja que afectaran un nombre reduït d'elements. Per constatar aquesta hipòtesi podem prendre una sèrie d'entrades del DIEC que representen, en termes generals, una estructura jeràrquica:

sentit *m.* Funció psicofisiològica per la qual l'home i els animals reben informacions externes o internes mitjançant un òrgan especialitzat. [DIEC]

gust *m.* Sentit corporal amb què es perceben els sabors. [DIEC]

oïda *f.* Sentit amb què es perceben els sons. [DIEC]

olfacte *m.* Sentit amb què es perceben les olors. [DIEC]

vista *f.* (...); sentit, òrgan, de la visió. [DIEC]

tacte *m.* Sentit amb què es percep la pressió o la tracció exercides sobre la pell o una mucosa i es coneix la forma i les dimensions dels objectes, les condicions de llur superfície: aspra o llisa, dura o tova, etc. [DIEC]

A grans trets, i fent abstracció de les diferències estructurals que presenta aquest conjunt de definicions, podem determinar que aquesta sèrie de mots presenta una estructura de descriptors com la que correspon a la figura 5.3.

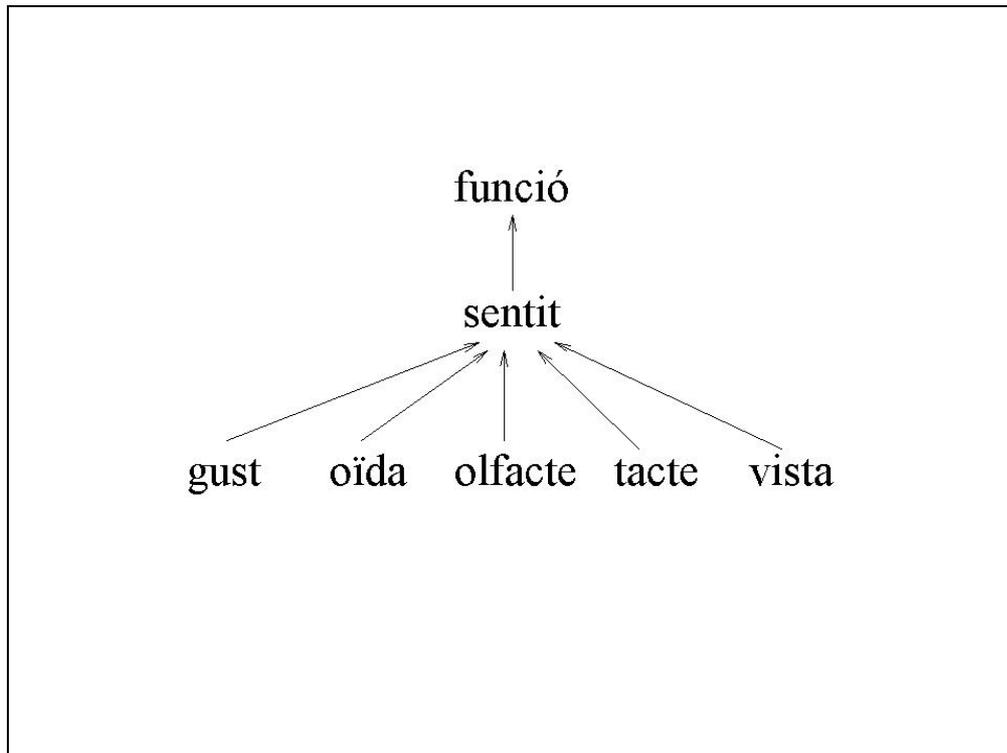


Figura 5.3. Jerarquia dels mots relacionats amb sentits corporals.

Sobre l'estructura de la figura 5.3 podem representar també les relacions que cada mot de la sèrie contrau amb altres entrades de diccionari, i fer la projecció de relacions semàntiques entre definicions. Les definicions que considero són:

òrgan *m.* Part d'un animal o d'una planta adaptada a l'acompliment d'una funció específica, com el cor, el fetge, l'estómac, etc. [DIEC]

percebre (...) *v. tr.* Rebre impressions, obtenir coneixença (de quelcom) per mitjà dels sentits. [DIEC]

impressió *f.* (...) Excitació fisiològica d'un sentit a part de la sensació desvetllada. [DIEC]

sensació *f.* Impressió que les coses causen en l'esperit per mitjà dels sentits; estat de consciència produït per un objecte extern o per un canvi en l'estat intern de l'organisme. [DIEC]

visió *f.* Funció del sentit de la vista. [DIEC]

so¹ *m.* (...) Sensació produïda en l'òrgan de l'oïda deguda a impulsions que ocorren regularment (oposat a *soroll*, degut a impulsions irregulars o confuses). [DIEC]

sabor *m. (o f.)* (...) Sensació percebuda pel sentit del gust. [DIEC]

redundància està motivada, principalment, per la presència d'elements desambiguadors com ara “sentit (de)” o “òrgan (de)” introduint elements de la definició que de fet ja es troben en la descripció lexicogràfica del component introduït (*gust, olfacte, etc.*).

Si aquestes definicions es reformulen a partir d'un vocabulari més restringit i que, complementàriament, estableixi un cert control sobre la possibilitat d'introducció d'elements redundants en les definicions la representació de les relacions internes d'aquesta informació pot canviar de manera sensible. A fi d'il·lustrar gràficament aquests canvis, simularé l'existència d'un diccionari d'aquest tipus per al català redactat a partir del vocabulari definidor que estableixo més endavant (§ 5.7, veg. Apèndix I.10). Es poden reformular les definicions examinades anteriorment de la següent manera (i veure-les representades gràficament a la figura 5.5:

sentit *m.* Funció per la qual es perceben sensacions mitjançant un òrgan especialitzat.

vista *f.* Sentit amb què es perceben les imatges.

oïda *f.* Sentit amb què es perceben els sons.

olfacte *m.* Sentit amb què es perceben les olors.

tacte *m.* Sentit amb què es perceben els contactes amb la pell.

gust *m.* Sentit amb què es perceben els sabors.

òrgan *m.* Part d'un ésser adaptada a l'acompliment d'una funció.

X percebre Y *v.* X Rebre impressions d'Y per mitjà dels sentits.

sensació *f.* Impressió que les coses causen en els sentits.

impressió *f.* Efecte que causa en l'organisme allò que és percebut pels sentits.

visió *f.* Funció de la vista.

olor *f.* Sensació produïda en l'olfacte.

sabor *m.* Sensació produïda en el gust.

so *m.* Impressió produïda en l'oïda per les vibracions elàstiques d'un cos que es propaguen en tots els medis materials en forma d'ones.

El fet que els mecanismes d'expressió de la llengua presentin graus diversos de redundància (gramatical o contextual) a fi de millorar, a través de la repetició, les possibilitats de comprensió comunicativa no implica, però, que un text tan altament estructurat com és un diccionari hagi d'incorporar la repetició en les seves descripcions. La redundància contextual no pot ser considerada, en si, com un avantatge descriptiu. Almenys en els casos que hem examinat fins ara, l'aportació informativa dels elements considerats redundants és nul·la i, en tot cas, correspon únicament a elements no necessaris que formen part del conjunt d'elements impropis. Si examinem aquestes dues definicions:

confessar *v. tr.* Reconèixer o declarar voluntàriament (els pecats) a un sacerdot a fi de rebre l'absolució. [DIEC]

penitència *f.* ESP. Sagrament en què són perdonats els pecats al pecador penedit que els confessa amb el propòsit d'esmena. [DIEC]

En la definició de *penitència* es pot observar la presència de diversos elements redundants, per referència a la definició de *confessar*. Si la redundància implica una millora de comprensibilitat sobre una fórmula del tipus: “sagrament en què són perdonats els pecats a qui els confessa”, per què no incorporar altres elements redundants com ara el *sacerdot* que també és argument de *confessar*? Mantenir que la redundància és una característica positiva de les definicions lexicogràfiques hauria d'implicar la formulació de mecanismes per a la seva expressió i, naturalment, fonamentar amb rigor on radica la millora aportada per l'aparició d'elements redundants.

Ja he mostrat com la redundància informativa en les definicions lexicogràfiques és un fenomen directament lligat a l'ús que el lexicògraf fa del vocabulari definidor disponible. Encara que la inclusió d'elements redundants és, en principi, independent del caràcter limitat o no del vocabulari definidor, sí que es pot afirmar que la inclusió d'elements redundants es pot veure afavorida si el nombre d'elements de vocabulari disponibles és alt. Conseqüentment, es pot afirmar que la limitació del vocabulari definidor pot contribuir al disseny d'estratègies per a la reducció de la redundància.

D'altra banda, està per demostrar que la redundància aporti claredat descriptiva a la lexicografia, mentre que és evident que no aporta cap millora estructural als diccionaris.

5.3 Un exemple de vocabulari definidor: els vocabularis controlats

La lexicografia anglosaxona, especialment la lexicografia orientada a l'aprenentatge de la llengua, ha introduït en els denominats diccionaris *EFL*¹⁰⁹ (“English as a foreign language”) una sèrie d'innovacions metodològiques destinades a millorar la comprensió dels diccionaris monolingües en els parlants no nadius. La més remarcable d'aquestes innovacions ha estat, sens dubte, l'adopció de conjunts tancats de vocabulari definidor, en diccionaris com el LDOCE i el CIDE.

La introducció d'aquesta estratègia per a la millora de la comprensió de les definicions ha posat de manifest de manera indirecta les implicacions del control de vocabulari definidor sobre l'estructura dels diccionaris (veg. § 5.4), i s'ha fet notar (Neubauer, 1987, 1989) que els diccionaris poden millorar, a partir de la delimitació apriorística del vocabulari de definició, la seva coherència estructural.¹¹⁰ Un dels efectes més notables assenyalats per Neubauer (1989) en aquest terreny és la minimització de la circularitat que, en relació a la totalitat del diccionari, suposa la utilització de vocabularis controlats. La reutilització d'aquests diccionaris per a l'adquisició automatitzada

¹⁰⁹ Gutiérrez (1999) utilitza la denominació genèrica DAL (“diccionarios para el aprendizaje de una lengua extranjera”) per a referir-se als diccionaris del castellà orientats a l'aprenentatge de la llengua. Mantinc al llarg d'aquest treball la denominació EFL ja que els diccionaris de l'anglès han estat els únics que han seguit l'estratègia d'utilitzar vocabularis de definició controlats de forma explícita.

¹¹⁰ Neubauer es refereix específicament a aquest aspecte: “It stands to reason that circularity in a dictionary cannot be totally avoided as long as all the words of a language are included in the dictionary, but lexicographers could at least try to avoid obviously circular definitions by monitoring the words used in the definitions even in those dictionaries where a defining vocabulary is not used”. [Neubauer (1989: 900)]

d'informació lèxica i per a la generació de bases de coneixement lèxic (veg. Boguraev & Briscoe, eds., 1989) ha mostrat, per comparació amb d'altres de vocabulari definidor no controlat, que la dispersió del vocabulari d'aquests darrers resulta un dels inconvenients més importants per al processament automatitzat de la informació lexicogràfica.

El problema de les desviacions estructurals en els diccionaris està força relacionat amb l'excés d'unitats presents en la metallengua de definició, és a dir, amb el conjunt de mots presents en la nomenclatura del diccionari que són utilitzats, al seu torn, en la redacció de les definicions. Aquest excés sol donar lloc a la introducció en els diccionaris d'elements que podem considerar, des del punt de vista formal, una minva de la capacitat descriptiva, que es produeix bàsicament en dos sentits diferents:

- a) La metallengua de definició recull sovint un nombre de termes que pot ser considerat excessiu, i que sol donar lloc a problemes de redundància informativa o agreujar les relacions circulars.
- b) Sovint un mateix lexema apareix utilitzat amb sentits diferents en les definicions lexicogràfiques sense que, en general, hi hagi cap mena d'explicació ni es consignï explícitament quin dels seus sentits s'utilitza en els casos d'ambigüitat.

Sembla clar que, com més variat és quantitativament i qualitativament el vocabulari definidor, més possibilitats hi haurà que els seus elements es vegin afectats per fenòmens de desviació de l'estructura de relacions (jeràrquiques i no jeràrquiques) i que, per tant, la coherència interna i el rigor metodològic del diccionari en sortiran perjudicats.

D'entrada, doncs, el fet de determinar explícitament quins són els elements que poden formar part del vocabulari definidor (amb la idea de reduir-lo) representa, si més no, una reducció de l'àmbit del problema que representen els fenòmens d'inadequació

descriptiva com ara la circularitat i la redundància.¹¹¹ Alhora, també es redueix la possibilitat de trobar duplicitats innecessàries,¹¹² etc.

El control del vocabulari definidor d'un diccionari, però, pot portar-se a terme amb graus diversos d'explicitud i rigor metodològic. En aquest sentit, Neubauer (1989: 900) distingeix entre les següents estratègies que defineixen diferents tipus de diccionaris d'acord amb aquest criteri:

- a) Absència total de control sobre el vocabulari utilitzat.
- b) Voluntat d'utilitzar un vocabulari senzill, sense explicitar-ne les restriccions.
- c) Utilització dels mots continguts en una llista específica, però també d'altres mots que no formen part de la llista en uns casos determinats.

¹¹¹ Svensén fa notar aquest valor afegit del vocabulari definidor: "The use of a defining vocabulary has the advantage that one can verify that a concept with a complex content is in fact being defined by means of less complex concepts. One can also define related concepts more consistently: the user can be sure that the concept *x* is always represented in definitions by the word *y*, and conversely that the concept *x* is meant whenever the word *y* is used. On the other hand, using a defining vocabulary does not guarantee easily intelligible definitions: in many cases it may be difficult to arrive at a natural form of expression". [Svensén (1993: 137)]

¹¹² Em refereixo a la possibilitat que en la metallengua de definició ocorrin dos sinònims. Tot i que no m'he referit explícitament a aquest fenomen, es pot il·lustrar a partir del cas de *profunditat*, *fondària* i *pregonesa* dins el DGLC. Tot i que *fondària* compta només amb la definició sintètica simple de 'profunditat', que també és un dels termes de la definició múltiple de *pregonesa*, en el diccionari podem trobar casos en què apareixen indistintament, com:

fregidora *f.* Recipient d'una certa fondària, proveït d'una escorredora, que serveix per a fregir-hi aliments. [DIEC]

escandallar *v. tr.* Amidar la profunditat (d'una capa d'aigua) amb l'escandall o un altre instrument. [DIEC]

escandall *m.* Instrument que consta d'un plom posat a l'extrem d'un cordill, que serveix per a determinar la pregonesa del mar, d'un riu, etc., en un punt determinat. [DIEC]

Tal vegada una aproximació interessant a aquesta qüestió fóra la suposició que això potser passa amb més freqüència en casos en els quals els dos pretesament sinònims no poden ocórrer indistintament amb el mateix significat en tots els contextos possibles.

d) Utilització exclusivament dels mots continguts en una llista específica, sense concretar el sentit en què s'utilitzen.

e) Utilització exclusivament dels mots continguts en una llista específica, només en els sentits especificats en aquesta llista.

És en el tipus e) on es fa més evident que els mètodes de control de vocabulari poden comportar un avenç qualitatiu important. A la pràctica, aquesta estratègia és aplicada amb diversos graus de rigor i explicitud. La utilització de cada unitat de la metallengua de definició amb un sentit específic i explícit, que en general es correspondrà amb el sentit més bàsic o usual,¹¹³ és, com mostro mes endavant (§ 8.5) una de les condicions per a una organització sistemàtica de la descripció semàntica.

Això ens introdueix al doble problema que representa la determinació de vocabularis controlats: en primer lloc, quines han de ser les seves unitats i, en segon lloc, quin és el sentit o sentits concrets del mot que cal actualitzar en les definicions. Si es vol que aquesta delimitació estigui motivada des del punt de vista empíric, sembla clar que cal fixar un i altre inventaris a partir de criteris més objectius que la mera introspecció del lingüista o del lexicògraf. D'aquí sorgeix la necessitat de disposar de dades provinents d'un corpus d'extensió suficient (això és, *representatiu*) i que, a més, aquestes dades siguin analitzades convenientment des dels punts de vista sintàctic i semàntic.¹¹⁴ En

¹¹³ En general se sol partir de la base que el caràcter bàsic o usual dels elements lèxics es pot determinar a partir de la seva freqüència en un corpus representatiu de la llengua. Aquesta assumció, tot i que és molt compartida entre els autors que s'han ocupat d'aquest tema, no està totalment exempta de certs problemes d'interpretació. En tot cas, la correspondència entre *bàsic* i *freqüent* cal matisar-la per referència al tipus de corpus de què es parteix per a l'establiment de la freqüència dels elements lèxics i també en funció de les necessitats que un vocabulari bàsic estigui destinat a cobrir.

¹¹⁴ El LDOCE és sensible a aquests condicionaments, tot i que no dóna gaires precisions i deixa d'explicar moltes qüestions que podrien resultar interessants: "The Defining Vocabulary was based originally on *A General Service List of English Words*, by Michael West, the only frequency list to take into account the frequency of *meanings* rather than the frequency of word forms. However, it has been updated with reference to more recent frequency information." [LDOCE: IX]

segon terme, val a dir que, en general, les propostes inicials de conjunts de vocabularis definidors reben molt de *feed-back* dels processos de redacció, de manera que no es pot determinar la seva compleció fins que el diccionari no es troba ja en fase final.

Les primeres temptatives (Cowie, 1999) d'establiment de conjunts d'aquest tipus amb finalitats aplicades a la lexicografia estan, com ja he dit, estretament vinculades a l'elaboració de materials per a aprenents estrangers.¹¹⁵ Cowie mostra la connexió existent entre els treballs en la línia d'establiment de conjunts de vocabulari bàsic per a l'anglès, portats a terme durant la primera meitat del segle XX, i els actuals diccionaris de tipus EFL:

The circumstances in which learners' dictionaries first appeared were exceptionally favourable, since not only were their editors –the founding fathers of EFL lexicography– (...), but all were contributors to, and beneficiaries of, innovative and fruitful programmes of research. Of these, none had a deeper or more widespread effect on the early history of the monolingual learners' dictionary than the so-called 'vocabulary control' movement of the 1920s and 1930s. In fact, it would be no exaggeration to say that vocabulary limitation gave birth to the learners' dictionary –and its parental influence endures to this day. [Cowie (1999: 3)]

El primer dels diccionaris, publicat el 1935, elaborat amb un vocabulari controlat fou el NMED. Les definicions d'aquest diccionari estan redactades a partir d'un vocabulari definidor de 1.490 mots, establert sobre criteris estadístics i semàntics.¹¹⁶ Amb posterioritat, West (1953) estableix una llista de vocabulari bàsic a partir de dades freqüencials explícites sobre un corpus de 5 milions d'ocurrències per a una part dels mots seleccionats (mentre que per a la resta el corpus fou de 2,5 milions d'ocurrències);

¹¹⁵ Dodd (1994) fa una visió històrica d'aquesta qüestió, i analitza les diferents propostes que s'han fet per a l'anglès en el segle XX. Alhora, també inclou una descripció de les principals obres lexicogràfiques que es basen en vocabularis controlats.

¹¹⁶ Em refereixo, per exemple, a la necessitat que West (autor del vocabulari) va remarcar d'incloure sèries de mots relacionades semànticament: segons la idea de West, si un vocabulari inclou el verb *menjar*, ha d'incloure també el verb *beure* i els substantius *gana* i *set*.

per a cada mot, West estableix la freqüència relativa (expressada en percentuals) de cadascun dels seus sentits. Els treballs de West en el terreny de la selecció de vocabulari bàsic¹¹⁷ foren els que formaren el nucli fonamental del vocabulari definidor dels diccionaris com el LDOCE i els que es publicaren posteriorment a partir de metodologies de descripció similars (OALD i CIDE).

Una de les obres lexicogràfiques més emblemàtiques constituïdes sobre un vocabulari controlat és el LDOCE, que es publicà l'any 1978. Per a l'explicitació de les definicions lexicogràfiques, el LDOCE s'autoimposa un vocabulari basat essencialment en el vocabulari de West, que consisteix en un conjunt d'aproximadament 2.000 mots d'ús corrent en anglès estàndard, i que els autors de l'obra anomenen *Longman Defining Vocabulary*. Un altre diccionari produït amb un vocabulari controlat és el CIDE, que explicita amb claredat els principis de selecció¹¹⁸ i que, a més, fa un ús extensiu d'aquest conjunt fins al punt que és utilitzat no només en les definicions, sinó també en els quadres gramaticals i en les notes d'ús.

¹¹⁷ West no és l'únic a treballar, durant la primera part del segle XX, en la determinació del vocabulari bàsic de l'anglès. Charles K. Ogden creà el BASIC (*British American Scientific International Commercial*) English, com una mena de selecció lèxica de 850 mots de l'anglès corrent que consistia en una proposta de llengua artificial. Ogden publicà en 1940 *The General BASIC English Dictionary*, amb 20.000 entrades descrites segons aquest sistema. Ogden es basà en una estratègia lògica per a la determinació d'aquest conjunt: agrupació de sentits pròxims, minimització de diferències, etc. A partir d'aquests criteris Ogden revisa i afina la selecció estadística; per exemple: *quick*, *rapid* i *fast* podrien formar part d'una selecció basada en criteris estadístics, però la desestimació de llurs diferències semàntiques a fi de retenir només el seu sentit bàsic condueix a la incorporació d'un sol mot al vocabulari bàsic.

¹¹⁸ “The words in this list have been carefully chosen according to these principles:

- contains words useful for explaining other words
- uses common words of high frequency
- uses words which have the same meaning in British and American English
- is easy for learners to understand
- avoids old-fashioned words
- avoids words which are confused with other words in English
- avoids words which are often confused with foreign words” [CIDE: 1702]

Tant el LDOCE com el CIDE parteixen d'un vocabulari controlat, d'uns 2.000 mots, que és essencialment el mateix. Les diferències fonamentals es troben en la supressió d'alguns substantius concrets, en la major especificació semàntica de la llista del segon diccionari, i en la incorporació en aquest de tots els mots derivats utilitzats en el vocabulari, tal com es pot veure a la figura 5.6, que recull una mostra comparativa de tots dos diccionaris referida als mots que es troben entre *cage* i *card*.

LDOCE	CIDE
cage <i>n</i>	cage
cake <i>n</i>	cake (FOOD)
calculate	calculate
	calculation
calculator	calculator
call	call (NAME)
	(TELEPHONE)
	caller
	(SHOUT/CRY)
calm <i>adj</i>	calm
	calmly
	calmness
camera	camera
camp <i>n, v</i>	
can <i>n, v</i>	can (ABILITY)
	(PERMISSION)
	(REQUEST)
	(POSSIBILITY)
candle	candle
	cannot
cap <i>n</i>	
capital <i>n</i>	capital (CITY)
	capital (LETTER)
	capitalize
	capitalization
captain <i>n</i>	
car	car
card <i>n</i>	card (STIFF PAPER)
	(GAME)

Figura 5.6. Mostra comparativa entre els vocabularis definidors del LDOCE i del CIDE

En l'aspecte quantitatiu cal remarcar, però, que els mots que usa el vocabulari definidor del LDOCE poden incorporar diverses categories gramaticals per a un mateix element, amb la qual cosa la llista excedeix lleugerament els 2.000 elements. Si tenim en compte les diferents categories gramaticals com a diferents realitzacions lèxiques, la llista en el LDOCE és de 2.170 elements. Una altra via per a la incorporació de noves unitats són aquells casos en què la flexió adquireix rellevància lèxica: *beside(s)*, *indoor(s)*, *provision(s)*, etc.; i la doble utilització dels numerals com a cardinals i ordinals: *four(th)*, *hundred(th)*, etc. A més, alguns dels processos¹¹⁹ de derivació més productius poden ser aplicats als elements del vocabulari, de manera que la quantitat de mots disponibles per al lexicògraf és força més alta.¹²⁰

A diferència del LDOCE, el CIDE utilitza una llista tancada en què tots els elements del vocabulari (obtinguts o no per derivació) hi són presents; la sagnia de línia de la figura 5.6 determina si es tracta d'un element derivat. La totalitat d'elements primaris és de 1.879 mots, i la de secundaris que s'hi relacionen és de 2.129 (comptant-hi tant els derivats com els diferents sentits especificats per a cada entrada de la llista). Això fa un total de 4.008 unitats del vocabulari definidor.

Una de les objeccions que s'han fet a la utilització de vocabularis controlats és que aquest sistema comporta en molts casos la formulació de definicions expressades de forma *poc natural*. S'ha objectat, també, que els vocabularis controlats no són adequats

¹¹⁹ Aquests processos es donen en forma d'una llista d'afixos que poden combinar-se amb entrades del vocabulari definidor. La llista és la següent: *-able, -al, -an, -ance, -ar, -ate, -ation, dis-, -ed, -en, -ence, -er, -ery, -ess, -ful, -ible, -ic, -ical, im-, in-, -ing, -ion, ir-, -ish, -ist, -ity, -ive, -ization, -ize, -less, -like, -ly, -ment, mid-, mis-, -ness, non-, -or, -ous, re-, -ry, self-, -th, un-, -ward(s), -work, -y*. Les propietats combinatòries de cada element de la llista es descriuen a l'apartat sobre formació de mots (LDOCE: B8), i la seves propietats semàntiques es donen en llistes específiques d'entrades: *word beginnings* (LDOCE: B7-B10), i *word endings* (LDOCE: B10-B14).

¹²⁰ No sembla haver-hi dades publicades sobre l'extensió *real* del vocabulari definidor de diccionaris com l'LDOCE. Les apreciacions que hi ha no són gaire exactes; Herbst (1986) situa aquesta xifra entre 5.000 i 10.000 mots.

per al tractament de determinats tipus de mots. Kirkpatrick (1985) es manifesta en aquest sentit:¹²¹

Like so many other things the defining vocabulary scheme is fine in theory, not so in practice. The worst feature of the defining vocabulary system is the constraints it places on the definer. A restricted vocabulary system definitely runs the risk of producing definitions phrased in a language which sounds most unnatural (...) The other problem is that it is virtually impossible to define some words, especially technical words, in terms of a limited vocabulary list. [Kirkpatrick (1985: 10)]

Herbst (1986) i Neubauer (1989) responen a aquesta crítica aportant una comparativa d'entrades de diccionari redactades a partir de diferents tipus de vocabularis de definició. A parer dels autors, queda demostrada l'adequació descriptiva dels vocabularis controlats i la *naturalitat* que es pot aconseguir amb la utilització d'aquest component de descripció en els diccionaris monolingües.

La manca de naturalitat en les definicions, en tot cas, no pot ser adduïda com una característica inherent als vocabularis controlats sense prendre en consideració el *sistema de definició* (veg. més endavant § 6.3) en què es produeixen els enunciats de definició. Comparem els següents exemples dels dos diccionaris a què ens hem referit fins ara:

rack 'In the past, **the** rack was a device to which people were tied and which stretched their bodies by pulling their arms in one direction and their legs in the other direction, usually used as a way of getting information from them'. [CIDE]

rack '[*the*+S] an instrument used to TORTURE people (=to cause them great pain) by stretching their bodies. [LDOCE]

¹²¹ A aquestes constatacions crítiques sobre els vocabularis definidors es pot afegir la de Michiels & Noël (1984) que, a partir d'una anàlisi comparativa d'un petit conjunt de dades procedents, respectivament, d'un diccionari amb vocabulari controlat (el LDOCE) i un de vocabulari *lliure* (el *Collins Concise English Dictionary*), extreuen conclusions parcials basades en el fet que tots dos diccionaris no difereixen essencialment en els seus problemes descriptius, especialment en l'assignació de *genus* i en les fórmules de definició.

En el primer cas, la rigorosa limitació al conjunt de definició, i la total renúncia a la incorporació de cap tipus d'element extern, condueix a una definició de poca capacitat evocadora; això es deu, en part, a una redacció no gaire afortunada, que omet alguns dels elements referencials relacionats més directament amb la tortura (el fet de *causar dolor*, per exemple); però també es deu a la dificultat que, dins del vocabulari definidor, té el lexicògraf per a descriure semànticament el referent i situar-lo en relació al camp d'activitat de què forma part. Així, recorre a fórmules com ara “usually used as a way of getting information from it”, que no recull en absolut els elements situacionals que són pertinents per a situar aquest objecte entre els aparells de tortura.

El sistema del LDOCE és diferent. En determinats casos com aquest, recorre a la inclusió de mots externs al vocabulari definidor (acompanyats d'una glossa¹²² de la definició corresponent) que són difícilment parafrasejables dins de la definició. Aquest sistema és utilitzat per a la referències de sinònims, mots en algun tipus de relació d'oposició, o per a la referenciació de mots que pertanyen a una mateixa sèrie lèxica. A part d'aquest mot, la definició del LDOCE no aporta nous elements en relació a la del CIDE, però és fonamental la circumstanciació introduïda per la preposició *by*, que introdueix “stretching their bodies”, ja que recull perfectament la relació existent entre l'aparell (*rack*) i la seva finalitat (*torture*).

Els vocabularis controlats també s'associen freqüentment a diccionaris escolars per a infants. Landau (1984) aborda la qüestió dels vocabularis controlats a partir d'aquesta perspectiva i es mostra molt crític amb el mètode:

The theory underlying school dictionaries, begun by Thorndike's analysis of the frequency of word use, is that dictionaries for children should be written in words that children can understand. (...) This may appear to be a truism, but it is not. One learns new words by

¹²² Aquesta glossa no reproduïx exactament l'enunciat de definició del mot:

torture (...) *v* [T] 1 to deliberately hurt someone to force them to tell you something, to punish them, or to be cruel. [LDOCE]

encountering them, either in speech or writing, and by making a shrewd guess, based on the context in which they occur and one's knowledge of similar words in similar contexts, as to what they mean. It is therefore not self-evident that a dictionary or any other book for children should avoid using words that challenge the reader to add a new word to his vocabulary. (...) If one criticism of school dictionaries is that they make some children reluctant to switch to adult dictionaries, another is that there are just too many school dictionaries for various grade levels. The second criticism has more merit than the first. [Landau (1984: 15-16)]

Ara bé, potser fóra bo de distingir l'objectiu de la intel·ligibilitat propi d'un diccionari escolar, que no implica necessàriament la utilització d'estratègies basades en el control estricte del vocabulari, de l'objectiu d'utilitzar aquestes estratègies per a millorar l'adequació descriptiva del treball lexicogràfic. De fet, un diccionari que s'adreça fonamentalment al públic en edat escolar s'ha de plantejar especialment, des d'un punt de vista més o menys rigorós, quin és el grau de variació del seu vocabulari definidor. En català el DIDAC constitueix un exemple d'aplicació d'algunes de les tècniques posades en circulació pel COBUILD i, tot i que no es planteja la necessitat d'un control explícit de vocabulari definidor, sí que declara en la seva introducció que les definicions han estat redactades amb la intenció de fer-les fàcilment comprensibles.

5.4 El vocabulari definidor i l'estructura dels diccionaris

Ja he dit a § 5.3 que la motivació principal dels vocabularis controlats no es basa en l'objectiu d'assolir una major coherència estructural del diccionari, sinó en raons d'ordre força més pràctic.¹²³ Sigui com sigui, és evident que la realització d'un

¹²³ No es pot oblidar que el LDOCE és un diccionari orientat bàsicament a un públic de *learners*, i que el seu desig és d'evitar les definicions que incorporin mots pretesament "difícils" per a aquest col·lectiu.

diccionari de vocació comercial amb l'aplicació sistemàtica d'aquesta mena de criteris de control representa, en termes estrictament metodològics, un avenç important pel que fa als mecanismes de minimització de les desviacions estructurals dels diccionaris d'ús corrent.

La delimitació de la metallengua de definició, però, tot i que contribueix a minimitzar alguns problemes que he assenyalat més enrere (veg. § 5.1 i § 5.2), no constitueix ni de bon tros una solució completa a la qüestió de la circularitat lexicogràfica, que es basa en principis inherents de la lexicografia d'orientació comercial. Sigui quina sigui l'extensió d'aquest vocabulari es podran establir relacions que, en algun punt, resultaran necessàriament circulars entre elements qualssevol de la metallengua.¹²⁴

De més a més, cal tenir en compte que el problema de la circularitat afecta de manera especial una gran part del lèxic més general de la llengua, que és justament el tipus de lèxic que amb tota probabilitat hauria de formar part d'un vocabulari controlat establert segons els principis que he mostrat. I encara, si el control del vocabulari definidor no va acompanyat d'un rigorós control o marcatge dels sentits utilitzats en les definicions, les incoherències estructurals del diccionari poden esdevenir encara més nombroses.

Della Summers, en la part introductòria del diccionari, és especialment clara pel que respecta a això: “Students of English graduating from a bilingual dictionary often have considerable difficulty in coping with a dictionary which gives explanations entirely in English. It is for this reason that Longman has developed the technique of writing definitions within a core vocabulary of 2000 base words” [LDOCE: VIII]. En aquest mateix sentit, veg. Neubauer (1987).

¹²⁴ Wierzbicka (1987: 5) comenta amb més o menys ironia aquest fet, referint-se precisament al LDOCE: “The authors could thus defend themselves against the charge of circularity as follows: «All right, we have defined *ask* in terms of *invite* and *invite* in terms of *ask*; we have defined *demand* in terms of *claim* and *claim* in terms of *demand*. These definitions are indeed circular and therefore they are not real definitions at all. However, the verbs in question (*ask*, *invite*, *demand*, *claim*, *request answer*, and so on) all belong to our defining vocabulary of 2000 words. We haven't really defined these 2000 because we can't define all words. But at least we have defined all the other words included in the dictionary (around 40000 or 50000) in terms of these basic 2000». [Wierzbicka (1987: 5)]

Sigui com sigui, el recurs a un vocabulari definidor pot resultar de gran interès en el cas dels diccionaris monolingües en general (i no només necessàriament en el cas dels dirigits a aprenents de la llengua). Entre els motius d'aquesta adequació es poden destacar:

1. Contribueix a fer definicions menys *cientifistes* (encara que sovint hagi estat atacat aquest sistema pel fet que produeix definicions massa *vagues* des del punt de vista conceptual, o que recorren a perífrasis excessivament complexes per referir-se a entitats que poden ser caracteritzades per un terme tècnic). Compareu els casos següents:

llavor *f* 1 *bot/agr* Embrió en estat de vida latent, acompanyat o no d'endosperma i de perisperma i protegit per l'episperma. [DLC]

seed [...] a usu. small hard object produced by most plants, from which a new plant of the same kind can grow, and which is used for planting. [LDOCE]

seed [...] a small esp. round or oval object which forms inside a fruit or the case of a vegetable such as a bean and from which, when it is planted, a new plant can grow. [CIDE]

Assumint com a certa la idea que l'evocació del referent constitueix una prova (si més no intuïtiva) d'adequació descriptiva en una definició lexicogràfica, és evident que la capacitat evocadora de la primera definició és escassa, almenys per a un parlant no especialitzat en botànica.

2. Organitza semànticament el diccionari i redueix els nivells de descripció semàntica: impedeix que mots molt específics formin part de les definicions i ajuda a fixar criteris estables en l'assignació de descriptors.

3. Facilita decisivament la construcció de bases de dades lèxiques a partir de diccionaris ja escrits o en procés de redacció, i la construcció d'eines per a assegurar la coherència de les obres lexicogràfiques durant el procés de redacció. De fet, la utilització de mitjans informàtics en les diferents fases de la redacció d'un diccionari ha permès o, si més no,

ha facilitat considerablement, la introducció de les tècniques basades en vocabularis definidors que han representat millores quantitatives i qualitatives.¹²⁵

Una part de les crítiques que s'han formulat en contra de la utilització de vocabularis controlats cal relacionar-les directament amb l'aplicació que es fa d'aquest conjunt dins del sistema de definició.¹²⁶

Indirectament, sembla també clar que la perícia del lexicògraf, com he mostrat anteriorment amb l'exemple de *rack* (veg. § 5.3), pot contribuir decisivament a aconseguir que una definició compleixi els requeriments d'adequació descriptiva que se li suposen.

¹²⁵ En especial, la utilització de configuracions informàtiques per a la fase de redacció ha possibilitat el control estricte de totes les informacions macroestructurals i microestructurals del diccionari, de manera que es poden garantir amb molta fiabilitat la coherència i la sistematicitat en l'aplicació dels criteris de redacció del mateix diccionari. Per a tenir una idea clara de la importància dels mitjans informàtics en els sistemes de control de vocabulari que tracto en aquest capítol, veg. Jansen & al. (1987).

¹²⁶ Des de la lexicografia pràctica, una de les crítiques internes més generals que s'han fet contra aquest sistema és la que exposa Hanks (1987) en la descripció dels criteris formulats per al COBUILD: “No attempt was made to set up a «restricted defining vocabulary» of a fixed number of words. Such vocabularies are a potential source of distortion, especially if they are not accompanied by equally strict controls on the meanings of each word used and the syntactic structures in which they are used. The fact that *stick* and *out* are in a learner's active vocabulary in one sense or another is no guarantee that that learner can understand *stick out* as a phrasal verb, while a reader who has access to the ordinary English word *beef* may wonder what distinction is being made when he or she reads that a steak is a piece of «meat from cattle».” [Hanks (1987: 1197)]

A la pràctica, el vocabulari de les definicions del COBUILD correspon plenament al d'un diccionari per a aprenents de la llengua, i no difereix essencialment d'un diccionari constituït amb vocabulari controlat, però selecciona els mots en funció de la familiaritat que presenten per als parlants no totalment competents. En l'exemple de *rack*, el COBUILD dona la següent definició:

The **rack** was a piece of equipment on which people were tortured in former times by being tied to it and then stretched, with their legs being pulled in one direction and their arms in the other.
[COBUILD]

Els mots *torture* i *former* no formen part dels vocabularis del LDOCE i del CIDE.

Hi ha, però, altres valoracions crítiques de l'ús de vocabularis controlats que tenen a veure amb qüestions més pròpiament estructurals del diccionari. Un vocabulari controlat no és per si sol un sistema de definició, i no implica necessàriament que els diccionaris construïts a partir d'aquests vocabularis s'alliberin totalment de les desviacions estructurals que són pròpies de la lexicografia convencional. Entre aquestes desviacions, la circularitat ocupa un lloc destacat. Al LDOCE es poden trobar força casos de circularitat estructural que afecten alguns elements de la metallengua de definició; naturalment, això es dona especialment en cert tipus de mots que es poden considerar de significat imprecís des del punt de vista jeràrquic:

feeling a belief or opinion, esp. one that is not based on reason. [LDOCE]

belief a feeling that someone or something is good or can be depended on; trust or confidence. [LDOCE]

trust firm belief in the honesty, goodness, worth, etc., of someone or something; confidence; faith. [LDOCE]

confidence a strong belief in the ability of a person, plan, etc., to do what is needed effectively and successfully.

faith firm belief; trust; complete confidence. [LDOCE]

Com ja he fet notar anteriorment, cal veure els vocabularis controlats com un factor que contribueix a minimitzar certs problemes estructurals del diccionari, més que no pas com una solució definitiva que permeti l'elaboració d'obres lexicogràfiques completament satisfactòries des del punt de vista de llur estructura. Per a la consideració d'aquests aspectes estructurals cal, més que no pas l'establiment apriorístic d'un vocabulari per a la redacció de les definicions, la formulació d'un sistema de definició complet que permeti l'eliminació o, almenys, la determinació clara del conjunt d'unitats afectades per elements de desviació estructural.

5.5 Condicions per a l'establiment d'un vocabulari definidor

Independentment de les limitacions que presenta la utilització de vocabularis controlats, la consecució de l'objectiu principal d'aquest treball exigeix la determinació *apriorística* d'un vocabulari definidor com a pas previ a l'establiment d'una metallengua de definició.

En principi, els únics experiments seriosos que s'han fet per a la determinació de vocabularis definidors es refereixen, com he mostrat a § 5.4, als vocabularis controlats establerts per a l'anglès. El fet que aquests vocabularis es refereixin exclusivament a diccionaris per a aprenents estrangers condiciona no només el tipus de discurs de definició, sinó també la selecció de les unitats que han de formar part del vocabulari.

La distinció entre els vocabularis anomenats *bàsics* i els vocabularis definidors dels diccionaris per a aprenents de la llengua és molt tènue. De fet, l'estratègia dels diccionaris EFL ha estat recórrer als vocabularis bàsics com a punt de partida per a l'establiment dels vocabularis controlats. L'establiment dels vocabularis bàsics es fa, en general, a partir de dades freqüencials sobre el lèxic per a garantir l'objectivitat en la selecció de les unitats. Com assenyala Sterkenburg (1989):

As an excellent objective criterion to help people attain a lexical proficiency level, and thus construct a basic vocabulary, frequency was and still is valid. Investigations into the basic vocabulary have been characterized by a firm belief in the objectivity of frequency countings. These frequency countings were regularly modified by reflections and gradations with regard to the grammatical words, spoken language, distribution, availability, and most of all by the differentiation of statistical methods of counting and calculation. [Sterkenburg (1989: 161)]

La familiaritat, que normalment s'enuncia com un dels principis de selecció dels vocabularis bàsics —i, per extensió, dels vocabularis definidors (Herbst, 1986;

Neubauer, 1989)— correspon a grans trets, des del punt de vista empíric, a la freqüència que els lexemes manifesten en un corpus general i representatiu de la llengua. En el cas dels vocabularis definidors, aquest principi de selecció s’ha de complementar amb la *utilitat* lexicogràfica que presenten els mots seleccionats com a bàsics. De tota manera, tal com afirma Neubauer (1989), la determinació d’un vocabulari definidor no es pot desvincular de la seva utilitat lexicogràfica :

it has to be stressed that the defining vocabulary is not something existing in the language waiting to be discovered by some empirical research. Rather it is a corpus of words that have shown their usefulness for defining other words in a dictionary. [Neubauer (1989: 903)]

Aquesta qüestió introdueix en la determinació d’un vocabulari definidor un problema de base: fins a quin punt es pot determinar aquest conjunt de forma totalment apriorística i prescindint de la mateixa tasca de redacció del diccionari? La qüestió no té una resposta senzilla, i no està exempta d’un cert caràcter circular.

D’una banda, els mots que presenten una freqüència més alta (i que, a grans trets, es poden considerar, en funció d’aquest criteri, més *bàsics*) no tenen per què formar un conjunt idoni per a la descripció lexicogràfica. D’una altra banda, els diccionaris construïts amb un vocabulari de definició “lliure” no poden, estrictament parlant, ser adoptats com un punt de referència exclusiu per a la constitució de vocabularis controlats. En darrer terme, l’establiment d’un vocabulari de definició no pot dur-se a terme de forma completament independent de la redacció del diccionari, que determinarà en última instància quines unitats del vocabulari cal incloure i quines cal desestimar.

Un cop assumida la conveniència i les limitacions de la determinació del vocabulari definidor, cal preguntar-se encara quin ha de ser el nombre d’unitats d’aquest conjunt i, sobretot, quina estratègia concreta cal seguir per al seu establiment. Pel que fa al nombre d’unitats, naturalment, és una qüestió que està en funció de la naturalesa, el tipus de nomenclatura i el públic a qui s’adreça el diccionari que l’hagi d’integrar. De

tota manera, es pot assumir l'existència d'un nucli bàsic de vocabulari general que podria ser compartit per diccionaris de tipus diferents. Com he mostrat anteriorment, els diccionaris anglesos de vocabulari controlat parteixen d'un conjunt d'unes 2.000 unitats, que en la pràctica s'amplia considerablement per l'admissió de derivats i sentits específics.

Per a les finalitats d'aquest treball, estableixo que el vocabulari definidor que ha de servir com a punt de partida estarà format per uns 3.000 mots. Aquesta xifra se situa per sobre de la dels vocabularis controlats i assegura, d'aquesta manera, la suficiència del conjunt obtingut com a vocabulari definidor.

Pel que fa a les estratègies de selecció, la qüestió tampoc no és senzilla. Encara que la determinació dels elements d'un vocabulari definidor és una opció subjectiva per un conjunt d'elements que representa una determinada manera de concebre la descripció semàntica del lèxic, es pot assenyalar que la introspecció (o el recurs a la intuïció) dels parlants competents d'una llengua no és suficient per a la determinació d'aquest conjunt d'elements. Calen sistemes més *objectius* que ens forneixin dades generals sobre el comportament dels lexemes a fi d'establir una selecció de base.

Cal tenir present, a més, que el conjunt d'elements que constitueix un vocabulari definidor ha de ser establert amb l'objectiu que cobreixi la finalitat bàsica de fornir els elements necessaris i suficients per a una descripció semàntica de tots els mots de la llengua.¹²⁷

¹²⁷ Svensén assenyala els requeriments d'un vocabulari definidor: "The construction of a general-language defining vocabulary is very demanding work. The words used must be representative at a certain level, and all concepts must be successively broken down into less complex ones. Then it is a matter of deciding where to stop. In principle, the process can be continued to the point where the content of the concepts is so simplified that they can no longer be expressed in words, but need some kind of logical symbolic language. Of course, this is not what is sought. The definitions are to consist of words; the question is how many are needed, and which they should be". [Svensén (1993: 137)]

L'existència de dades freqüencials sobre el comportament de les unitats lèxiques de la llengua, i la demostració empírica que un nombre reduït d'unitats concentra la major part de les ocurrences, ha portat sovint a la conclusió —Dodd (1994) n'és un exemple— que es poden establir els conjunts de vocabulari de definició a partir de repertoris de freqüències. Les dades que ens aporta el DFDCS corroboren aquesta tendència a la concentració, tal com es pot veure a la figura 5.7. Les dades freqüencials d'aquest quadre s'han calculat a partir de l'*ús*,¹²⁸ que el DFDCS recull com un tipus de freqüència modificada pel repartiment de cada lema en relació a les divisions tipològiques del CTILC (Rafel, 1996). Com més uniforme és el repartiment d'un lema, més acostat és el seu *ús* a la seva *freqüència*. L'*ús* de cada lema es calcula independentment en relació al corpus literari i al no literari del CTILC. Per a les dades d'aquest quadre, he fet la suma de tots dos valors per a cada lema i l'he anomenada *suma d'usos*. Els resultats es donen per a cada tram de 1.000 lemes ordenats decreixentment en funció de la suma d'usos, per als primers 10.000 lemes.

Una altra qüestió és, però, que la consideració de la freqüència —o, fins i tot, de l'*ús*, com a freqüència modificada— en un determinat corpus com a únic factor de selecció d'un vocabulari definidor sigui una opció metodològicament defensable. La naturalesa del corpus, la seva estructura, determina substancialment les informacions freqüencials que proporciona, i aquestes dades poden variar en funció de múltiples paràmetres, de

¹²⁸ Els conceptes de *dispersió* i *ús* (Muller, 1965) foren introduïts en l'estadística lèxica per Juilland i Chang-Rodriguez (1964). El concepte de dispersió es calcula a partir d'una fórmula matemàtica basada en la introducció del concepte previ de *freqüència teòrica* d'un lema en cada partició del corpus. Aquesta freqüència teòrica és la que caldria esperar com a real en un corpus ideal on cada lema estigués uniformement repartit. La diferència entre la *freqüència teòrica* i la *freqüència real* de cada lema en cada partició del corpus (és a dir la *variança*) és elevada al quadrat, i dona com a resultat un nombre positiu; l'arrel quadrada del qual, dividida pel nombre de particions del corpus dona com a resultat la *desviació tipus*. A partir d'aquí, s'obté el *coeficient de variació* dividint la variança per la freqüència teòrica. Aquest coeficient de variació es divideix per l'arrel quadrada del total de particions del corpus menys u, i dona com a resultat la *dispersió* d'una unitat lèxica. La dispersió és un valor entre 0 i 1 que, multiplicat per la freqüència absoluta d'un lema, ens dona com a resultat el seu *ús* dins del corpus considerat.

tipus quantitatiu (extensió del corpus, tractament de les dades) i de tipus qualitatiu (temàtics, cronològics, de mitjà, d'autor, etc.)¹²⁹

Rang (per suma d'usos)	Interval suma d'usos	Freq. acumulada	% Freq acum,
1-1000	3157141.68 3587.41	40211892	78.52%
1001-2000	3585.84 1739.06	3205325	84.78%
2001-3000	1738.25 1010,29	1745081	88.18%
3001-4000	1009.13 668.67	1119666	90.37%
4001-5000	668.35 474.11	787983	91.91%
5001-6000	473.94 363.71	587575	93.05%
6001-7000	363.69 288.98	463726	94.14%
7001-8000	288.97 233.59	380509	94.70%
8001-9000	233.64 191.12	312855	95.31%
9001-10000	191.09 159.25	263063	95.83%
		49077675	

Figura 5.7. Rang dels lemes del CTILC.

Per altra banda l'objectiu específic de la selecció de mots és l'establiment d'un vocabulari capaç de produir *totes* les perfrasis de definició d'un diccionari. Això condiona definitivament els requisits que ha de reunir el conjunt final. Neubauer (1987) introdueix la idea que només la interacció entre els diccionaris considerats com a corpus i, alhora, com a referències d'ajustaments de la selecció poden constituir elements de rellevància per a la millora metodològica de la lexicografia:

the best corpus for our statistical analysis of language for the purpose of dictionary improvement are the dictionaries themselves and this could make possible a kind of two-

¹²⁹ Per a una perspectiva general de com aquests factors, quantitatius i qualitius, poden influir en les llistes de freqüència, veg. Rafel (1996).

way interaction: on the one hand, the monolingual dictionaries can serve as corpora for the compilation of raw data, on the other hand, the resulting data could lead to a modification of the dictionary entries and therefore change the original corpus. [Neubauer (1987: 51)]

Neubauer també afirma que no es pot recórrer a l'extracció directa del vocabulari definidor a partir exclusivament de llistes de freqüència, sinó que cal modificar i adaptar a les necessitats de la lexicografia la informació que poden proporcionar-nos aquestes fonts.¹³⁰ Tal com assenyala Herbst (1986: 101): “A list based purely on frequency analyses would have hardly served the purpose” de l'establiment del vocabulari definidor del LDOCE, ja que aquestes llistes inclouen mots amb significats molt similars, i n'exclouen d'importants per a la definició d'altres mots.

Es pot admetre, doncs, que aquests tipus de *modificacions* sobre els criteris freqüencials vénen donades (o exigides) per les necessitats de la praxi lexicogràfica. Les inclusions o les supressions es determinen a partir de criteris aproximadament semàntics, que solen tenir relació amb les necessitats descriptives,¹³¹ amb les estructures paradigmàtiques que es produeixen en la semàntica dels mots seleccionats,¹³² o amb la proximitat semàntica dels lexemes.¹³³ En general, es pot dir, tal com assenyala Herbst (1986: 115), que la selecció de les unitats a incloure en el vocabulari definidor hauria de basar-se alhora en

¹³⁰ Tal com assenyala Neubauer: “Frequency counts are always dependent on the kind of corpus used for the research in question, and there is no guarantee that a high count for a particular item necessarily means that it is a good candidate for a Defining Vocabulary. Similarly, a word included in a basic vocabulary list for the purpose of language teaching may have been included because a tourist would encounter the word at the first sight”. [Neubauer (1987: 50)]

¹³¹ Aquestes necessitats comporten la inclusió de mots que, tot i que no formen part d'una selecció freqüencial, són requerits per a la descripció d'éssers, objectes o propietats del món real. Exemples d'aquest tipus de mots podrien ser: *circumferència* m, *receptacle* m, *cilíndric* adj., *elàstic* adj., etc.

¹³² Aquest criteri provocaria la inclusió de *set* si prèviament s'ha incorporat *gana*, i de *beure* si prèviament s'ha incorporat *menjar*.

¹³³ Aquest criteri determinà, en el cas del vocabulari de West, l'eliminació de *rapid* i *fast*, que tot i que tenen una freqüència elevada, són molt pròxims semànticament a *quick*.

els criteris de l'*expected familiarity* i el de *lexicographical usefulness*: “Frequency lists (and also analyses of the vocabulary introduced in text-books) will serve as important sources to determine what these items are”, però en determinats casos la selecció freqüencial s’ha de corregir en funció del criteri d’utilitat lexicogràfica.

La satisfacció completa del criteri d’utilitat lexicogràfica no es pot desvincular de la utilització efectiva del vocabulari en la redacció d’un diccionari, ja que aquesta utilització és el que en darrer terme determina, sobre la selecció freqüencial, la incorporació de tal o tal element no seleccionat en principi, o bé l’exclusió de tal o tal element innecessari per a la descripció lexicogràfica.

L’aproximació que duc a terme en aquest treball és diferent: la pertinença lexicogràfica no es determina al nivell del vocabulari sinó al nivell de la metallengua de definició, tal com mostro més endavant (veg. § 6.1). A partir d’aquest supòsit, crec que, un cop establert un vocabulari definidor sobre una selecció freqüencial, és factible d’arribar, amb l’aplicació de certes operacions de validació i de regularització, a la pertinença lexicogràfica de cadascun dels elements del conjunt obtingut.

5.6 Un vocabulari definidor per al català

Tenint en compte les consideracions analitzades en la secció precedent, cal ara descriure l’estratègia i els passos seguits, en aquest treball, en l’establiment d’una primera proposta per a la constitució d’un vocabulari definidor per al català. El primer objectiu a aconseguir és la determinació, en funció del grau de disponibilitat, dels materials de referència per a l’establiment previ del conjunt d’unitats candidates a formar part d’aquest vocabulari. En aquest sentit, es poden prendre com a referència les estratègies que s’han utilitzat per a la determinació de vocabularis controlats.

A partir de l'anàlisi de les condicions que ha de reunir un vocabulari definidor que he fet a § 5.5, es poden establir els següents requeriments que han de guiar la selecció de les unitats que n'han de formar part. Una proposta de vocabulari definidor per al català, establert sobre bases empíriques i no merament introspectives, hauria de tenir en compte elements de naturalesa diversa:

1. La freqüència d'aparició dels lexemes en un corpus suficientment gran i representatiu
2. La repartició d'aquestes unitats en el corpus
3. La freqüència d'aparició dels lexemes en el discurs lexicogràfic

A fi d'obtenir unes dades que tinguin en compte aquests tres factors condicionants he partit de dos corpus de dades diferents: el CTILC (a través dels seus resultats publicats al DFDC) i un corpus format per la totalitat de les definicions del DIEC (excepte les xifres numèriques, els noms taxonòmics, i les fórmules químiques que apareixen de vegades en les definicions). Les dades freqüencials estrictes (1 i 3) provenen del càlcul de la freqüència dels lexemes en aquestes dues fonts, mentre que el paràmetre de la repartició està implícitament recollit en les dades extretes del DFDC. La magnitud que en aquest diccionari expressa més pròpiament aquest paràmetre és l'anomenada *dispersió* (veg. Rafel, 1996; i també més enrere nota 128), que indica el grau d'uniformitat en la repartició de cadascun dels lemes del CTILC, en relació a les diferents divisions tipològiques del corpus.

El càlcul de l'índex de dispersió depèn de les divisions tipològiques del corpus; el CTILC presenta, per a cada lema, dos índexs de dispersió, corresponents respectivament a la distribució en el corpus no literari (dividit temàticament) i a la del corpus literari (dividit en gèneres literaris). A causa del fet que els dos corpus estan, doncs, organitzats a partir de diferents paràmetres, resulta incerta la consideració conjunta dels dos índexs de dispersió de forma combinada. Per aquest motiu he optat per recollir la repartició dels lemes seleccionats a partir de la suma d'*usos* (veg. més enrere § 4.5).

Per a portar a terme la prospecció de les unitats he tingut en compte, també, el text de les definicions del DIEC. Aquest text ha estat etiquetat (amb informació morfosintàctica i lematització) automàticament¹³⁴ d'acord amb els estàndards d'anotació morfosintàctica d'EAGLES que foren adaptats al català en el marc del projecte PAROLE. Malgrat l'existència de petites diferències entre els sistemes d'etiquetatge entre el corpus del CTILC i el del DIEC, els resultats d'una i altra font han pogut adaptar-se amb facilitat per establir les operacions de vinculació necessàries entre les dades del DFDCC i les dades freqüencials aportades pel corpus de les definicions del DIEC.

La prospecció d'aquests repertoris de dades m'ha permès extreure tres conjunts de base d'un nombre d'unitats similar (aproximadament 5.000 lemes):

1. Conjunt A. Lemes de freqüència més alta del DFDCC. (veg. Apèndix I.1).
2. Conjunt B. Lemes de suma d'usos (ús literari + ús no literari) més alta del DFDCC. (veg. Apèndix I.2).
3. Conjunt C. Lemes de freqüència més alta a les definicions del DIEC. (veg. Apèndix I.3).

El nombre de lemes corresponent a cadascun d'aquest grups varia lleugerament en funció de les característiques que reuneixen els primers 5.000 lemes de cadascun dels grups. En el cas del conjunt A, he seleccionat els lemes que presenten $F > 672$ (5.007 lemes en total); en el cas del conjunt B, he seleccionat els lemes que presenten una suma d'usos $> 474,10$ (5.000 lemes en total); en el cas del conjunt C he seleccionat els lemes que presenten $F > 15$ a les definicions del DIEC (5.190 lemes en total).

Prenent com a base aquests tres conjunts es pot establir la hipòtesi que la importància d'un mot depèn, a més de la seva freqüència, del fet que es trobi en dos o més conjunts. Per establir fins a quin punt pot existir coincidència o divergència entre els diferents

¹³⁴ Per a l'etiquetatge del corpus de definicions del DIEC s'ha utilitzat un etiquetador basat en tècniques de relaxació (*relaxation labelling*) amb restriccions ponderades de tipus estadístic (*bigrams*), extretes d'una mostra de 800.000 paraules del CTILC etiquetades i amb validació manual. L'eina d'etiquetatge dona una precisió del 95%.

conjunts, procedeix en dues fases: la primera referida a la comprovació de coincidències i especificitats en els conjunts del DFDCC, a través de l'establiment dels següents grups:

1. Conjunt $D = A \cap B$. Lemes comuns als conjunts A i B. (veg. Apèndix I.4)
2. Conjunt $D' = A \setminus B$. Lemes exclusius del conjunt A. (veg. Apèndix I.5)
3. Conjunt $D'' = B \setminus A$. Lemes exclusius del conjunt B. (veg. Apèndix I.6)

Un cop establerts aquests tres grups, comparo cadascuna de les llistes amb el llistat C, corresponent als lemes del DIEC, a fi de detectar-ne les coincidències, de manera que s'obtenen tres nous llistats:

1. Conjunt $E = C \cap D$. Lemes comuns a C i D (veg. Apèndix I.7)
2. Conjunt $E' = C \cap D'$. Lemes comuns a C i D'. (veg. Apèndix I.8)
3. Conjunt $E'' = C \cap D''$. Lemes comuns a C i D''. (veg. Apèndix I.9)

Els tres conjunts resultants d'aquesta operació tenen extensions molt desiguals: $E = 3.189$ lemes, $E' = 123$ lemes, i $E'' = 86$ lemes.

El conjunt seleccionat com a punt de partida per a l'establiment d'un vocabulari definidor es planteja com el conjunt format per $E \cup E' \cup E''$ (veg. Apèndix I.10). Aquest conjunt és equivalent a la reunió de les interseccions dels conjunts A i C, d'una banda, i B i C, de l'altra:

$$\mathbf{E \cup E' \cup E'' \equiv (A \cap C) \cup (B \cap C)}$$

El conjunt resultant està format pels 3.398 lemes que superen els filtres de selecció descrits anteriorment. El fet d'haver tingut en compte tant dades freqüencials d'un corpus de referència com del corpus aportat per les definicions lexicogràfiques del DIEC garanteix, alhora, el caràcter bàsic i l'adequació de la mostra per a finalitats de descripció lexicogràfica.

D'acord amb el que he dit a § 4.5, aquest vocabulari ha estat obtingut a partir d'uns procediments que no garanteixen la pertinença lexicogràfica de cadascun dels seus

elements. De tota manera, crec que el nombre d'unitats i l'estratègia de selecció assegurin que un subconjunt d'aquest vocabulari podria ser utilitzat com a vocabulari definidor en un projecte lexicogràfic. Els objectius d'aquest treball, però, van més enllà: el conjunt seleccionat serà considerat el punt de partida per a la determinació d'una metallengua de definició lexicogràfica integrada en un sistema de definició per a un diccionari monolingüe del català.

6 Vocabulari definidor i sistema de definició

6.1 D'un vocabulari definidor a una metallengua de definició

A § 5.6 m'he ocupat de l'establiment d'una llista de mots seleccionats a partir de criteris freqüencials. Tal com he mostrat, criteris d'aquest tipus són els que han servit de base per a la selecció dels vocabularis definidors que han estat utilitzats en la redacció de diversos diccionaris de l'anglès. Per tant, es pot partir de la base que, en principi, les unitats lèxiques seleccionades a § 5.6 poden constituir un conjunt apte per a l'establiment d'un vocabulari definidor. Cal, però, precisar diversos aspectes de la meua proposta:

1. De l'establiment d'aquesta llista no se segueix necessàriament que el conjunt seleccionat sigui adequat (en el sentit que cadascun dels elements que el componen sigui necessari, i que el conjunt de tots els elements sigui suficient) per a la descripció lexicogràfica d'un diccionari general.
2. Un vocabulari definidor s'integra en un *sistema* de definició que defineix els criteris generals de la descripció lexicogràfica. El mer establiment d'una llista de mots (encara que estigui, com en aquest cas, motivada des del punt de vista empíric) no representa necessàriament un avenç sobre les capacitats descriptives d'un diccionari ni sobre la seva coherència estructural.
3. En funció de les observacions precedents crec necessari, d'una banda, prendre el conjunt seleccionat com a punt de partida per a la determinació d'un vocabulari

definidor que satisfaci les condicions del punt 1; d'altra banda, crec també necessària l'explicitació d'una sèrie de condicions bàsiques en les quals s'integri el vocabulari definidor i que satisfaci les necessitats descriptives d'un diccionari general monolingüe.

La diferència que estableixo en aquest treball entre un vocabulari definidor i una metallengua de definició resideix en la integració d'aquest vocabulari en un sistema de definició. Un vocabulari de definició, independentment de la seva adequació i fonamentació empírica, és una llista de mots que poden conduir a descripcions diverses en funció del sistema de definició en què s'integri. Anomeno metallengua de definició, doncs, un vocabulari definidor integrat en un sistema explícit que determina els aspectes de l'aplicació d'aquest vocabulari en un projecte de descripció lexicogràfica o, si es vol, en un diccionari. La diferència entre vocabulari i metallengua de definició és, doncs, en certa manera, paral·lela a la que es pot establir entre lèxic i llengua. El lèxic no és per si sol una llengua si no està integrat en un sistema de funcionament gramatical precís.

El procés de redacció de cada diccionari determina els aspectes que fan referència a les seves eines de descripció; confereix una determinada *forma externa* a l'explicitació del significat lèxic a partir de les estratègies de tractament i de presentació de la informació. En funció d'aquestes tries, el resultat final pot variar entre un ventall més o menys ampli de possibilitats. No hi ha dubte que, entre totes les decisions sobre el sistema de descripció semàntica, l'establiment d'un vocabulari definidor és probablement allò que determina en un grau més alt la forma final del discurs lexicogràfic i, per tant, la manera com un diccionari explicita la descripció del significat. La mera determinació del vocabulari definidor d'un diccionari no basta, però, per a cobrir les necessitats que planteja el diccionari com a estructura, sinó que cal el concurs d'un sistema de definició que estableixi els criteris d'organització de les descripcions semàntiques.

Com ja he mostrat, l'existència de cercles viciosos és una característica inherent a la lexicografia convencional i al seu mateix mètode d'explicitació del significat a través d'enunciats lingüístics. De tota manera, la tendència dels diccionaris cap a la normalització metodològica ha de representar una sèrie d'avenços en la línia d'un major

rigor estructural dels diccionaris. Aquesta normalització exigeix l'anàlisi del discurs de definició en tant que metallengua. Ja m'he ocupat de delimitar la transcendència estructural que sobre els diccionaris poden tenir fenòmens com ara la circularitat o la redundància, i crec que es pot determinar que la coherència descriptiva d'un diccionari no pot basar-se exclusivament en el control de les jerarquies lèxiques, sinó en models de representació més complexos. En aquest sentit, crec que és necessari de plantejar-se la normalització de la descripció semàntica des del punt de vista de totes les relacions que pot contreure la nomenclatura del diccionari amb la metallengua de descripció.

L'objectiu de dotar els diccionaris d'un major rigor formal no implica, ni en la distinció de constituents (intrínsecs i extrínsecs) ni en la formulació d'enunciats definidors, la pèrdua d'intel·ligibilitat ni la renúncia al sistema descriptiu de la lexicografia basat en la paràfrasi. Aquestes condicions han estat socialment assumides com a fonamentals en el mètode lexicogràfic i, per tant, formaran part de la meua proposta de tractament.

L'objectiu de determinar una sèrie de criteris sobre la forma i l'estructura de les definicions a fi d'aconseguir un grau satisfactori de coherència en el discurs lexicogràfic ha d'adequar-se a dos eixos descriptius bàsics, que són aplicables a la definició lexicogràfica: l'eix paradigmàtic i l'eix sintagmàtic. L'eix paradigmàtic determina la selecció de les unitats de descripció, i pot identificar-se aquí amb el vocabulari definidor. Per altra banda, l'eix sintagmàtic regula la combinatòria d'aquests elements en la construcció d'expressions, i estableix les regles per a la construcció de descripcions adequades. La naturalesa d'aquests dos components d'una metallengua de definició és, doncs, fortament diferent i cadascun dels dos s'integra complementàriament en l'altre.

La determinació del primer d'aquests components depèn, en un diccionari monolingüe, de la llengua descrita, i requereix a més el concurs de dades empíriques sobre la freqüència dels lexemes; el segon, en canvi, pot ser enunciat amb un cert grau d'independència, ja que estableix, de fet, indicacions metodològiques que fan referència a l'estructura dels diccionaris en general. L'adaptació d'aquestes condicions generals a

una llengua concreta no es pot considerar un problema de rellevància teòrica, sinó que s'associa als projectes lexicogràfics concrets.

D'acord amb les consideracions precedents, portaré a terme (veg. més endavant § 7) la determinació de les unitats que conformen l'eix de descripció paradigmàtic a partir del vocabulari definidor proposat a § 5.6. En aquesta secció i les següents (§ 6.2 i § 6.3) m'ocupo de precisar el concepte paradigmàtic de *sistema de definició* i d'assenyalar les condicions que ha de complir tot sistema de definició lexicogràfica.

Un cop acceptada la idea que, més enllà de la mera delimitació d'un vocabulari definidor, cal determinar un sistema explícit que satisfaci les necessitats de la descripció lexicogràfica, convé dir d'entrada que els treballs que s'ocupen d'aquesta qüestió són escassos, i en general no transcendeixen el llindar de les observacions de tipus més general. Aquesta falta d'anàlisis més detallades coincideix amb el fet que els diccionaris no solen explicitar entre els seus criteris d'elaboració quines són les pautes generals del seu sistema de definició, encara que aquest pugui tenir un grau més o menys alt de formalització.

En qualsevol cas, un sistema de definició lexicogràfica estableix unes pautes de descripció que mai no incorporen totes les possibilitats de la gramàtica d'una llengua i dels seus diferents modes d'expressió. Wiegand (1999*b*), aplicant elements de la teoria textual a l'anàlisi de la definició, formula d'una manera sintètica la caracterització d'un sistema de descripció lexicogràfica:

A lexicographic description method is a set of rules to be applied in an orderly way, which lay down in a more or less strict fashion (depending on the type of dictionary) in which way a lemma sign of a specific class is to be described within the framework of the dictionary article. [Wiegand (1999*b*: 159)]

En la línia de les propostes prèvies a l'establiment d'un sistema de definició es pot

considerar la proposta de Weinreich (1962), que formula, a partir de la idea de controlar i caracteritzar la metallengua de definició,¹³⁵ un sistema de descripció lexicogràfica compost de diferents *estrats*, en què cada estrat representa el conjunt que formen els mots definits amb el conjunt de mots de l'estrat anterior. En el cas de l'estrat més bàsic (l'estrat 0) és clar que cal construir forçosament les definicions de manera circular o “ostensiva”.¹³⁶

Weinreich defineix, doncs, els seus estrats de la següent manera:

Stratum 0: terms definable only circularly and by ostensive definition.

Stratum 1: terms whose definitions contain only stratum-0 terms, but without circularity.

Stratum 2: terms whose definitions contain only stratum-0 and stratum-1 terms, without circularity.

Stratum *n*: terms whose definitions contain only terms of strata 0, 1, 2, ... *n* - 1.” [Weinreich (1962: 38)]

¹³⁵ Weinreich és el primer lingüista a posar un èmfasi especial en la idea de vocabulari definidor, i fa la primera formulació explícita dels problemes pràctics que comporta per a la lexicografia la utilització d'una metallengua de definició “incontrolada”: “Conventional lexicography apparently believes that the defining metalanguage should contain at least the entire object language. This belief may be unwarranted. Ideally we might wish for an «absolute» metalanguage which is entirely independent of the object language. But since this ideal in semantics is illusory (in contrast to phonetics), we should seek ways to make the metalanguage less rich, rather than as rich as, or richer than, the object language.” [Weinreich (1962: 37)]

¹³⁶ El terme *definició ostensiva* prové de la filosofia. Wittgenstein (veg. Brand, 1975) l'utilitza aplicat a les situacions en què hom defineix un mot a partir de l'acció de mostrar el seu referent en la realitat; per tant, es tracta de definicions que no expliquen el significat, tret que el significat es conegui prèviament. Per a una panoràmica dels sentits del terme en la filosofia lògica contemporània, veg. Robinson (1954). Pel que fa a l'ús que fa Weinreich (1962) del terme, sembla equiparable amb un tipus de definició deductiva, en què el significat es pot deduir de la presentació acumulada de mots que pertanyen a la classe del definit: “Orienting a map corresponds in semantic description to the demonstration of some denotata. This is also known as ‘ostensive definition’. To be effective, ostensive definitions must produce not only positive instances of denotata of a term, but also negative counter-instances (= positive instances of denotata of a synonymous term). Thus, a single swatch of red does not yield a sufficient ostensive definition of *red*; the limits of the meaning must be established by showing various kinds of red and also examples of what is similar to red but not red (e. g. orange, pink, reddish black). Denotata may be presented physically, or being named; thus, a series of colored swatches, or the terms *red*, *blue*, *yellow*, etc., constitute two varieties of ostensive definition of *color*.” [Weinreich (1962: 31)]

La proposta de Weinreich presenta nombrosos punts de contacte amb les que posteriorment elaboraren Martin (1974, 1977a i 1977b) o Calzolari (1977) (veg. més endavant § 6.2) però, al contrari d'aquestes, no està orientada a la detecció d'elements primaris a partir de l'anàlisi de diccionaris existents, sinó més aviat a la caracterització apriorística d'un sistema de definició.

Weinreich proposa, de fet, una mena de control de vocabulari de naturalesa diferent¹³⁷ al dels vocabularis controlats del LDOCE o el CIDE. Al contrari d'aquests, el sistema proposat per Weinreich no estableix un vocabulari que es basa en un conjunt tancat i predeterminat d'elements, sinó que posa les bases d'una metallengua indefinida per naturalesa, i que es configura en una sèrie de conjunts estratificats que es determinen en el procés de redacció del diccionari. Weinreich redueix els problemes de l'estructura semàntica dels diccionaris a l'anomenat estrat 0, que previsiblement estarà constituït per un conjunt reduït de mots. S'acosta, per tant, a la idea d'un conjunt de primitius semàntics que constitueixen la base de descripció lexicogràfica, amb la diferència metodològica important, però, que Weinreich assumeix la circularitat com un mètode lícit i necessari en la descripció lexicogràfica; i justifica aquesta posició a partir del fet que cal caracteritzar el significat d'un terme per contrast amb el dels termes que tenen un significat més pròxim:¹³⁸

The consistency of lexicography could be improved if dictionary makers were held to the assumption that the terms of a language are, on the whole, complementary (...). This assumption implies that the most important case to deal with in semantic description is one in which, where the signification of one term ends, that of another begins. On the whole, a semantic description should aim not at "absolute" definitions, but at definitions which

¹³⁷ Weinreich (1962: 37) però, considera també el mètode de control de vocabulari aplicat posteriorment per aquests diccionaris i, encara que el considera menys rigorós, l'assenyala com més viable.

¹³⁸ Aquest és un dels principis bàsics del model teòric en què es mou Weinreich, basat en la continuïtat i discontinuïtat semàntica, i que origina allò que ell mateix anomena *Continuity of the Definition System*.

delimit the meaning of a term from that of terms with similar meanings (synonyms). The circularity which results should be frankly admitted, not as a vice, but as a guiding principle of lexicography. [Weinreich, 1962: 30]

Weinreich utilitza el criteri freqüencial en el seu sistema estratificat com a corresponent a les nocions més imprecises sobre la naturalesa més o menys general, més o menys bàsica, d'un mot. En aquest aspecte, Weinreich segueix la línia d'identificació entre la freqüència i el caràcter més o menys bàsic d'un mot que he mostrat a § 5.5. Aquesta progressió del particular al més general en l'estructura dels diccionaris ha estat formulada molts cops des de punts de vista i finalitats diferents; i també es troba a la base del *principi de descomposició* de Mel'čuk (1988a).

Una altra de les estratègies proposades per Weinreich, més pròxima metodològicament als vocabularis controlats, consisteix en la utilització d'un conjunt reduït de mots disponibles per a la redacció de les definicions lexicogràfiques, assajant així d'arribar a un tipus de reducció similar a la del *Basic English* (veg. nota 117). En qualsevol cas, aquesta aproximació parteix de la base que probablement 800 mots o menys foren suficients per a la descripció completa del vocabulari d'una llengua, la qual cosa queda bastant lluny quantitativament del vocabulari definidor dels diccionaris redactats fins ara.

El model proposat per Weinreich presenta dificultats d'aplicació que el mateix autor reconeix i, de fet, no ha estat aplicat en cap projecte lexicogràfic.

Més en la línia de la descripció tècnica que no pas en la de l'aportació teòrica, Svensén (1993: 138) il·lustra el concepte de sistema de definició amb el model següent (aplicat en el *Svesnk ordbok*):

At the centre is a defining vocabulary (DV). The items in this are assumed to have known meanings. They are nevertheless defined in words, but these definitions are circular and the concepts involved cannot be decomposed into less complex units. The words in the DV are not necessarily the shortest, simplest and commonest words in the language. These are

often polysemous, vague in meaning, and less suitable for the purpose. [Svensén (1993: 138)]

Els mots que no pertanyen al DV han de ser definits, directament o indirectament, a partir dels mots que hi pertanyen. Els mots que no pertanyen al DV poden ser dividits en dues categories:

a) FDV (*fully defined vocabulary*), conjunt al qual pertanyen un nombre limitat d'unitats, que estan definides exhaustivament amb descripcions completes que descriuen semànticament i sintàcticament cada lexema. Cadascuna d'aquestes unitats pot ser descomposta successivament en elements del DV (sovint a través d'altres elements de FDV).

b) PV (*paraphrased vocabulary*), conjunt de mots caracteritzat per descripcions més aproximatives mitjançant paràfrasis de definició; els seus elements poden ser igualment descompostos en elements del DV.

Aquestes propostes, però, tot i que estableixen les bases d'un sistema es limiten a assenyalar l'organització interna de la metallengua, sense referències explícites a les condicions descriptives que han de verificar-se en les definicions. En l'àmbit de la caracterització global dels sistemes de definició, Benson & al. (1986) assenyalen 12 principis bàsics de la definició lexicogràfica. D'aquests 12 principis, només 6 fan referència directa a allò que jo en aquest treball anomeno definició lexicogràfica; de la resta, un fa referència a la utilització complementària dels sinònims, un altre a l'ordenació de sentits, un altre als exemples lexicogràfics, un altre als patrons de comportament sintàctic, un altre a la utilització de marques d'ús i, finalment, un altre al tractament dels significats connotatius. L'aportació de Benson & al. (1986: 205-226) pot resumir-se en els següents punts:

1. Definició referencial:¹³⁹ a partir de la mateixa part de l'oració; sempre que sigui possible, un mot hauria de ser referencialment definit a partir de la seva mateixa part de l'oració.
2. La funció del metallenguatge:¹⁴⁰ la utilització de les descripcions metalingüístiques ha de ser restringida al mínim indispensable.
3. Especificitat de les definicions: idealment, les definicions haurien de donar només els trets necessaris per diferenciar el referent dels altres membres de la seva classe general.
4. Claredat: una definició ha de ser tan clara (en el sentit de *no ambigua*) com sigui possible.
5. Consistència: mots semblants han de ser definits de manera semblant, i mots relacionats, de manera relacionada.
6. Objectivitat: una definició lexicogràfica no pot contenir opinions o elements subjectius que puguin ser motiu de discussió ideològica.

Un dels plantejaments més explícits per a la caracterització d'un sistema de definició és el que fa Mel'čuk (1988a), que es basa en els següents principis:

1. Principi d'univocitat: el llenguatge de definició no ha de contenir termes ambigus o sinònims.
2. Principi d'adequació: a la definició d'una unitat lèxica L, cada component ha de ser necessari, i el conjunt de tots els components ha de ser suficient, per a identificar únicament L en tots els seus usos imaginables.

¹³⁹ La definició referencial és equivalent a allò que Ilson (1987) anomena definició analítica (veg. més enrere § 4.2).

¹⁴⁰ Benson & al. (1986) usen aquest terme en el sentit de “metallengua de signe” de Rey-Debove (1967, 1969, 1971, 1978).

3. Principi de descomposició: la definició d'una unitat lèxica L ha de contenir només termes que siguin semànticament més simples que L.

4. Principi del bloc màxim (*maximum block*): si una definició conté una frase lliure composta de les unitats lèxiques $L_1 + L_2 + \dots + L_n$ que és semànticament equivalent a una unitat lèxica L, aleshores la frase s'ha de reemplaçar per L.

5. Principi del pont semàntic (*semantic bridge*): les definicions de qualsevol parell d'unitats lèxiques del mateix vocable¹⁴¹ han d'estar explícitament relacionades, ja sigui per un pont semàntic¹⁴² o per una sèrie de ponts semàntics.

6. Principi d'uniformitat: dos vocables pertanyents al mateix camp semàntic s'han de presentar d'acord amb el mateix esquema: (i) les unitats lèxiques relacionades d'aquests vocables han d'aparèixer en el mateix ordre en cadascun dels dos vocables; (ii) les distàncies semàntiques entre les unitats lèxiques relacionades s'han de representar idènticament.

D'aquests 6 principis, els quatre primers fan referència directa al sistema d'expressió de les definicions, mentre que la resta fan referència a les relacions que mantenen les definicions en l'àmbit de l'article o del conjunt d'articles que formen un camp semàntic determinat. Mel'čuk (1988a) formula, a més, onze criteris *heuristics* com a compleció d'aquests principis bàsics. Aquests criteris fan referència a tres aspectes diferents de la descripció lexicogràfica:

- a) l'estructura d'una definició particular (condicions d'equivalència),
- b) la distinció de sentits en un article,
- c) l'ordenació dels sentits en un article.

¹⁴¹ Mel'čuk (1988a: 169): "A *vocable* is the set of all lexical units that (i) their signifiers are identical and (ii) the signifieds of any two units are linked (directly or indirectly)".

¹⁴² Mel'čuk (1988a: 169): "A *semantic bridge* between lexical units L_1 and L_2 is the non-trivial component common to their definitions, which formally expresses the semantic link felt between the two".

El sistema de definició descrit per Mel'čuk, a partir de l'aplicació sistemàtica de l'anomenat principi de descomposició, porta com a resultat de la descripció lexicogràfica a la determinació dels *primitius* semàntics. Com se segueix del sistema de definició en què es basen els DEC, un sistema de definició representa, també, la determinació de la naturalesa dels elements primaris que l'integren.

A partir de la presentació que n'he fet, es constata que els sistemes de definició revisats en aquesta secció giren entorn de dues qüestions: l'establiment dels elements primaris de la metallengua i l'expressió de les condicions del sistema de formulació de les definicions. M'ocupo específicament de la primera qüestió a § 6.2 i de la segona a § 6.3.

6.2 Els elements primaris d'un sistema de definició

Un dels temes tractats amb més recurrència per la lexicografia teòrica que s'ha ocupat de la metallengua de definició ha estat la possibilitat d'utilitzar algun tipus d'elements primitius (lèxics o semàntics) en la descripció lexicogràfica. Sintèticament, es poden observar, en relació a aquesta qüestió, dues tendències; d'una banda, la que entén els primitius com els components mínims d'una metallengua de definició que presenta diversos graus de complexitat; d'una altra banda, la que postula l'existència d'un conjunt limitat de primitius *universals* que constitueixen *per se* una metallengua de descripció semàntica.

La primera de les dues tendències es pot vincular amb els DEC de la MTT, però també amb els autors que s'han ocupat de l'establiment del conjunt d'elements primaris a partir de l'anàlisi i del processament de la informació semàntica dels diccionaris convencionals, de manera que aquests s'utilitzen com a fonts per a la detecció dels lexemes que presenten un grau més alt de recurrència en les definicions lexicogràfiques.

La segona de les dues tendències que he assenyalat es basa en la idea que una teoria de la descripció del lèxic (i en concret de la descripció del significat lèxic) ha de dotar-se a

priori d'una metallengua de descripció i d'uns mecanismes d'anàlisi basats en elements teòrics externs a la concreció del llenguatge que es vol descriure. Aquests elements, pel que fa a la determinació de la metallengua de descripció, poden prendre en consideració factors d'ordre extralingüístic com ara la cognició humana, la comparació de les diferents llengües, el caràcter més o menys bàsic dels lexemes, etc.). Una metallengua postulada a partir d'aquest mètode ha de ser, en tot cas, validada com a eina de descripció lexicogràfica. A causa de l'orientació d'aquest treball, només m'ocuparé d'il·lustrar sumàriament la tendència basada en els primitius universalistes, mentre que centraré l'atenció en la les metallengües de descripció basades en el llenguatge natural.

Gold (1983) s'ha referit explícitament a la possibilitat d'integrar els primitius en la descripció lexicogràfica, i la posa en relació amb la proposta de Weinreich (1962) que he esmentat anteriorment (veg. § 6.1):

In order to *minimize* the inconvenients of natural metalanguage, the lexicographer might try to limit is as few terms as possible, i. e. to a set of lexical primitives accepted axiomatically and symbolizing first-order experiencies, on which all definitions would ultimately fall back. Definitions could then be classified according to the primitiveness of their components: level-0 definitions consisting solely of undefined primitives: level-1 definitions, of components defined solely by level-1 definitions (or level-0 and level-0 (*sic*) definitions, etc. [Gold (1983: 140)]

El sistema de definició dels DEC, de fet, correspon (Apresjan, 2000)¹⁴³ a la formulació de la denominada escola de Moscou que, juntament amb la de Polònia, està orientada a la descripció lèxica a partir de la utilització de primitius semàntics. Aquestes dues

¹⁴³ Apresjan (2000) fa referència a les dues principals escoles de semàntica lèxica, que tenen en comú el fet de formular les seves descripcions a partir de formulacions lexicogràfiques: “The question of the language of explications, or semantic metalanguage, has been discussed over the past thirty to forty years in all the advanced lexicographies of the world. (...) the ideas of greatest interest are those set forth on this subject by two contemporary semantic schools: the Moscow School (...) and the Polish School.” [Apresjan (2000: 215)]

escoles representen, respectivament, les dues tendències a què m'he referit més amunt. D'acord amb els postulats teòrics de l'escola de Moscou, el metallenguatge de descripció lexicogràfica és un subllenguatge de la llengua descrita, el qual integra una part petita i unificada del seu vocabulari. La base d'aquest subllenguatge són els primitius semàntics, que s'obtenen a través de la descomposició successiva dels components de la definició lexicogràfica del DEC:

With the aid of the metalanguage, complex semantic units of the object language (grammatical as well as lexical) are reduced to a fixed structure of semantic primitives by a process of hierarchical breakdown. The resulting explications (definitions) have a specific theoretical status: using them as a basis we may establish systematic paradigmatic links between various units of the language and formulate rules for the interaction of linguistic meanings. [Apresjan (2000: 224)]

Mel'čuk (1984: 5), en aquesta mateixa línia, assenyala que, per a l'evitació de la circularitat lexicogràfica, en les seves definicions “on prête une attention particulière à n'utiliser dans celles-ci que des lexèmes définis eux-mêmes sans faire appel au lexème considéré”. Per tant es pot dir que “du point de vue formel, cela signifie que la définition d'un lexème dans le DEC représente sa *décomposition* lexicale, qui, poussée à la limite, nous conduit aux *éléments sémantiques primitifs*”. Aquest criteri es fonamenta sobre el *principi de descomposició*, que implica que la definició d'una unitat lèxica només pot contenir termes que siguin semànticament més simples que aquesta; i s'entén que L_1 és semànticament més simple que L_2 quan L_2 pot ser definit en termes de L_1 , però no a l'inrevés. Aquest principi implica la *descomposició* de cada significat complex en significats progressivament més simples, de manera que els primitius són un resultat de la tasca lexicogràfica i no vénen donats a priori.¹⁴⁴ Vegem-ne un exemple:

¹⁴⁴ En paraules de Mel'čuk: “According to Decomposition Principle, the meaning of the lexical unit defined is always decomposed, in its definition, into simpler meanings. Applied methodically, such decompositions guarantee the absence of vicious circles in the system of definitions and eventually lead to semantic primitives.” [Mel'čuk (1988a: 175)]

COMPLIMENTER, verbe

1. *X complimente Y pour Z* = X loue Y, des qualités, de Z de Y ou des actions de Z de Y dans le but de faire plaisir à Y.

(...)

2. *X complimente Y pour Z* = X prononce un petit discours ou des paroles où X félicite Y à propos de Z et où d'habitude X complimente_1 Y.

Les aproximacions basades en primitius es plantegen l'estructura dels diccionaris com una entitat de plena coherència lògica, és a dir, completament alliberada del problema de la circularitat. Això s'aconsegueix a partir d'un conjunt d'expressions bàsiques, no definides, que forma el nucli de la metallengua de definició, i que en constitueix la base axiomàtica. Un diccionari construït a partir d'aquesta idea, doncs, està en situació d'evitar la circularitat, ja que qualsevol expressió definidora ha de contenir o bé aquest tipus d'elements axiomàtics, o bé lexemes que són definibles directament o indirectament pels axiomàtics.

Aquesta possibilitat ha estat posada de manifest des del camp de la lexicologia per Picoche (1977),¹⁴⁵ que remarca especialment la dificultat que representa l'establiment d'una llista discreta d'aquests axiomes:

les difficultés commencent lors qu'on essaye de dresser la liste de ces **mots-axiomes**, auxquels Pascal a donné le nom de **primitifs** et que certains linguistes américains appellent "*atomic predicate*". On pense en principe qu'ils sont en nombre restreint et limité, mais personne ne les a jamais comptés, quoique le problème soit posé de longue date. [Picoche (1977: 149)]

¹⁴⁵ Picoche parteix de la base que aquests primitius tenen i han de tenir una base lèxica, és a dir, que es tracta de *mots* no definibles sobre els quals es fonamenta la descripció lexicogràfica: "Un dictionnaire qui se voudrait logiquement organisé devrait donc accepter l'idée qu'il existe dans la langue des mots comparables à ce que sont les axiomes en mathématiques; que ces mots sont complémentaires, la définition de l'un finissant où celle d'un autre commence, mais non analysables; qu'il serait honnête de ne pas chercher à en donner une définition en trompe-l'oeil; que ces mots, en nombre forcément réduit et limité, sont à la base de la définition de tous les autres." [Picoche (1977: 148-149)]

Cal considerar bàsic en aquest plantejament metodològic, doncs, el problema de la determinació estricta del nombre i tipus d'unitats a què cal recórrer per arribar a l'inventari d'elements de què ha d'estar compost el conjunt axiomàtic, i que han de complir la condició general de ser suficients i la condició particular de ser necessaris per a la caracterització jeràrquica de tot el lèxic que s'inclouï en la nomenclatura del diccionari.¹⁴⁶

Entre les dificultats que representa proposar un criteri acceptable per a la detecció d'aquests primitius (que, per imposició del mètode lexicogràfic han de tenir el mateix estatus que la resta dels mots de la llengua) destaquen, segons Picoche:

L'établissement d'une métalangue composée de primitifs univoques exigerait donc qu'on opte pour un mot pas concept en se débarrassant des synonymes, et que parmi tous les sens d'un polysème, on en choisisse arbitrairement un à l'exclusion de tous les autres. Énorme effort lexicologique qui n'est même pas entrepris et qui aboutirait peut-être à un progrès sur le plan de l'analyse sémique, mais sûrement pas sur celui du naturel et de la substituabilité des définitions. [Picoche (1977: 150)]

Diversos autors, com Martin (1974, 1977a i 1977b) o Calzolari (1977), han partit de la idea que el factor d'implicació en les definicions i l'estructuració jeràrquica del lèxic que se'n deriva porten de manera natural a l'establiment dels primitius semàntics. Des del punt de vista lexicogràfic, l'establiment d'un conjunt de primitius podria configurar la base axiomàtica del lèxic, i permetria d'evitar la circularitat lexicogràfica. Martin pretén d'aïllar els primitius a partir dels elements d'un graf d'inclusions que presenten una alta concentració de nodes, és a dir, els "cims" dels grafs.

¹⁴⁶ Pot ser per aquest motiu que les úniques realitzacions lexicogràfiques fetes segons aquests principis no corresponen a la idea de diccionaris generals, sinó més aviat de recopilacions fragmentàries com ara els substantius referits a objectes d'ús comú (Wierzbicka, 1985), a tipus gramaticals com ara els verbs d'acte de parla (Wierzbicka, 1987) o, en el millor dels casos, recopilacions sistemàtiques de verbs en general (veg. Lewicka & Bogacki 1983). El ritme de publicació dels DEC tampoc no permet equiparar-los a

La proposta de Martin consisteix a aïllar els primitius semàntics sobre la base “empírica” dels diccionaris i convertir aquests primitius en axiomes lexicogràfics indefinibles, de manera que totes les definicions estiguin integrades per elements que, al seu torn, siguin definits per mots reductibles a aquests axiomes i el diccionari en el seu conjunt sigui, doncs, equiparable a un sistema lògic.¹⁴⁷

El sistema de detecció dels primitius lexicogràfics a partir de l’anàlisi de diccionaris existents es pot considerar un precedent de la lexicografia computacional. De fet, els treballs de Martin sobre l’associació de descriptors a les entrades lexicogràfiques a partir de llur definició es poden considerar un clar precedent d’allò que després han estat les eines de taxonomització de diccionaris per a l’extracció d’informació semàntica desenvolupades en l’àmbit de la lexicografia computacional.¹⁴⁸

Cal reconèixer, però, que aquest tipus de sistema de deducció parteix d’un pressupòsit implícit sobre el qual es poden fer algunes reserves importants, ja que presenta algunes limitacions: els primitius lexicogràfics ocupen els nodes més “alts” i recurrents de les jerarquies lèxiques que s’estableixen en un diccionari. La possibilitat d’extreure els primitius semàntics dels diccionaris parteix de la base que cal taxonomitzar *tot* el

diccionaris generals, tot i que des del punt de vista de la selecció de la macroestructura no segueixen cap criteri determinista.

¹⁴⁷ Segons Martin (1977b), la legitimitat de l’equiparació d’un diccionari a un sistema lògic s’ha d’avaluar a la vista d’algunes reserves importants, a saber: que el diccionari no és un sistema tancat, sinó obert sobre l’univers referencial; que els diccionaris s’adrecen a usuaris que tenen coneixements totalment independents de l’obra consultada; i que la lexicografia no està sotmesa als mateixos principis d’ordenació que els sistemes lògics, sinó “aux aléas de l’ordre alphabétique”. De tota manera, Martin (1977b: 59) opina que “en dépit de toutes ces réserves importantes, le chemin qui conduit des primitifs supposés aux autres vocables, s’il est artificiel, n’est pas dénué d’interêt. Même si pratiquement la procédure lexicographique a un caractère intuitif et pragmatique, le problème théorique de la cohérence ne peut être ignoré.”

¹⁴⁸ Martin (1977b) es proposava la utilització dels mitjans informàtics per a la seva recerca; pel que sembla no va materialitzar mai les seves intencions en aquest sentit. Amb posterioritat a Martin (1977b) s’han publicat un cert nombre de treballs en aquesta mateixa línia, com per exemple Calzolari (1988) o Fontenelle (1992).

diccionari a partir de l'anàlisi de les seves relacions d'inclusió, una altra cosa és si això és adequat metodològicament i útil pràcticament, a la vista dels diferents tipus de relacions que es poden projectar a partir de l'anàlisi de les definicions de diccionari.

Quan, en aquesta mateixa secció, m'he referit a les estratègies de determinació dels primitius semàntics, he fet esment del fet que algunes de les metallengües basades en universals troben en la lexicografia (fonamentalment teòrica però no exempta de resultats aplicats) una fórmula per a la validació dels seus mecanismes d'anàlisi.

En les aproximacions universalistes, hom parteix de la base que els requisits comuns a la lexicografia i a la descripció lingüística es poden complir difícilment si el lexicògraf utilitza de manera exclusiva el llenguatge natural com a eina de descripció. Els dos principis bàsics de la descripció semàntica: l'adequació i la certesa lògica de l'enunciat d'equivalència esdevenen, en la lexicografia pràctica, dos objectius difícils d'assolir en els diccionaris d'ús corrent, tal com demostren moltes anàlisis de diccionaris.

Si la comesa principal d'una definició és la construcció d'un enunciat veritable, cal entendre com a desviacions no del tot intrínseques a la lexicografia, sinó associades a la praxi de la lexicografia convencional,¹⁴⁹ aquells casos en què la definició lexicogràfica formula enunciats falsos, o bé incorpora especificacions supèrflues, o bé dóna un sentit massa extens per a un mot.

¹⁴⁹ Sigui quina sigui l'extensió i la naturalesa d'una obra lexicogràfica, cal tenir present que l'adequació descriptiva exigeix donar compte d'una sèrie de fenòmens en una determinada direcció. Altrament, la descripció contindrà elements que la faran inadequada. Tal com assenyala Wierzbicka (1996): "The description of a word's meaning may vary, legitimately, in completeness from one work to another, but it should not differ in its basic content. A «definition» is meant to represent the truth about a word's meaning, and there is only one such truth, whether it is to be presented in a research paper devoted to one particular word or in a dictionary intended for a general audience, including various dictionaries addressed specifically to «children», «learners», «students», and so on." [Wierzbicka (1996: 259)]

El lexicògraf haurà aconseguit el seu objectiu quan hagi integrat dins de la seva definició allò que és *invariant* en el significat d'un mot, és a dir que satisfaci el doble requeriment de reunir els aspectes pertinents per a la seva caracterització, i alhora no permetre la proliferació no fonamentada de sentits, que condueixen a la manca d'adequació descriptiva.

Aquests pressupòsits caracteritzen els sistemes de definició basats en els primitius universals, com les propostes de Wierzbicka (1972, 1985, 1996), que dedica alguns dels seus treballs a l'establiment d'un reduït grup de primitius semàntics indefinibles.¹⁵⁰ Tot i que a Wierzbicka (1972) no hi ha referències explícites a la lexicografia, en treballs posteriors (Wierzbicka, 1985, 1987) tria aquesta disciplina per a portar a la pràctica les seves descripcions semàntiques i fa una adaptació del seu conjunt de primitius semàntics per a la descripció lexicogràfica, amb la qual cosa aquest conjunt esdevé la base axiomàtica de la seva metallengua de definició. Wierzbicka (1996) fixa un conjunt de primitius *universals* de descripció a partir de la consideració de llengües diferents i de la valoració particularitzada de cadascun.

Això dona lloc a una proposta de lexicografia de base mentalista,¹⁵¹ exhaustiva en la seva anàlisi i que s'explicita a través d'una metallengua establerta amb uns criteris de

¹⁵⁰ Diu Wierzbicka (1972): "In the course of the seven years I have spent studying linguistic facts in search of the semantic primitives, my number of suspects has systematically diminished. I am of the belief that there are somewhere between ten and twenty of them. Below I present a list of my present favourites:

want	something
don't want (diswant)	someone (being)
feel	I
think of	you
imagine	world
say	this
become	
be a part of"	

¹⁵¹ És especialment indicatiu d'aquest punt la defensa que fa Wierzbicka (1985) de la intuïció i els mètodes introspectius d'anàlisi semàntica.

màxima simplicitat, que rebutja les classificacions científiques i que només considera adequadament descriptives aquelles definicions que expressen les intuïcions dels parlants ordinaris.

El sistema de definició proposat per Wierzbicka (1985) consisteix, en conjunt, a fer allò que ella mateixa anomena el *full portrait* d'un terme.¹⁵² Aquest *full portrait* defuig els criteris de distinció mínima:

To avoid arbitrariness in the choice of other words to compare a given word with, one must abandon the diacritical approach in favour of the Aristotelian model of a full, positive representation of a concept. [Wierzbicka (1985: 38)]

La manera de defugir la circularitat en les definicions consisteix en l'aplicació d'una *anàlisi reductiva*, que postula que “all concepts must be defined via a set of undefinable semantic primitives” (Wierzbicka, 1985: 45), i que en les definicions només hi poden aparèixer aquells mots que es defineixen exclusivament amb indefinibles.

El mètode de definició d'objectes i classes d'objectes posat en pràctica per Wierzbicka (1985) opera a partir de la conceptualització dels lexemes i per l'exposició detallada de totes les seves característiques en relació al seu propi coneixement del món o per consulta a parlants nadius. El resultat és una definició de tipus compositiu, que agrupa les diferents característiques del definit i el determina en tots els seus aspectes més o menys rellevants. Vegem-ho amb un exemple, que reproduïx el contingut de l'entrada per a la classe dels cotxes (*cars*) (Wierzbicka, 1985: 143):

¹⁵² L'aplicació d'aquest criteri d'exhaustivitat informativa dona lloc a definicions d'una extensió inusual, la qual cosa que no representa la inclusió d'informació enciclopèdica: “To demand that the semantic description of a concept should be exhaustive does not mean to abandon distinction between dictionaries and encyclopaedias. A good dictionary definition must be exhaustive in the sense that it states all the components of the concept, not in the sense that it includes all the available knowledge about the denotata.” [Wierzbicka (1985: 40)]

CARS

A KIND OF THING MADE BY PEOPLE

FUNCTION

a thing of this kind is made for a few people - not many people but more than one -
to be able to go by means of it from one place to another

without having to make any bodily effort

fast and comfortably

being able to take more things with them than they would be able to carry

being able to get quickly to a place far away

being able to go anywhere where there is a road

[...]

MODE OF OPERATION

a thing of this kind can move because of something that happens in some part of it
a person sitting inside can cause it to happen

[...]

CONSTRUCTION

things of this kind have wheels underneath,

on both sides, at the front and at the back,

so that they don't have to be supported against anything else when they are not
moving

with the people inside being able to sit and do nothing

things of this kind have a higher part where the people sit and a lower part in front
of it

the bottom part of these parts is as close to the ground as it can be without coming
into contact with the ground in places where the ground is uneven

[...]

En tot cas, l'única obra pròpiament lexicogràfica d'aquesta autora es refereix als *speech act verbs* (Wierzbicka, 1987), i és on desenvolupa l'aplicació més sistemàtica de la seva concepció de l'activitat lexicogràfica. La formulació de les definicions és també compositiva de les diferents característiques dels verbs que descriu, i que conceptualitza a partir de les premisses, associacions i accions que es poden deduir de la seva utilització, de manera que el significat d'un verb qualsevol s'expressa a partir de fórmules com la següent:

ACCUSE

(...)

Meaning

I say: person X did Y

I assume that everyone would say that doing Y is bad
I assume that something bad should be said about person X because of that
I feel something bad thinking about it
I want someone to show that this is not true if they can
I assume no one could
I say this because I want to cause something to happen because of Y that will be bad for X

Per a Wierzbicka (1987), els components utilitzats en aquestes definicions són de dos tipus bàsics: el 'dictum' (*I say...*) i la 'illocutionary purpose' (*I say this because...*);¹⁵³ de més a més, el significat dels *speech act verbs* pot incloure diverses assumpcions, emocions, pensaments i intencions, que es troben a la part central de la definició.

El problema que plantegen, però, els primitius *universals* de Wierzbicka és que constitueixen un discurs d'explicitació que, encara que formalment es pot identificar com a lexicogràfic, cau fora de l'àmbit estricte de la lexicografia. Això fa que difícilment les seves aportacions puguin ser incorporades a la normalització metodològica dels diccionaris.

Cal assenyalar que els plantejaments basats en primitius, especialment les estratègies componencialistes (o basades en la determinació dels primitius a partir de la

¹⁵³ Wierzbicka (1987) assenyala que: "The «dictum» represents the overt content of the utterance, whereas the «illocutionary purpose» represents the speaker's (purported) intention in making that utterance. For example, if I say to somebody «I inform you that I am leaving», the dictum is «I say: I am leaving», whereas the illocutionary purpose is «I say this because I want to cause to know it». Of course my real purpose in making that utterance may be different from the purported one, but the semantic formula aims at the purported purpose, encoded in the meaning of the verb *inform*, not at the real purpose, which may or may not be the same." [Wierzbicka (1987: 18)]

descomposició dels enunciats de definició) poden contribuir a la coherència estructural dels diccionaris, integrats en els seus sistemes de definició. Més endavant (veg. § 7.4 a § 7.7) tracto de la determinació dels elements primaris per a un diccionari general monolingüe del català. Abans, però, i a partir de les observacions analitzades anteriorment, faré una proposta de síntesi d'un sistema de definició per a un diccionari general.

6.3 Les condicions d'un sistema de definició

A partir de l'anàlisi d'allò que s'ha dit sobre els principis de la definició i sobre els sistemes de definició, constatem que aquesta qüestió s'ha abordat des de punts de vista diversos i que, a partir de les observacions precedents de § 6.1 i de § 6.2, és difícil arribar a caracteritzar un sistema de definició com el conjunt de condicions que vinculen la metallengua de descripció amb els mecanismes d'explicitació del significat de les definicions lexicogràfiques.

Un sistema de definició concret està lligat al tipus, l'orientació i la finalitat de cada diccionari. En aquest sentit, és cert que per a cada diccionari es pot determinar un conjunt diferent d'assumpcions que regulen la forma final del seu discurs de descripció del significat, en funció de l'estructura de constituents de l'article.

Però també és cert que sembla possible l'enumeració d'una sèrie de condicions universals que ha de satisfer un sistema de definició concebut amb criteris actuals, que tingui en compte les aportacions que des de la lexicografia teòrica s'han fet en aquest terreny. En tot cas un sistema de descripció ha de contenir una sèrie de mecanismes que permetin assegurar determinats requisits descriptius. Encara que el grau de profunditat i de rigor de cadascun d'aquests requisits no pot tenir un mateix valor per als diferents tipus de diccionaris monolingües, a continuació assajaré de proposar un conjunt de

condicions bàsiques que podrien fonamentar el sistema de definició d'un diccionari. En molts casos, aquestes condicions són conseqüència directa dels aspectes de la descripció lexicogràfica que he tractat específicament en aquest treball, en especial de les relacions que mantenen els diversos elements estructurals del diccionari entre ells i de les projeccions semàntiques que poden fer-se a partir de les definicions. Crec, finalment, que aquestes condicions no han estat determinades —almenys de manera conjunta— per la lexicografia teòrica com a reguladores d'un sistema de definició i que, per tant, constitueixen en conjunt una de les aportacions significatives d'aquest treball.

1. Condió de regularitat. En un sistema de definició lexicogràfica, les regularitats s'han de tractar regularment. Es poden establir diversos casos de regularitat, a diferents nivells; per exemple, les alternances d'ergativitat que es produeixen en els verbs, que originen de manera sistemàtica un sentit causatiu i un de no causatiu per al mateix verb; les alternances metonímiques del tipus continent-contingut, etc. Aquests casos, i altres que sorgeixen en la descripció del lèxic, poden ser tractats de maneres molt diverses en els diferents diccionaris; cal, però, que se segueixin uns patrons deliberats i sistemàtics.

2. Condió de distinció de constituents. En un sistema de definició lexicogràfica, els constituents *extrínsecs* i els *intrínsecs* han de trobar-se distingits. Una definició lexicogràfica ha de donar informació sobre la manera com el definit s'insereix en els enunciats lingüístics, per tant, la determinació explícita de la seva naturalesa sintàctica ha de quedar recollida en la definició, a través dels constituents corresponents.

3. Condió d'adequació. En un sistema de definició lexicogràfica, la informació continguda en la definició ha de ser necessària i suficient per a la caracterització semàntica del mot.

4. Condió d'homogeneïtat discursiva. En un sistema de definició lexicogràfica, els recursos expressius no són tots els disponibles en la llengua. Cal determinar un discurs estilístic homogeni que va des de qüestions bàsiques de tria lèxica (ús de *ésser* o de *ser*; *cap a/vers/envers*, etc.) fins a gramaticals (*per/per a*, estructures admeses, etc.). També es fa necessari l'establiment dels recursos d'expressió discursiva amb què una definició

es refereix als seus mateixos components (ús de persones gramaticals, diferenciació d'actants semàntics, etc.).

5. Condició de síntesi descriptiva. En un sistema de definició lexicogràfica, una definició perifràstica ha d'aparèixer un sol cop en el diccionari, i cal evitar l'aparició de definicions repetides o semànticament equivalents, a través d'un tractament coherent i normalitzat de les sèries sinonímiques.

6. Condició d'univocitat. En un sistema de definició lexicogràfica, una definició ha de ser unívoca. En un grau màxim, cada un dels elements del vocabulari de definició ha d'estar desambigüat, ja sigui mitjançant la referenciació sobre el mateix diccionari dels sentits amb què es construeixen les definicions o bé decidint *a priori* quin sentit únic de cada mot es vol aplicar. En general, i com a compliment mínim d'aquesta condició, cal que un diccionari expliciti els seus mecanismes de desambiguació, tant dels elements que formen part de les definicions perifràstiques com de les sintètiques.

7. Condició de projecció coherent de relacions. En un sistema de definició lexicogràfica, cal evitar la circularitat estructural, de manera que el conjunt de relacions projectades per les definicions lexicogràfiques donin lloc a estructures coherents des del punt de vista semàntic.

8. Condició de transparència. En un sistema de definició lexicogràfica els components d'una definició han de ser considerats individualment, sense valors que en funció de la co-dependència dels elements introdueixin opacitats, com ara l'ús d'expressions fraseològiques no indicades explícitament, etc.

9. Condició de limitació del vocabulari definidor. En un sistema de definició lexicogràfica, els mots disponibles no són tots els de la llengua, i cal imposar els mecanismes de limitació adequats per a la determinació dels elements disponibles en la formulació de definicions.

10. Condició d'integració. Un sistema de definició lexicogràfica ha d'estar integrat amb els mecanismes d'explicitació referits a altres elements estructurals del diccionari.

La capacitat descriptiva de totes aquestes característiques haurà de ser major o menor en funció que la mateixa definició incorpori informacions que solen anar fora; en concret, em refereixo, d'una banda, al sistema d'expressió de les subcategories dels lexemes (la transitivitat dels verbs, comptabilitat dels substantius, etc.); de l'altra, a les fórmules d'expressió de les característiques combinatòries del definit (règims preposicionals, col·locacions, etc.).

Com es pot suposar, aquests principis poden aplicar-se amb fórmules de representació i nivells de profunditat molt diferents segons els diccionaris, però crec que es poden considerar com qüestions que el sistema de definició d'un diccionari qualsevol, de manera implícita o explícita, ha de tenir presents. Moltes de les condicions exposades fan referència directa o indirecta al conjunt del vocabulari definidor. Com ja he fet notar a § 6.1 Aquest conjunt no és lliure, sinó que està sotmès a les limitacions que imposa el fet de constituir el component fonamental d'una metallengua de definició. En els capítols següents (§ 7 i § 8), em concentro en l'estudi de les unitats d'aquesta metallengua, en funció, preferentment però no únicament, de les dues condicions bàsiques que la determinen de manera directa: la condició d'univocitat (condició 6) i la condició de limitació (condició 9).

7 Una metallengua de definició per al català

7.1 El vocabulari proposat com a punt de partida

Ja he dit (veg. més enrere § 5.3) que els vocabularis constituïts amb els principis aplicats a § 5.6 no han estat concebuts específicament per a una utilització en lexicografia general, sinó, en primer terme, com un recurs orientat a l'aprenentatge de segones llengües. Només complementàriament aquests recursos han estat utilitzats en obres lexicogràfiques destinades a aprenents estrangers i això ha requerit adaptacions estretament vinculades amb els processos de redacció dels diccionaris que els han aplicat.

L'aplicació dels vocabularis controlats a la lexicografia, juntament amb la reflexió metodològica que ha suposat l'aparició de la lexicografia computacional, ha comportat que es valorin les millores estructurals que comporta l'ús de vocabularis definidors amb algun tipus de restricció quantitativa. La importància del rigor estructural del diccionari —pel que fa a la projecció de les relacions de diversa naturalesa vehiculades per la descripció semàntica que conté— ha estat posada de relleu per l'aplicació de mitjans informàtics a la redacció, a la publicació i a l'extracció d'informació dels repertoris lexicogràfics.

De tota manera, el vocabulari proposat (veg. Apèndix I.10) no és sinó un mer punt de partida per a la formulació d'una proposta més elaborada que garanteixi la coherència i les possibilitats de tractament posterior de les obres lexicogràfiques. Tal com he dit a § 6.1, per a la consecució d'aquest objectiu cal estructurar convenientment el vocabulari i

integrar-lo en un sistema de definició que conformi una veritable metallengua de definició.

Entenc que un requisit indispensable per a una utilització rigorosa d'aquest vocabulari en tasques lexicogràfiques és la caracterització semàntica de les seves unitats. Dit d'una altra manera, cal especificar, per a cadascun dels mots que conformen el vocabulari, en quin sentit o sentits pot ser utilitzat com a element de descripció semàntica. D'aquesta manera, el vocabulari passa de ser una mera llista de mots disponibles a tenir una estructura que presenta nombrosos punts de contacte amb la d'un diccionari i que integra una descripció semàntica precisa de cadascuna de les seves unitats.

Per a la consecució de l'objectiu de l'anotació semàntica del vocabulari de descripció he seguit una estratègia basada en el mètode de descripció lexicogràfic: he incorporat a cada element del vocabulari la informació de l'article que hi correspon al DIEC i he aplicat una sèrie de mecanismes d'anàlisi que m'han portat a l'articulació d'una metallengua de descripció que pot ser usada com a base de la descripció lexicogràfica.

L'aplicació d'aquestes operacions utilitzant com a punt de partida un diccionari convencional (DIEC), senzilla des del punt de vista de la seva implementació pràctica, revela, d'entrada, dos fets:

1. L'existència d'elements redundants: en el vocabulari proposat hi ha, de fet, grups de sinònims (en sentit ampli: fórmules d'expressió equivalents, etc.) susceptibles de ser reduïts a una sola unitat.
2. L'existència d'un nombre de sentits limitat (respecte al conjunt de sentits possibles) que realment s'utilitzen (o que són susceptibles de ser utilitzats) en la redacció de definicions.

Pel que fa a la primera d'aquestes constatacions, cal partir de la base que una llista de lexemes basada en dades freqüencials depèn fortament del tipus de corpus de la qual s'extreu (veg. més enrere § 5.5); en aquest sentit, la *representativitat* d'un corpus de referència ha de ser presa com un paràmetre que cal orientar a la finalitat que es

persegueix, és a dir, la d'un vocabulari caracteritzat bàsicament pel fet de contenir els elements necessaris i suficients que permetin l'explicitació de qualsevol significat d'un lexema. A partir de les experiències portades a terme per a d'altres llengües amb l'ús de vocabularis controlats, es pot inferir que els lexemes del conjunt seleccionat són suficients per a la determinació dels significats dels lexemes. Cal, però, considerar a la vista d'una anàlisi més aprofundida la condició de la necessitat de cadascun d'aquests elements en un conjunt bàsic de descripció. Eliminar la redundància del vocabulari de partida serà, doncs, un dels objectius bàsics per a la consecució d'una metallengua de definició. Això suposa determinar les opcions que cal prendre per a eliminar la presència de fórmules diverses per a l'expressió del mateix significat: eliminar sinònims (*semblant/similar*) o bé determinar les fórmules habituals de definició entre expressions intercanviables, com ara *semblant a / similar a / que sembla / que s'assembla a*.

Pel que fa a la segona constatació, cal entendre que la vinculació de cada element del vocabulari amb el seu sentit o sentits concrets ha de filtrar aquells sentits que no es consideren bàsics per a la descripció de la resta de les unitats. Aquesta vinculació, que explícito mitjançant paràfrasis de definició, suposa determinar els significats actualitzables dels lexemes que s'usaran en l'expressió de definicions: resoldre ambigüitats, evitar la concurrència de sentits diversos en una mateixa expressió lèxica, etc.

La consideració d'aquestes característiques, que no es compleixen directament a partir del vocabulari aïllat, requereixen la conversió d'aquest vocabulari en una metallengua de descripció que doni prioritat als aspectes estrictament lexicogràfics.

La realització de les operacions necessàries per a la integració del vocabulari en un sistema de definició constitueix el primer pas en la determinació de la metallengua de descripció. Ja he mostrat que la tria d'aquest conjunt no obeeix a criteris lexicogràfics, sinó que més aviat està fonamentada en paràmetres que poden determinar el grau de *generalitat* o el caràcter més o menys *bàsic* de cada lexema. Aquesta condició no

implica necessàriament que el conjunt presenti un grau satisfactori d'adequació descriptiva com a metallengua de descripció lexicogràfica.

Aquesta orientació pot plantejar certes limitacions, si s'examina a la llum de consideracions exclusivament lexicogràfiques:

pell *f.* Tegument resistent i flexible que recobreix el cos de l'home i dels animals.[DIEC]

xàpol *m.* Aixada ampla, sovint amb el màneg curt i corbat, que serveix per a cavar o per a arreplegar. [DIEC]

grum *m.* Cera blanca amb què es dona l'última capa a les atxes, a les candeles, etc. [DIEC]

En els exemples precedents, he subratllat els mots que apareixen en les definicions i que no formen part del conjunt proposat com a punt de partida. Una de les constatacions que es produeixen d'una manera més evident és que en determinats casos és el descriptor (semàntic i sintàctic) de les definicions precedents un dels elements que no figura en la proposta prèvia de vocabulari definidor (veg. Apèndix I.10), ja sigui pel fet que els definits corresponen a nivells molt baixos en l'organització jeràrquica (com en el cas de *xàpol*), o bé perquè (com en el cas de *pell*) el DIEC tria un descriptor menys *bàsic* que el definit. En el cas de la definició de *pell* es pot assenyalar la inadequació descriptiva en la tria de descriptor, que s'hauria evitat amb la utilització d'una fórmula de definició basada en un vocabulari controlat. Els dos casos restants, però, són diferents.

En el cas de *xàpol*, un diccionari elaborat amb vocabulari controlat tendirà a l'assignació de descriptors de nivells més alts de la jerarquia, la qual cosa condueix a la projecció de taxonomies més inespecífiques. Tal com fan notar Ide & Véronis (1994):

The most pervasive problem in automatically extracted hierarchies is the attachment of terms too high in the hierarchy (...). In dictionaries with a controlled definition vocabulary (such as the LDOCE), the problem of attachment at high levels of the hierarchy results also from a lack of terms from which to choose (...). As a result, hierarchies extracted from LDOCE are consistently flat. [Ide & Véronis (1994: 149)]

La definició de *grum*, d'altra banda, és un exponent de la tècnica de definició consistent a enumerar elements que compleixen alguna de les condicions del definit (cosa que fa referència a l'*extensió* més que no pas a la *intensió* del definit). Aquesta tècnica tendeix a introduir en la definició elements amb graus d'especificitat elevats i que, per tant, solen restar fora del vocabulari proposat. En el cas de *grum*: *Cera blanca amb què es dóna l'última capa a les atxes, a les candelles, etc.*, l'última part d'aquesta definició no introdueix una caracterització global dels elements denotats, sinó que es limita a establir-ne una enumeració (incompleta, en aquest cas) que denota directament cadascun d'aquests elements. El DIEC, per la seva relació històrica amb el DGLC, utilitza sovint aquest recurs en les seves definicions. Weinreich (1962) anomena *part ostensiva* de la definició aquests components referits a la naturalesa *extensional* (Svensén, 1993) de la semàntica del definit o d'un component de la definició.

Aquestes dues constatacions primàries (descriptors de nivells baixos i elements ostensius en la definició) poden ser identificades com els factors de diferenciació principal pel que fa a les estructures que projecten els diccionaris en funció del fet que recorrin o no a vocabularis limitats. Aquestes diferències poden posar en dubte l'opció d'introduir en els criteris de selecció de la mostra les dades freqüencials extretes del DIEC, en la mesura que aquest diccionari presenta una notable dispersió del vocabulari de definició. De tota manera, cal tenir en compte el fet que he recorregut únicament a la selecció del lèxic que presenta una freqüència més alta i que, per tant, és més recurrent en les definicions del DIEC. Una de les proves que es poden aportar en el sentit de la conveniència d'aquesta opció metodològica ha estat la relativa facilitat amb què he pogut portar a terme la regularització (o reformulació) de les definicions a què em refereixo a § 7.5.

Abans de procedir a la determinació de la metallengua de definició cal abordar, però, una qüestió que afecta les possibilitats d'aplicació del conjunt a establir. La meua proposta en aquest treball es refereix a la formulació d'una metallengua per al tractament del lèxic general del català. Per tant, no he pretès un sistema que tingui entre els seus objectius la descripció de cap llengua d'especialitat, en el sentit que el lèxic que

solen integrar aquestes llengües ha de ser objecte d'una estratègia més pròpia de la terminologia que no pas de la lexicografia general. Sovint els elements d'un vocabulari terminològic requereixen unes descripcions semàntiques precises des del punt de vista extralingüístic que solen incorporar elements de tipus científicotècnic que cauen fora de l'abast d'un vocabulari definidor concebut per a la caracterització del lèxic general.

Això no vol dir, però, que el vocabulari definidor d'un diccionari monolingüe no especialitzat no sigui suficient per a la caracterització de determinades unitats lèxiques que en els diccionaris convencionals sovint incorporen definicions tècniques, a falta d'estratègies més adequades de descripció, o per la incapacitat del lexicògraf de determinar el significat sense adoptar plantejaments científicotècnics com a referència. Sense voler estendre'm massa sobre la naturalesa i la funció d'aquestes definicions, crec que es pot dir genèricament que no compleixen les condicions descriptives d'un diccionari general, que ha de recórrer preferentment a les classes considerades naturals o socialment acceptades, que en ocasions no són totalment coincidents amb les classificacions científiques o les caracteritzacions tècniques:

molsa f. Planta de la divisió dels briòfits, fixada al substrat per mitjà de rizoides, de caulidis amb fil·lids disposats a l'entorn i espores que originen un protonema filamentós. [DIEC]

ferro m. Metall de transició, blanc grisenc, mal·leable, dúctil, molt tenaç, ferromagnètic, que fàcilment es rovella en contacte amb l'aire humit, molt difós a la natura, on constitueix aproximadament el 5% de la litosfera (símbol, Fe; nombre atòmic, 26; pes atòmic, 55,85). [DIEC]

Els dos exemples precedents corresponen a mots del lèxic general que estan definits a partir d'explicacions més tècniques que no pas lingüístiques. En principi, parteixo de la base que aquest tipus de mots poden definir-se a través de fórmules més pròximes a la caracterització lingüística. Una altra qüestió és fins a quin punt els termes científicotècnics (especialment els que presenten una freqüència més alta, poden o no ser definits a través d'un vocabulari general. Tal com fa notar Svensén (1993):

A technical-language definition of a technical term is often more detailed and complex than a general-language definition. [Svensén (1993: 134)]

Aquesta observació, però, no manifesta la inadequació de les definicions basades en una metallengua general. De fet, es pot concloure que aquestes corresponen d'una manera més natural als propòsits d'un diccionari monolingüe de tipus general.

7.2 Processament del vocabulari de partida

L'adaptació que he portat a terme sobre el vocabulari proposat es basa, com he assenyalat a § 7.1 en la vinculació del conjunt de mots seleccionats com a punt de partida amb els articles lexicogràfics del DIEC que hi corresponen. Per a cada lema seleccionat en la proposta de conjunt del vocabulari, he obtingut l'article o els articles del DIEC que s'hi refereixen com a entrada de diccionari. Així, he passat de tenir una llista de vocabulari a tenir un conjunt d'entrades de diccionari amb el seu corresponent article lexicogràfic. Aquest canvi permet realitzar un procés sobre el conjunt del vocabulari, amb del doble objectiu, d'una banda, d'incorporar al vocabulari definidor informació de naturalesa semàntica (a fi de fixar els sentits del vocabulari proposat que són rellevants des del punt de vista lexicogràfic) i, d'una altra banda, de determinar quines unitats del vocabulari proposat són rellevants des d'aquest mateix punt de vista.

Les operacions de selecció portades a terme sobre el conjunt d'articles corresponents al vocabulari proposat m'han permès l'establiment d'un conjunt tancat de vocabulari definidor anotat semànticament, que té, respecte de la metallengua de descripció global d'un diccionari, un caràcter axiomàtic (veg. més endavant § 8.1), en el sentit que li confereix Weinreich (1962).

Aquestes operacions de selecció s'han portat a terme sobre el conjunt d'articles del DIEC seleccionat a partir del vocabulari definidor proposat a § 5.6. S'obté així, com a punt de partida, una mostra de diccionari formada pel conjunt d'unitats d'un vocabulari

controlat i que, per tant, es pot considerar suficient per a la descripció de la totalitat dels lexemes.

En una primera anàlisi de les dades, he exclòs del conjunt els mots que es defineixen a partir de sinònims sintètics; en tots els casos el cap de la sèrie sinonímica (veg. Rafel, 1988) corresponent a aquest mot formava part també del vocabulari, en cas contrari hauria inclòs l'article corresponent al cap de sèrie. Aquesta operació prèvia ha comportat l'exclusió de mots del tipus:

abundós -osa *adj.* Abundant.

assolir *v. tr.* Aconseguir.

En el cas que alguna definició del conjunt contingués el mot definit sintèticament, he substituït en el text de la definició corresponent el sinònim pel cap de la seva sèrie sinonímica. Un cop eliminades aquest tipus de redundàncies en el conjunt del vocabulari, he realitzat sobre el conjunt d'articles les següents operacions, il·lustrades al diagrama de flux de la figura 7.1:

1. Eliminació de les entrades que no s'utilitzen en almenys una de les definicions del conjunt donat.
2. Selecció de sentits (en els articles polisèmics) actualitzables en les definicions, i eliminació de les definicions corresponents a sentits que no s'utilitzen.
3. Regularització metodològica de les definicions i reformulació de cada definició utilitzant exclusivament elements del conjunt.

Aquestes tres operacions estan, de fet, estretament vinculades, de manera que el resultat de l'aplicació de l'una pot conduir a la necessitat d'aplicar novament l'altra. Per exemple, és possible que la regularització metodològica i la reformulació de l'operació 3 impliqui que es produeixin casos susceptibles de ser tractats per 1 o 2 (que hi hagi mots que desapareguin del conjunt de definicions o que algun sentit deixi de ser

utilitzat). Les operacions 1 i 2, per tant, són de fet operacions cícliques que s'apliquen sobre resultats d'aplicacions prèvies, fins que el conjunt resultant no té redundància.

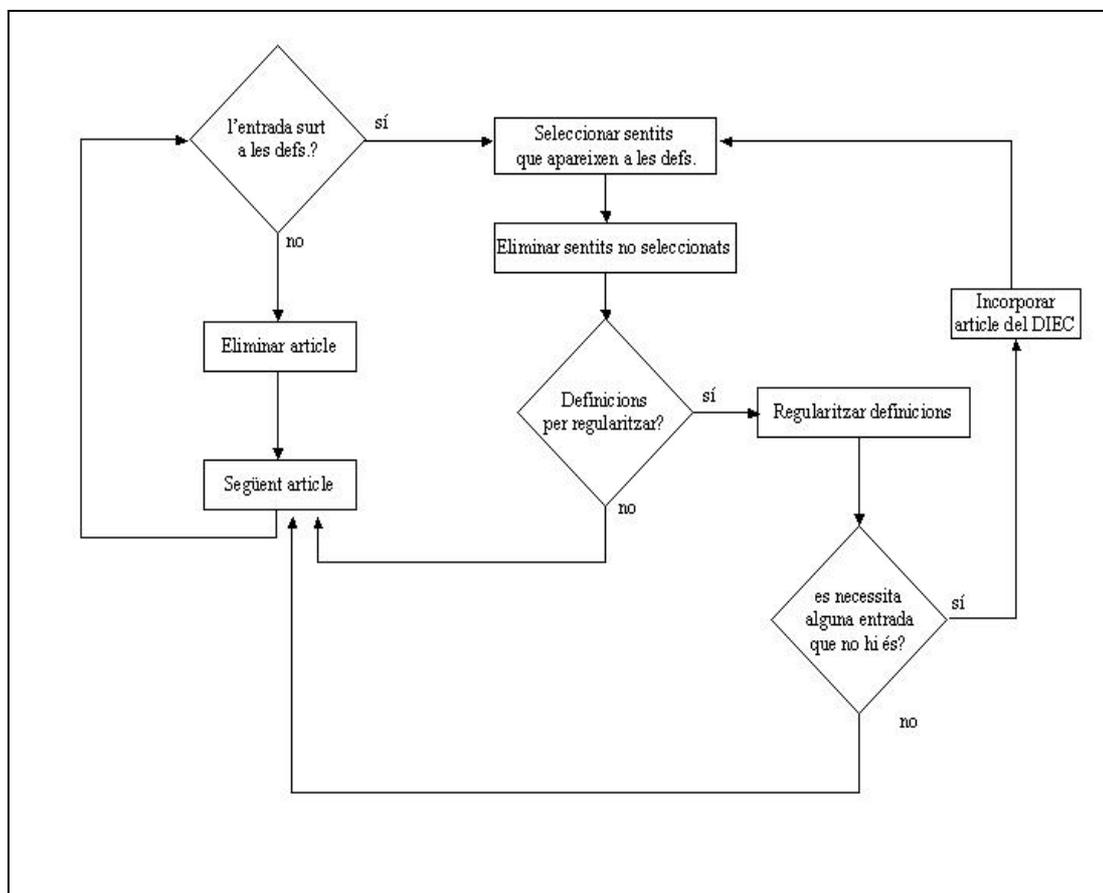


Figura 7.1. Procés d'obtenció del vocabulari definidor.

El procés acaba quan es compleixen de manera conjunta les següents condicions:

1. Totes les definicions del vocabulari estan redactades amb els mots (i en una de les accepcions corresponents d'aquests) que formen part del conjunt del vocabulari.
2. Totes les entrades del vocabulari apareixen almenys un cop en alguna de les definicions regularitzades del conjunt.
3. Totes les definicions han estat regularitzades.

Així, aquest vocabulari es pot constituir en un repertori de descripció que conforma el nivell primari de la metallengua de descripció semàntica d'un diccionari. A § 7.6 dono la informació detallada de la manera com aquest conjunt s'integra en un sistema de definició global per a la totalitat del lèxic del diccionari.

El procés de determinació de la metallengua de definició pot il·lustrar-se a través d'un exemple concret. L'article corresponent al mot **fita** *f.* fou eliminat del vocabulari pel fet que aquest mot no apareix en cap de les definicions de mostra:

fita *f.* Pedra o altre senyal clavat a terra que assenyala el límit d'una heretat, d'una contrada, etc., d'una distància a recórrer.

Aquesta eliminació correspon de fet a la primera revisió sistemàtica. El motiu pel qual aquest mot no figura entre el vocabulari definidor proposat és, doncs, que ja en el creuament inicial entre el vocabulari definidor de partida i les seves definicions, preses directament del DIEC (sense cap anàlisi ni regularització complementària), es va detectar com a no necessari per a la descripció semàntica de cap dels elements de la llista. Aquest no era el cas, en principi, del mot *contrada*, que apareix a la definició de *fita* i que té el seu article corresponent en la mostra de partida:

contrada *f.* Extensió territorial determinada per límits naturals o polítics.

Contrada apareixia en les següents definicions a més de la de *fita*:

accent *m.* (...) Entonació particular, manera de pronunciar peculiar dels habitants d'una contrada. *Parlar amb accent empordanès. Accent castellà.*

bestiar *m.* Conjunt de bèsties d'una mateixa espècie, esp. de les que es crien per al servei i la utilitat de l'home; conjunt de totes les que es crien en una casa, en una contrada, etc. *S'hi fa molt de bestiar, en tal contrada. Bestiar banyut. Bestiar cabrum, boví. Bestiar de llana. Bestiar de peu rodó.*

capital *adj.* (...) *f.* Ciutat o vila on resideix el govern; ciutat o vila principal d'una contrada. *París és la capital de França. Una capital populosa. Anar a la capital.*

districte *m.* Qualsevol de les demarcacions en què es divideix una contrada, una província, una població, per facilitar-ne l'administració, el govern, etc.

exòtic -a *adj.* Introduït d'una contrada estrangera. *Una planta exòtica. Una droga exòtica.*

localitat *f.* Part circumscripita, punt, d'una contrada, d'una regió. *És una localitat malsana.*

migdia *m.* (...) Sud; part meridional d'una contrada, d'un continent, etc. *El migdia de França.*

muntanya *f.* (..) Part muntanyosa d'una contrada (per oposició a *plana*). *Anar a muntanya. Criat a muntanya.*

país *m.* (...) PER EXT. Contrada. *Vins del país.*

racó *m.* (...) Lloc apartat, ocult, d'una casa, d'una ciutat, d'una contrada, etc. *Un casalt ple de racons i amagatalls. En un racó de la serra.*

ratlla *f.* (...) Límit d'una contrada, d'un país, etc. *A la ratlla de França.*

terra *f.* (...) Porció de la superfície de la Terra considerada en ella mateixa, o com a pertanyent a un individu, a un poble; contrada, país; pàtria. *Les terres àrtiques. Ve de llunyes terres. A la meua terra no hi ha aquest costum. Fruites de la terra. Cada terra fa sa guerra. Terres de llengua catalana: Catalunya, València, les Balears...*

El mot *contrada*, malgrat que apareix en les definicions precedents, també ha estat eliminat del vocabulari proposat. La causa de l'eliminació de *contrada*, tot i que és més complexa pel que fa als seus precedents que la de *fita*, és també que en algun punt del cicle d'adaptació del vocabulari definidor ha passat a ser un mot no utilitzat en cap de les definicions de la mostra revisada. Això s'ha produït per motius diversos:

1. L'entrada corresponent a l'article que conté *contrada* no apareix en cap de les definicions de la mostra: el mot *racó* no ha estat usat en cap definició de la mostra de partida, per la qual cosa (com en el cas de *fita*) s'elimina de la mostra l'article que li correspon.
2. El sentit a què es refereix la definició que conté *contrada* no apareix en cap de les definicions de la mostra. Aquest és el cas d'*accent*, *capital*, *migdia*, *muntanya*, *ratlla* i *terra*. Els sentits en què apareix el mot *contrada* no han estat seleccionats. Per tant, s'eliminen de la mostra les definicions corresponents i, amb aquestes, les aparicions corresponents del mot *contrada*.
3. El sentit a què es refereix la definició que conté *contrada* ha estat seleccionat, però les definicions en què apareix han estat regularitzades i el mot en qüestió ha estat

eliminat de la definició. Aquest és el cas de *districte*, que apareixia a la definició de *registre*, però que ha estat eliminada en el cicle de regularització de les definicions.

4. El sentit a què es refereix la definició que conté *contrada* és una definició sintètica i, per tant, es considera preferent el seu sinònim. Aquest és el cas de *país*.

Els casos de *bestiar*, *exòtic*, i *localitat*, presenten un grau de complexitat més elevat. El motiu és que els sentits corresponents d'aquests mots (o bé tot l'article) presenten, al seu torn, motius que han portat a la seva eliminació i, en conseqüència, al fet que *contrada* desaparegués de la mostra. El mot *bestiar* apareix en els següents sentits no seleccionats:

barca *f.* (...) Transbordador fluvial de mercaderies, de vehicles i de bestiar en indrets on no hi ha pont.

cort *f.* (...) Habitació, tancat, destinat al bestiar. *La cort dels bous. Una cort de porcs.*

cria *f.* (...) Cada nou cap de bestiar en un ramat durant el temps que mama. *Té un ramadet de poques cries.*

engegar *v. tr.* (...) [2] Aviar (el bestiar).

espatlla *f.* (...) Qualsevol dels dos quaters davanters del bestiar boví, oví, cabrum i porcí.

mà [pl. *mans*] *f.* (...) Qualsevol dels quatre peus del bestiar de l'escorxador un cop tallats.

passatge *m.* (...) Exacció pel trànsit de bestiar o d'esclaus.

picar *v. tr.* (...) Un gos de pastor, mossegar un cap de bestiar per tal de fer-lo fugir d'un lloc o de fer-lo tornar al ramat.

picar *v. tr.* (...) El pastor, tirar una pedra (a un cap de bestiar) per tal de fer-lo fugir d'un lloc.

senyal *m.* (...) Tall, osca o forat que hom fa a l'orella d'un cap de bestiar per reconèixer-lo i evitar que es barregi amb els animals d'un altre ramat.

El mot *bestiar* també ha estat eliminat en la fase de regularització de les següents definicions:

menar *v. tr.* Conduir o portar (esp. bestiar, un vehicle de tracció animal).

palla *f.* Tija de blat, d'ordi, de sègol i d'altres grans, seca i separada del gra; col·lectivament, les tiges de cereals, esp. de blat, després de la batuda, emprades com a farratge, per a fer el llit del bestiar, per a embalar, per a emplenar, etc.

Un dels articles on apareixia *bestiar* ha estat eliminat pel fet que l'entrada corresponent no apareixia a cap definició:

corral *m.* Lloc tancat, en les cases o en el camp, esp. el destinat a tenir-hi bestiar.

Pel que fa a la resta de definicions en què apareixia el mot *bestiar* (i que corresponen a les entrades **ramat** *m.*, **pastor** –*a m. i f.*, i **pasturar** *v. intr.*) es dona el cas que les úniques aparicions d'aquests mots en definicions d'entrades de la mostra es refereixen als casos eliminats més amunt, o bé són recurrents respecte d'aquest grup, per la qual cosa han de ser eliminats els següents sentits:

ramat *m.* Reunió d'un cert nombre de caps de bestiar que es crien i es fan pasturar plegats.

pastor **pastora** *m. i f.* Persona que mena bestiar a pasturar.

pasturar *v. intr.* El bestiar, menjar l'herba dels camps.

pasturar *v. intr.* (...) || tr. Pasturar el bestiar l'herba fresca dels prats.

pasturar *v. intr.* (...) || tr. Portar a pasturar.

Per altra banda, el mot *exòtic* apareix als articles de **palmera** *f.* i de **diamant²** *m.*:

palmera *f.* Arbre exòtic (*Phoenix dactylifera*), de fulles coriàcies pinnatipartides i de baies oblongues i comestibles, anomenada també palmera de dàtils. | Nom donat a una gran diversitat d'espècies de la família de les arecàcies o palmàcies.

diamant² *m.* Nom de diversos ocells exòtics de la família dels estríldids, de mida petita i colors vius.

Pel que fa a **diamant²** *m.*, l'article desapareix pel fet que correspon a una entrada no utilitzada en cap definició de la mostra. Per altra banda, *palmera* només apareix a la definició sintètica de *palma* i, per tant, s'elimina pel fet que es prefereix el cap de sèrie.

Finalment, el mot *localitat* apareix a les següents definicions:

etapa *f.* Localitat on les tropes en marxa s'aturen per passar la nit; distància a recórrer per a arribar a l'etapa.

indret *m.* Lloc determinat, que hom considera, en una localitat, en un objecte; part determinada del cos; passatge determinat d'una obra.

lliura *f.* Unitat de pes de valor molt variable segons els països i les localitats.

missió *f.* (...) || Sèrie continuada de prèdiques, exercicis, etc., que es fa durant alguns dies, sovint per sacerdots d'una altra localitat.

ordinari -ària *adj.* (...) || *m.* (...) | Individu que va i ve regularment d'una localitat a una altra, i fa les comissions que li encomanen. ||

ple plena *adj.* (...) || *m.* (...) | Concurrencia que ocupa totes les localitats d'un teatre, tots els seients d'un vehicle, etc.

punt *m.* (...) | Indret geogràfic, localitat.

Els sentits corresponents d'*etapa*, *indret*, *missió*, *ordinari*, *ple* i *punt* no han estat seleccionats i per tant s'han eliminat. Per altra banda, *lliura*, no apareix en cap definició de la mostra, i per tant s'ha eliminat l'article corresponent. Finalment, la definició d'*indret* s'ha regularitzat i el mot *localitat* n'ha desaparegut.

Com a resultat de totes aquestes operacions, es conclou que cap dels casos en què apareixia el mot *contrada* podia considerar-se primari per a la descripció lexicogràfica. En alguns casos, les verificacions que cal dur a terme transcendeixen l'àmbit dels articles en què apareix el mot *contrada* i cal recórrer a l'aplicació de cicles d'anàlisi més complexos. Però, com podem veure, els criteris són simples i s'apliquen sempre en funció de les mateixes condicions.

A continuació analitzaré amb una mica de detall quines han estat les implicacions de l'aplicació cadascun d'aquests processos considerats particularment.

7.3 Eliminació d'articles innecessaris

El criteri que s'ha aplicat per a l'eliminació d'articles innecessaris és el següent:

A és un article innecessari si l'entrada corresponent A' representa un lexema que no apareix, sota cap forma flexional o variant apreciativa, en cap definició lexicogràfica de la mostra de diccionari.

El resultat de l'aplicació d'aquesta condició en cicles successius d'operació ha reduït considerablement el conjunt total del vocabulari. El procés d'eliminació d'articles s'ha fet en dues fases:

a) eliminació dels articles corresponents a entrades que no s'utilitzen en la definició (o definicions) de cap element del vocabulari, i

b) eliminació dels articles corresponents a entrades que no s'utilitzen en les definicions regularitzades (veg. més endavant § 7.5).

Mentre que la fase *a)* s'ha aplicat un sol cop a partir del contingut del fitxer inicial corresponent a les entrades preseleccionades com a vocabulari, la fase *b)* ha estat objecte d'una aplicació cíclica sobre els resultats obtinguts de les operacions descrites a § 7.4 i § 7.5.

La fase *a)* ha portat a eliminar del vocabulari mots que presenten un grau d'especificitat elevat, en el sentit que no s'utilitzen en les descripcions semàntiques de cap unitat. Així, s'han eliminat del vocabulari mots com ara:

franc *m.* Unitat monetària de diversos països europeus (França, Bèlgica, Luxemburg, Suïssa, etc.) i africans (Burundi, Congo, Costa d'Ivori, Txad, etc.).
Una peça de 20 francs. | Moneda d'un franc. || Nom donat a diferents monedes medievals franceses d'or i d'argent.

Complementàriament a aquesta especificitat conceptual de l'element eliminat, s'han bandejat també els casos en què es produeixen connotacions (generalment de tipus negatiu) estretament associats al mot definit i al seu significat conceptual. En general, aquest cas no representa l'aplicació d'un criteri de selecció específic, ja que aquest tipus de mots (precisament a causa de la seva connotació) no solen formar part de cap definició lexicogràfica, encara que el diccionari no recorri a tècniques de restricció de vocabulari:

xerrar *v. intr.* Parlar molt i sense substància, pel sol gust de parlar, fora de propòsit. *Es passa tot el dia xerrant amb el veïnat. Xerra més que setze.* || Determinats ocells, emetre llurs crits insistents. *Els pardals xerraven a son lloure.*

|| *tr.* Dir, contar (alguna cosa), indiscretament, mancant a la reserva deguda.
D'això, no li'n diguis res, que ho xerraria.

També com a resultat de l'aplicació de la fase *a)* d'aquesta operació he eliminat homògrafs que no es troben a les definicions del diccionari; en alguns casos la presència d'aquests homògrafs és conseqüència del fet que la mostra de què partia incorporava tots els articles que es corresponien amb alguna de les formes gràfiques seleccionades prèviament, sense fer una selecció categorial que lligués els codis de lema del CTILC (etiquetari utilitzat en el marcatge de les unitats lèxiques de la proposta de vocabulari) i els codis utilitzats al DIEC en els seus articles lexicogràfics:

beure *m.* Beguda (acc. 1).

lleu *m.* Pulmó, freixura.

lleu *adv.* Aviat.

sòlid *m.* Unitat monetària romana d'or, continuada pels bizantins i imitada per alguns pobles bàrbars.

lluny -a *adj.* Llunyà.

mel *m.* Os de la galta.

En alguns altres casos aquest mateix tractament s'ha aplicat a homògrafs categorials de mots que han restat a la metallengua de descripció. Com a resultat d'això s'han eliminat articles del tipus:

marca *f.* País de frontera. *Marca d'Ancona. Marca Hispànica.*

mica *f.* Qualsevol dels minerals del grup de les miques. || *pl.* Grup pertanyent als fil·losilicats monocònics de composició molt variada, els quals són constituents principals de les roques ígnies i metamòrfiques, i es caracteritzen perquè se separen en làmines transparents més o menys elàstiques.

potà *f.* Canana².

roca *f.* En els entremesos de la processó del Corpus, escenari mòbil o plataforma transportable que sosté elements decoratius, tramoies i figures o actors que representen escenes hagiogràfiques, bíbliques o al·legòriques.

suc *m.* Cop pegat amb el cap.

sucre *m.* Unitat monetària de l'Equador. | Moneda d'un sucre.

tasca *f.* Barra de sorra a l'entrada d'un port.

coll *m.* Depressió a la carena d'una serralada o d'un contrafort, per la qual se sol passar per anar d'un vessant a l'altre.

coll m. Qualsevol de les sèries en què es divideix una baralla de cartes. || Símbol que identifica una carta amb la seva sèrie.

Un cop eliminats els articles afectats per la fase *a)* he aplicat el cicle d'operacions descrit a § 7.2, que integra els processos presentats en aquesta secció i a § 7.4 i § 7.5. El conjunt de tots aquests processos representen la fase *b)* descrita més amunt. La regularització de les definicions, en la pràctica, fou portada a terme un cop vaig haver aplicat tot el procés de selecció de sentits a tots els articles del diccionari. Com es pot deduir de l'exemple de *fita* que he mostrar anteriorment, tant la selecció de sentits com la regularització de definicions han suposat l'eliminació d'un nombre molt considerable d'articles lexicogràfics seleccionats en la mostra inicial. L'alt nivell d'interdependència entre els elements del vocabulari i les fórmules de definició dels articles fa que l'aplicació de l'operació d'eliminació d'articles tingui un caràcter cíclic, lligat sempre a altres processos de modificació dels articles corresponents al vocabulari definidor.

7.4 Selecció de sentits

El criteri que s'ha aplicat per a la selecció de sentits és el següent:

El sentit S de l'article P se selecciona si el lexema P' apareix, sota qualsevol forma flexional o variant apreciativa, amb el sentit S en una o més definicions lexicogràfiques de la mostra de diccionari.

Els sentits no seleccionats s'eliminen.

La selecció de sentits s'ha fet per referència al sentits utilitzats en el mateix conjunt seleccionat. No he aplicat, per tant, criteris que poden resultar subjectius a la pràctica (com ara la selecció del sentit més "corrent", més "bàsic", o del primer sentit del diccionari). Una possible selecció de sentits basada en els criteris d'ordenació i

marcatge de les accepcions en el DIEC hauria conduït a resultats força diferents. Sovint, per exemple, la primera accepció d'un mot en un article del DIEC no conté el sentit més utilitzat en el mateix diccionari:

observar v. tr. Seguir exactament (allò que està prescrit). *Observar una regla. Observar una llei*, allò que prescriu la llei. *Observar les prescripcions del metge. Observar els dejunis.* || Mirar amb una atenció seguida. *Observar els passos d'algú. Observar els moviments de l'enemic. Observar els estels amb el telescopi.* | Notar, advertir. *He observat que no arriba mai a l'hora. Li vaig fer observar que el compte era equivocat.* || Servir-se dels instruments de fer les observacions nàutiques.

produir v. tr. Fer aparèixer, exhibir, esp. en estil forense. *Produir testimonis.* | pron. Presentar-se, explicar-se, captenir-se davant la gent, de tal o tal manera. *Produir-se un actor (en públic).* || Generar, donar naixença (a algú o a alguna cosa); causar que (quelcom) sigui, s'esdevingui, originar; manufacturar; fornir (un guany). *La terra produeix herba, fruits. Les glàndules que produeixen la saliva. Catalunya produeix bons vins. És una finca que no li produeix gaire.*

Per a cada un dels mots del vocabulari, doncs, he verificat les seves aparicions en les definicions lexicogràfiques de la mostra, i he seleccionat els sentits que més s'ajustaven al significat amb què apareixia en les definicions.

En general, la selecció de sentits ha donat com a resultat sentits no marcats específicament amb cap de les *marques d'ús* que utilitza el DIEC per a les seves definicions. En alguns casos, però, el resultat d'aquesta operació ha comportat seleccionar sentits que el DIEC recull marcats com a *figuradament (fig.)*, com en el cas d'*aconseguir*, *ardent* i *fonament*, que apareixen al diccionari de forma recurrent amb els sentits que indico:

aconseguir v. tr. Arribar caminant, corrent, nedant, etc., a ajuntar-se (a una persona o a una cosa que va al davant). *Si no camineu més de pressa, no els podreu pas aconseguir. Si no us torbeu pel camí, els aconseguireu abans d'arribar a la collada.* || Empaitar. || Arribar a tocar llançant un projectil, ventant un cop. || Haver ja nascut o no haver encara mort en l'època (d'un esdeveniment), en l'època en què vivia (una persona). *Havia aconseguit la Guerra dels Set Anys. Jo havia aconseguit el seu besavi.* | Arribar a tenir (una edat determinada). *Aquell vell ha aconseguit la vuitantena.* || FIG. Algú, arribar a fer (allò que es proposa), a obtenir (allò que desitja). *Volien obrir la porta, però no ho aconseguiren. No aconseguireu pas que*

desisteixi. Mai no aconseguireu el seu consentiment. Aconseguir honors, riqueses, aplaudiments, l'aprovació.

ardent *adj.* Que crema, incandescent. *Carbons ardents.* | FIG. Ple de foc, de passió. *Inspiració ardent.*

fonament *m.* Basament de les parets i dels pilars d'un edifici que estan sota terra i transmeten al terreny el pes de la construcció. *Posar els fonaments d'un edifici. (...)*

|| FIG. Part essencial sobre què recolzen les altres; allò sobre què recolza un judici, una apreciació, etc. *L'educació és el fonament de la civilitat. És una opinió sense fonament. Aquesta sospita no té cap fonament.*

Algun dels sentits seleccionats també està, en el DIEC, marcat amb el codi de restricció de significat *especialment (esp.)*, com en el cas d'*amagar* i *harmonia*; o amb el seu corresponent d'extensió de significat *per extensió (p. ext)*, com en el cas de *braç*.

amagar *v. tr.* Sostreure a la vista, impedir de veure. *Aquell núvol ens amaga la lluna.* | ESP. Posar (quelcom) en un lloc retret a fi que altri no ho vegi o li sigui difícil de trobar-ho. *Vaig amagar la clau darrere la cortina, entre els papers. Amagaren els diners sota terra, sota una pedra. Amagaren el fugitiu en un armari.*

harmonia *f.* Justa adaptació, entre elles, de les parts o coses que formen un tot; concordança d'opinions, de maneres, d'interessos, etc., acord, bona correspondència, entre dues o més persones. *L'harmonia de les diverses parts d'un organisme. L'harmonia del cos humà. Dos amics que viuen en bona harmonia. Sempre han viscut en bona harmonia.* || Acord de sons agradables a l'orella. *Cants plens d'harmonia.* | PER EXT. Successió de sons, de mots, que produeixen un efecte agradable a l'orella. *L'harmonia d'un vers, d'un període. Harmonia imitativa.* | ESP. Estructura d'una peça musical en allò que es refereix a la composició, progressió i modulació dels acords (en contraposició a melodia i ritme); ciència de l'estructura, progressió i modulació dels acords.

braç *m.* Part compresa entre l'espatlla i el colze (en oposició a avantbraç). | PER EXT. Qualsevol de les dues extremitats toràciques de l'home, esp. la part compresa entre l'espatlla i la mà.

La presència d'aquestes marques, en realitat, no ha suposat durant el procés de selecció l'aplicació de criteris favorables o desfavorables basats en el marcatge del DIEC, tot i que en molts casos, les acepcions marcades han estat eliminades per l'aplicació dels criteris de selecció:

vestir *v. tr.* Posar robes adaptades al cos (d'algú), per cobrir la seva nuesa, per resguardar-lo de la intempèrie. *Vestir una criatura. Està impedit, i l'han de vestir.* |

pron. La nena ja es vesteix sola. Vestir-se de pressa. | Vestir-se de frac. (...) || PER ANAL. Posar una coberta, ornaments, etc., (a una cosa). Vestir una ampolla amb palla. Vestir l'altar. (...) || PER ANAL. Cobrir amb roba els corròns d'algunes màquines d'aprest per tal de no produir una pressió massa dura, sinó una mica elàstica, sobre el producte tèxtil que es fa passar entre ells.

agafar *v. tr. (...) | FIG. Aprofitar quelcom com a pretext o ocasió, valer-se d'algú per salvaguardar-se, per reeixir en una empresa. S'agafaran en això per no concedir-nos-ho. Agafa't per aquí, que no cauràs. Vaig agafar-me al comandant, i així ho vaig aconseguir.*

arma *f. (...) | FIG. Mitjà de què hom se serveix per a vèncer, per a aconseguir alguna cosa. La raó i el dret no són sempre armes suficients.*

Hi ha, a més, altres marques d'ús en el DIEC que indiquen informacions sobre el registre, la vigència, la connotació, etc. que solen impedir l'aparició en les definicions dels sentits marcats:

arreplegar *v. tr. (...) | POP. Ésser atacat (per una malaltia). El meu cosí ha arreplegat una pulmonia.*

cavall¹ *m. (...) || POP. Heroïna.*

còmic² -a *adj. (...) | OBS. Actor.*

composició *f. (...) | ABS. Disciplina teòrica de la música que ensenya les regles per a compondre. Fa anys que deixà els estudis de composició.*

composició *f. (...) || ANT. Conveni o esmena que, d'acord amb el sistema penal de la venjança privada, restableix la pau que ha pertorbat una ofensa criminal entre persones o entitats mitjançant les indemnitzacions que hom considera suficients.*

comte comtessa *m. i f. (...) | ANT. En el règim feudal, senyor amb jurisdicció sobre un comtat.*

comú -una *adj. (...) m. (...) | ANT. Imposició general extraordinària proporcional als béns posseïts pagada per tots els veïns de determinats comuns o municipis catalans per tal d'equilibrar l'actiu i el passiu municipals.*

barra *f. (...) | ABUS. Pal o faixa. Les barres catalanes. La bandera de les quatre barres.*

mil *adj. (...) || HIPERB. Un nombre indeterminat considerable. Mil gràcies! Sempre troba mil excuses. T'ho he dit mil vegades!*

munt *m. ARC. Muntanya (emprat sobretot en la formació de noms geogràfics). Sant Llorenç del Munt.*

donar (...) *v. tr. (...) || EN REC. Lliurar (una persona) a algú. Va donar-me la seva filla en matrimoni.*

jutge jutgessa *m. i f. (...) || f. EN REC. Muller d'un jutge.*

pla² plana *adj.* (...) || *adv.* **pla [o pla bé]** Prou. *Ell pla ha fet tal cosa. Ell se n'és pla bé guardat.* | IRÒN. Expressa negació. *Ell pla ho farà.*

riure¹ (...) *v. intr.* (...) | IRÒN. Començar a rompre's la tela, les sabates, etc., a conseqüència de l'ús excessiu o de la mala qualitat del material.

accent *m.* (...) | *pl.* POÈT. Veu, cant, llenguatge. *Els dolços accents del rossinyol.*

acer *m.* (...) | POÈT. Arma d'acer. *L'acer homicida.*

costella *f.* (...) || VULG. Qualsevol dels llistons dels quals es compon el dibuix en els telers de maquineta.

escórrer (...) *v. tr.* (...) | VULG. Tenir un orgasme.

Finalment, cal fer notar, tal com es pot veure a l'exemple de *fita* presentat a § 7.2, la implicació que representa l'operació de selecció de sentits amb la d'eliminació de sentits innecessaris. L'eliminació de tal o tal sentit d'un article ha comportat sovint la desaparició en el conjunt de definicions de tal o tal entrada, de manera que això ha provocat l'eliminació subsegüent en l'article referit a aquesta entrada.

L'operació de selecció de sentits presenta també implicacions amb l'operació de regularització de definicions. En alguns casos, la fórmula que s'ha triat per a la regularització de tal o tal definició ha conduït a la desaparició, en el conjunt de definicions, de tal o tal sentit seleccionat per a un determinat mot. Això ha provocat consegüentment l'eliminació d'aquest sentit (que havia estat prèviament seleccionat) en l'article corresponent.

7.5 Regularització de les definicions

El criteri que s'ha aplicat per a la regularització de les definicions és el següent:

La definició D del sentit S de l'article P està regularitzada si: (i) satisfà totes les condicions del sistema de definició i (ii) conté exclusivament en la seva definició mots que formen part del conjunt d'entrades de la mostra de diccionari, en algun dels sentits descrits en la mostra.

Paral·lelament a l'adequació de les definicions dels elements del vocabulari, la regularització ha representat, en termes generals, la normalització de les definicions pel que fa als problemes que aquestes presenten des del punt de vista de la tècnica lexicogràfica i que es concreten en els següents casos principals:

a) Parts ostensives (o extensionals, inclosa la utilització d'*etcètera* a les definicions); el DIEC, com ja he exposat anteriorment, conté nombrosos casos d'elements ostensius, habitualment en forma de "l·listes" d'elements dins de les seves definicions. Aquest fet, que des del punt de vista del mètode lexicogràfic pot ser considerat com una desviació (en el sentit que no descriu estrictament el significat conceptual del definit, sinó que es refereix a un o més elements amb què el seu referent pot tenir una determinada relació), té a més una incidència directa sobre el nombre d'unitats en el corpus de definicions. La substitució d'aquests elements ostensius (que sovint contenen l'expressió *etcètera*) ha permès una notable reducció del nombre d'unitats necessari per a la descripció del conjunt de mots. Així, per exemple, definicions del tipus:

tira *f.* Tros llarg i estret de roba, cuir, paper, etc. [DIEC]

temps [inv.] *m.* (...) Forma mental que assumeix l'experiència de la durada i de la successió, el temps essent concebut com a transcorrent contínuament i uniformement i mesurat per fenòmens successius que ocorren a intervals regulars (com la revolució aparent de l'esfera celeste causada per la rotació de la Terra). [DIEC]

Passen a tenir la forma:

tira *f.* Tros llarg i estret d'alguna cosa de poc gruix.

temps *m.* (...) Forma mental que es dóna a l'experiència de la durada i de la successió de les coses, mesurat per fenòmens successius que s'esdevenen a intervals regulars.

En determinats casos hi ha, però, elements ostensius que s'han de deixar a la definició, si es vol evitar una caracterització no estrictament terminològica o científista; així, definicions com ara:

negre -a *adj.* Del color del sutge o del carbó, completament fosc. [DIEC]

blanc -a *adj.* Del color de la neu, que reflecteix tots els raigs de l'espectre combinats. [DIEC]

Passen a tenir la forma:

negre -a *adj.* Del color del carbó.

blanc -a *adj.* Del color de la neu.

b) Elements de metallengua de signe: *dit de...*, *nom donat a...*, etc.

combustible *adj.* Dit de les substàncies capaces de combinar-se amb l'oxigen per tal de ser cremades. [DIEC]

fill filla *m. i f.* Nom que es dona a una persona o animal amb relació al seu pare i a la seva mare. [DIEC]

Passen a tenir la forma:

combustible *adj.*(...) [Substància] capaç de combinar-se amb l'oxigen per tal de ser cremada i produir energia.

fill *m.* (...) [~ d'X] Algú de qui X és pare o mare.

c) Utilització de la primera persona com a marca de coreferència:

interès [pl. -essos] *m.* (...) || Sentiment que alguna cosa desvetlla en nosaltres, el qual ens mou a prestar-li una atenció especial, a ésser-hi favorables o desfavorables. [DIEC]

objecte *m.* Cosa material en tant que s'ofereix als nostres sentits. [DIEC]

senyal *m.* (...) Cosa sensible de la qual ens servim per a denotar quelcom, per a indicar el moment de fer quelcom. [DIEC]

voluntat *f.* Facultat de voler, potència de l'ànima que ens mou a fer o a no fer una cosa. [DIEC]

Passen a tenir la forma:

interès *m.* (...) [~ d'X per Y] Sentiment² que Y provoca en X, el qual fa que X manifesti una atenció especial per Y.

objecte *m.* (...) Cosa material en tant que s'ofereix als sentits¹.

senyal *m.* (...) Cosa sensible que serveix per a expressar quelcom.

voluntat *f.* (...) [~ d'X] Estat¹ de la ment d'X que mou X a fer o a no fer una cosa.

d) Utilització de fórmules de caracterització quantitativa del referent: *cadascun*, *qualsevol*, etc.

xifra *f.* Qualsevol dels signes que serveixen per a representar els nombres en un sistema de numeració. [DIEC]

Passa a tenir la forma:

xifra *f.* (...) Signe que serveix per a representar¹ un nombre.

e) Reducció de les definicions múltiples a definicions simples.

diferenciar *v. tr.* Fer diferent; separar o distingir (una cosa) d'una altra constatant-ne les diferències. [DIEC]

econòmic -a *adj.* Relatiu o pertanyent a l'economia. [DIEC]

Passen a tenir la forma:

diferenciar *v.* (...) [X ~ Y de Z] X distingir Y de Z examinant-ne les diferències².

econòmic *adj.* (...) Relatiu a l'economia.

Una de les conseqüències directes del procés de regularització d'accepcions ha estat la incorporació a la mostra de diccionari dels mots que, sent necessaris per a la descripció d'aquests elements, no havien estat incorporats en els processos de selecció freqüencial. En molts casos, la presència en les definicions de mots que no pertanyien a la nomenclatura de la mostra de diccionari s'ha pogut evitar mitjançant paràfrasis o a través dels criteris de regularització aplicats. Quan això no ha estat possible, he incorporat els articles corresponents del DIEC que, al seu torn, han estat sotmesos als processos descrits anteriorment.

En la llista d'aquests casos, que dono a continuació, incorporo en negreta els mots que ja eren inclosos en la llista de freqüència del DIEC (veg. Apèndix I.3), però que no es van seleccionar per a la mostra de partida perquè eren exclusius d'aquesta procedència:

accessible , <i>adj.</i>	constituent , <i>adj.</i>	model , <i>m.</i>
acompliment , <i>m.</i>	corb , <i>adj.</i>	pertorbar, <i>v.</i>
arbust , <i>m.</i>	elàstic , <i>adj.</i>	preferir, <i>v.</i>
arranjament , <i>m.</i>	extremitat , <i>f.</i>	receptacle , <i>m.</i>
arranjar , <i>v.</i>	favor , <i>m.</i>	recipient, <i>m.</i>
cavitat , <i>f.</i>	fermentació , <i>f.</i>	rectangle , <i>m.</i>
cessació , <i>f.</i>	filament , <i>m.</i>	rectangular , <i>adj.</i>
cilindre , <i>m.</i>	fluid , <i>m.</i>	rovellar, <i>v.</i>
cilíndric , <i>adj.</i>	grandària , <i>f.</i>	tija , <i>f.</i>
circumferència , <i>f.</i>	llenyós , <i>adj.</i>	vespre, <i>m.</i>
concernir , <i>v.</i>	mamífer , <i>adj.</i>	
conducte , <i>m.</i>	mòbil , <i>adj.</i>	

La majoria dels mots incorporats en el procés de regularització d'accepcions figuraven a la llista de mots més freqüents de les definicions del DIEC, la qual cosa confirma la seva rellevància des del punt de vista de la descripció lexicogràfica i contribueix a avalar l'adequació de la mostra resultant.

7.6 Integració en un sistema de definició

La sèrie d'operacions aplicades sobre el vocabulari de definició i la informació corresponent del DIEC han donat com a resultat una metallengua de definició presentada en forma de diccionari (veg. Apèndix II) que reuneix les següents propietats:

- a) Qualsevol element del vocabulari definidor és usat almenys en una de les definicions.
- b) Cadascun dels sentits recollits en un article és usat almenys un cop en una de les definicions.

c) El conjunt de definicions no conté elements lexemàtics que no estiguin recollits com a entrades.

d) Els elements lèxics de les definicions no s'usen amb sentits que no estiguin recollits en el seu article corresponent.

Un cop finalitzat el processament per a l'obtenció d'una metallengua de definició, el vocabulari definidor proposat inicialment ha quedat reduït a un conjunt de 1.089 *entrades*, que representen un total de 1.172 sentits (o elements de metallengua) diferents. En tant que metallengua de definició, la dada de referència que indica el total de les seves unitats és el nombre de sentits, que es diferencien en cas que formin part d'una mateixa entrada.

Aquesta operació ha donat com a resultat la base d'una metallengua de definició que conté exclusivament aquells mots que són utilitzats en la definició de les entrades de diccionari que hi corresponen. Des d'un punt de vista tècnic, doncs, aquesta metallengua de definició està integrada per elements *primaris* que poden ser considerats el nucli del sistema descriptiu d'un diccionari general. No he considerat elements centrals del vocabulari els mots gramaticals (les preposicions febles, els pronoms, etc.) i altres tipus de mots (veg. més endavant § 8.2) pel fet que he volgut restringir l'experiment als mots que poden ser definits per paràfrasis substituïbles.

El resultat és, doncs, un inventari lèxic en forma de diccionari i *tancat* pel que fa a les relacions entre la nomenclatura i el corpus de les seves definicions. Com a estructura, aquest inventari no es diferencia fonamentalment d'un diccionari convencional, en el sentit que s'hi produeixen relacions de circularitat entre els elements de la seva metallengua. La característica diferencial de l'inventari proposat és el reduït nombre d'elements que l'integra, i el fet que la tria d'aquests elements s'ha orientat a l'establiment d'una metallengua de definició.

L'objectiu que m'he fixat és integrar aquestes dades en un sistema de descripció del significat en lexicografia per a un diccionari general sense especificitats terminològiques. Les experiències amb vocabularis controlats descrites anteriorment

permeten suposar que la mostra proposada pot ser utilitzada com un vocabulari de definició especialment orientat a la lexicografia. El nombre d'unitats final se situa quantitativament entre certs vocabularis *bàsics* i els vocabularis controlats, amb la particularitat que la seva adequació en lexicografia ha estat contrastada pel procés d'obtenció. Aquest fet queda demostrat per la presència de mots que tenen una utilització recurrent en lexicografia i que no presenten les característiques freqüencials que solen tenir els elements dels vocabularis bàsics.

Les possibilitats d'utilització de l'inventari proposat com a vocabulari controlat, anàleg als que s'han utilitzat en els diccionaris anglesos, queda garantida pel fet que ha estat seleccionat a partir dels mateixos criteris que els vocabularis controlats per a l'anglès. Per tant, es pot suposar que reuneix els elements suficients per a la descripció de tots els sentits de les unitats lèxiques d'un diccionari general, i que el conjunt proposat està en condicions d'assumir la contrastació empírica d'aquesta adequació i servir de base a la construcció d'un diccionari monolingüe complet.

La intenció amb què ha estat concebut aquest inventari de descripció no és, però, la de ser utilitzat a la manera dels vocabularis controlats, com un repertori únic d'elements de descripció disponibles. Ja he mostrat amb anterioritat (veg. § 5.3) que els vocabularis controlats presenten, des del punt de vista de la descripció semàntica i de la projecció estructurada de les relacions lèxiques del diccionari, alguns problemes que són difícilment aïllables del mètode de control del vocabulari: associació de descriptors de nivells alts, definicions inespecífiques, etc.

Un diccionari fet a partir d'un vocabulari controlat projecta una estructura de relacions internes similar a la de la figura 7.2; aquesta figura correspon a un conjunt d'elements $\{A, \dots, Z\}$ que estan exclusivament definits a base de correlats dels elements $\{A, B, C, D, E\}$, que formen el vocabulari de definició.

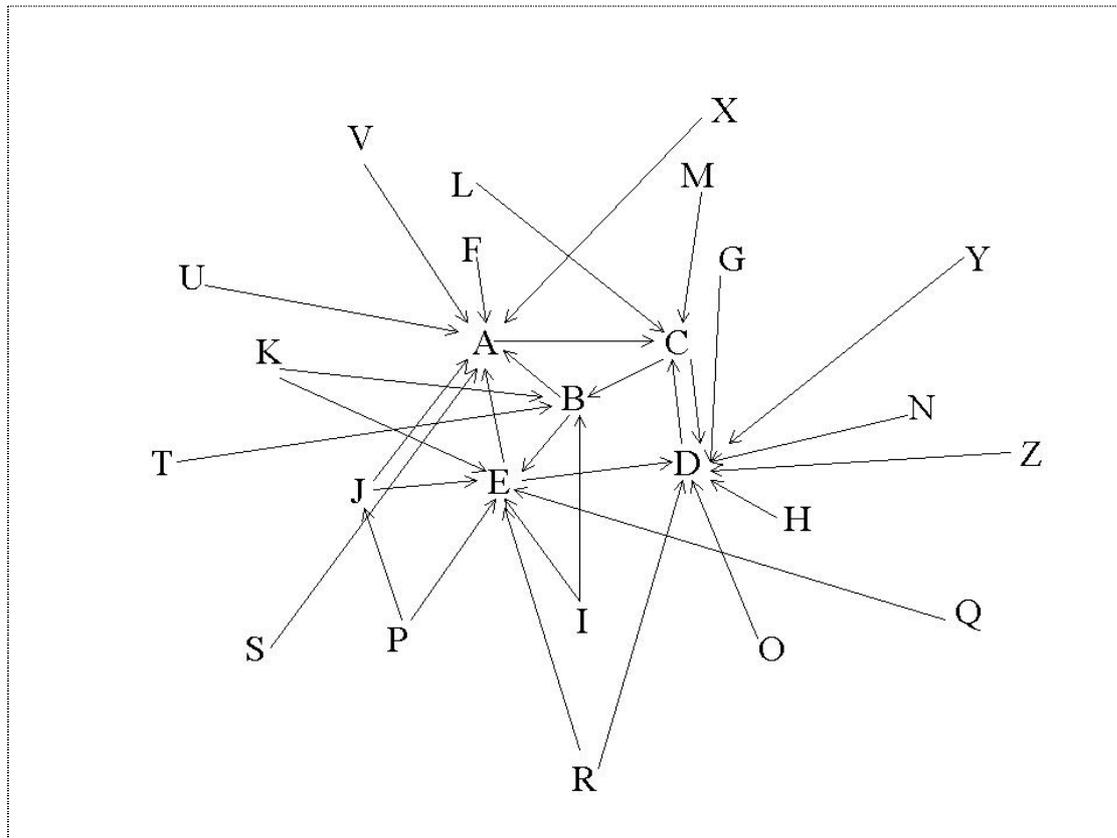


Figura 7.2. Representació d'un diccionari basat en vocabulari definidor controlat.

En la representació gràfica de la figura 7.2 es pot veure com tots els elements del conjunt de descripció hi fan referència directa i com la circularitat queda reduïda al subconjunt establert com a vocabulari de definició. L'alt nivell de recurrència que presenta, però, el vocabulari definidor respecte a la resta d'elements del conjunt té unes conseqüències determinades en la descripció lexicogràfica, que ja he assenyalat parcialment (veg. més enrere § 5.4). Una conseqüència especialment destacable és la indistinció dels nivells inferiors d'estructuració jeràrquica del lèxic i, per tant, la projecció de taxonomies lèxiques inespecífiques.

L'estructura representada en la figura 7.2 correspon a una descripció de la nomenclatura d'un diccionari que limita el seu vocabulari definidor a un conjunt predeterminat d'elements. A la figura 7.3. el subconjunt {A, B, C, D, E} continua concentrant les

sistema estructurat de descripció lexicogràfica que permeti el tractament normalitzat de les estructures que projecta un diccionari i que, alhora, no es vegi determinat negativament per les limitacions descriptives que suposa l'ús d'un vocabulari controlat.

Aquest sistema, que parteix d'una estructura similar a la proposada per Weinreich (1962), es basa en un nivell 0, considerat axiomàtic, en què es permet la circularitat. Successivament, la descripció del significat es fonamenta en una sèrie ordenada de n nivells que no presenten aquest tipus de problemes estructurals o descriptius. La projecció estructural d'aquest sistema es recull a la figura 7.4.

El nivell 0 d'un sistema de definició ha de complir la condició de ser un conjunt d'elements de metallengua caracteritzats circularment i tancat des del punt de vista de les seves unitats de descripció. Per tant, aquest nivell 0 es correspon amb el conjunt que he obtingut a partir de les operacions realitzades sobre la mostra de diccionari de partida.

Per a l'elaboració estructurada d'un diccionari amb un vocabulari de definició a partir del conjunt proposat cal formular l'existència de diversos nivells en el vocabulari definidor i entendre que cadascuna de les definicions del diccionari pot pertànyer a un i només a un nivell de descripció. Així, s'obté una sèrie ordenada de nivells de descripció:

nivell 1: les definicions només contenen mots que pertanyen al nivell 0.

nivell 2: les definicions només contenen mots que pertanyen al nivell 0 o al nivell 1.

nivell 3: les definicions només contenen mots que pertanyen al nivell 0, al nivell 1, o al nivell 2.

...

nivell n : les definicions només contenen mots que pertanyen al nivell 0, al nivell 1, al nivell 2, ... o al nivell $n-1$.

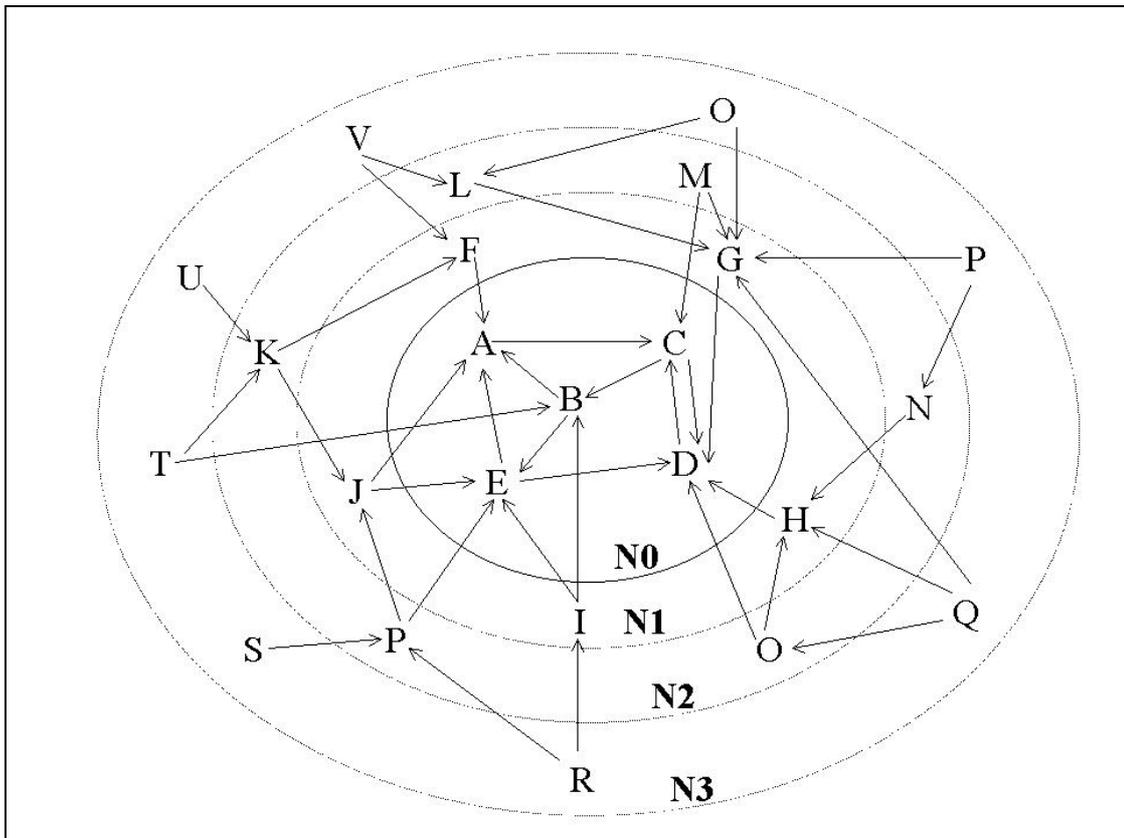


Figura 7.4. Representació d'un diccionari basat en un sistema de definició estructurat en nivells de descripció.

Representada gràficament, una estructura com la que acabo de proposar té la forma de la figura 7.4. Comparant aquesta figura amb la figura 7.3, es pot veure que es tracta d'una mateixa estructura en què simplement s'han marcat els nivells que es refereixen a les restriccions de formulació de definicions que acabo d'assenyalar.

El resultat d'un sistema de descripció d'aquest tipus presenta l'avantatge de reduir al nivell 0 els problemes de circularitat estructural del diccionari i, alhora, de no fonamentar exclusivament la descripció en un conjunt tancat; aquest sistema, de fet, estableix els mecanismes a partir dels quals el vocabulari de definició del diccionari pot projectar-se en funció de les necessitats de descripció semàntica.

Portat a un diccionari concret, aquest sistema assegura que les estructures de relació semàntica que projecta no presenten les desviacions pròpies de la lexicografia convencional i, per tant, no només contribueixen a una major adequació descriptiva del diccionari, sinó que, complementàriament, també en fan una eina de gran valor per a l'extracció d'informació de cara a l'elaboració de lèxics electrònics i de recursos lingüístics en general.

Els mecanismes de control informàtic que ja s'han integrat plenament a la redacció de diccionaris permeten la viabilitat d'un sistema de descripció com aquest, que condueix a la formulació de diccionaris més estructurats.

8 Metallengua i sistema de definició

8.1 Estructura i naturalesa de la metallengua

L'establiment del nivell 0 de descripció i la integració d'aquest en un sistema de definició suposen un avenç fonamental en la determinació d'una metallengua de definició suficientment adaptable a les necessitats descriptives, i no basada exclusivament en un conjunt d'elements limitat apriorísticament sobre bases freqüencials.

La manera com el sistema de definició proposat integra la incorporació d'aquest vocabulari en una metallengua de definició més àmplia fa evidents els paral·lelismes d'aquests nivell 0 amb els sistemes axiomàtics de la lògica (veg. per exemple Sacristán, 1964). Els elements que el conformen són, cadascun, irreductibles en el sentit que formen part del conjunt mínim necessari per a la descripció de les seves mateixes unitats; d'altra banda, aquests mateixos elements són la base que permet la formulació organitzada de totes les descripcions del diccionari, en la forma que he descrit a § 7.6.

Seguint el model de descripció proposat inicialment per Weinreich (1962), el nivell 0 s'integra en un sistema de descripció lexicogràfica que dóna lloc a la possibilitat de constituir un diccionari màximament *coherent* des del punt de vista semàntic, i que manté intacta la seva principal característica distintiva: la descripció basada en la llengua natural.

Des d'un punt de vista lògic, es pot concebre aquest nivell 0 com un conjunt de proposicions (els articles que en formen part) que enuncien una sèrie de *teoremes* que permeten la formulació del conjunt de teoremes que conforma la descripció

lexicogràfica del català. Aplicat als diccionaris, aquest sistema axiomàtic té la particularitat de concentrar tots els fenòmens de circularitat que es produeixen com a efecte associat a la praxi lexicogràfica, la qual cosa permet la redacció d'un diccionari amb un grau de coherència i d'evitació de desviacions molt més alt que el dels productes lexicogràfics convencionals.

Convé aclarir que les unitats de descripció que conformen la metallengua de definició no es poden identificar *directament* amb les entrades del diccionari, sinó amb cadascun del *sentits* (o *accepcions*) que cada entrada té definits explícitament. Cada unitat de metallengua és, doncs, un *signe* (en termes saussureans: una unió indissociable d'un *significant* amb un *significat*).

Així doncs, tot i que han estat formulats com a articles de diccionari, els elements de la metallengua no constitueixen una *part estructural identificable* del diccionari, sinó el conjunt d'elements que configura la seva descripció semàntica. Per raons pràctiques, he donat al nivell 0 una forma similar a la d'una sèrie d'articles lexicogràfics, però les informacions que contenen aquests articles no pretenen ser una descripció *exhaustiva* de cada entrada, sinó el conjunt nuclear de la metallengua de descripció del diccionari, és a dir, del conjunt més ampli sobre el qual es basa la descripció lexicogràfica del significat lèxic.

Dit d'una altra manera, el nivell 0 descrit no es correspon a cap conjunt d'articles en un diccionari ja redactat, sinó que constitueix la base del seu sistema de descripció semàntica. En un diccionari ja redactat, les entrades que apareixen al nivell 0 hauran de contenir, a més d'aquestes informacions, tots els sentits no recollits en aquest nivell bàsic i, per tant, recolliran definicions que hauran de ser formulades a partir d'unitats de nivells superiors.

D'altra banda, sí que es pot dir que les unitats del nivell 0 (i de la resta de nivells de descripció) formen una part no estructural del diccionari, en el sentit que les accepcions corresponents a cada nivell de descripció han de ser necessàriament, per raó de

l'aplicació coherent del mètode de descripció lexicogràfic, presents en els articles corresponents del diccionari ja redactat.

Cada entrada d'aquest nivell 0 té com a correlats possibles en el vocabulari de definició cadascun dels sentits en què es divideix l'article. Convencionalment, es poden identificar de forma no ambigua¹⁵⁴ cadascun d'aquests sentits definidors amb una etiqueta que correspon al mot d'entrada i un superíndex que indica el sentit concret de què es tracta. Per tant, a una entrada qualsevol li corresponen, com a correlats, tants elements potencials del vocabulari definidor com accepcions té el seu article.

Ja he mostrat a § 7.6 que cadascuna de les unitats del vocabulari definidor pertany a un *nivell de descripció* que s'estableix en funció del nivell més alt de les unitats de definició que formen la seva paràfrasi de definició, de manera si *i* és el nivell més alt de tots els elements que componen una definició, l'element de metallengua descrit per aquesta definició pertanyerà al nivell *i+1*.

L'únic nivell en què no es verifica aquest procés és en el nivell 0, ja que les definicions que pertanyen a nivell 0 corresponen exclusivament a sentits que pertanyen també a aquest mateix nivell axiomàtic. Alhora, l'existència d'aquest nivell axiomàtic permet també d'evitar l'existència de relacions circulars en les descripcions pertanyents a nivells superiors a 0.

D'acord amb aquests principis de descripció, es pot inferir que el nombre de nivells que s'obtenen en una descripció lexicogràfica normalitzada a partir d'aquests principis reflecteix la complexitat semàntica dels lexemes descrits, i representa de manera adequada les relacions semàntiques que mantenen els lexemes entre si. L'indicador d'aquesta aproximació componencialista a la descripció del significat lèxic no té relació directa ni necessària amb la quantitat de vegades que apareix un lexema en les descripcions de la resta. Si bé resulta evident que en el seu conjunt hi ha d'haver

¹⁵⁴ M'ocuparé amb més detall d'aquesta qüestió a § 8.5.

necessàriament una sèrie d'elements lèxics pertanyents al nivell més alt que no apareixeran en la descripció de cap altre element, també és cert que no necessàriament els elements de nivells inferiors (1, 2, 3, etc.) han d'aparèixer més representats en les definicions que els de nivells superiors. Allò que determina, de fet, la major o menor freqüència d'aparició d'un element en el total de descripcions del diccionari no és, doncs, el nivell a què pertany, sinó el seu grau de necessitat en la descripció semàntica de la resta de sentits del diccionari. Complementàriament, si el nivell a què pertany un sentit no té relació directa amb la seva freqüència d'aparició en la resta de les definicions sinó amb la seva complexitat semàntica, es pot comprendre fàcilment que el caràcter d'aquesta metallengua no és, doncs, el d'una jerarquia semàntica que ordena el grau de generalitat dels seus elements, sinó una organització que situa els elements en una estructura que té relació amb la seva complexitat intrínseca. Idealment, una jerarquia és capaç d'organitzar únicament (Soler, 1992a) els anomenats *descriptors* de les definicions en una estructura especialment dirigida a l'evitació de les circularitats de descriptors del diccionari.

El model d'organització jeràrquica, però, és insuficient per reflectir la complexitat de tota la descripció lèxica dels elements lingüístics. En primer lloc, pel fet que en molts casos les definicions dissocien les propietats semàntiques i sintàctiques del descriptor en lexemes diferents (veg. Ilson, 1987, i els comentaris que he fet més enrere § 4.2). I en segon lloc, perquè (com he mostrat a § 5.1) la definició lexicogràfica presenta, a més del descriptor pròpiament dit, diversos nuclis de descripció que resulten de gran importància en la caracterització del significat.

L'estructura de diccionari que s'obté com a resultat de l'aplicació del sistema de definició proposat correspon a una xarxa semàntica de relacions entre les descripcions dels lexemes.

8.2 Elements complementaris de la metallengua

Les definicions que descriuen les unitats del nivell 0 contenen alguns elements que no formen part de cap dels nivells de descripció proposats en aquest treball. Aquests elements pertanyen a dos tipus diferents:

a) Variables, que identifiquen els components extrínsecs de les definicions, i que es representen mitjancament les lletres majúscules *X*, *Y* i *Z*.

b) Mots pertanyents a certes classes especials.

Un i altre tipus tenen, en el sistema proposat, la característica comuna de pertànyer a conjunts tancats de la metallengua (no ampliables, pel fet que ja cobreixen totes les necessitats descriptives de la metallengua de definició), susceptibles d'aparèixer en definicions corresponents a qualsevol nivell de descripció. En el cas dels mots, es tracta d'elements funcionals de la metallengua que, malgrat que tenen entrada pròpia al diccionari, aquesta no es vincula amb la metallengua de descripció.

Aquesta diferència de tractament respecte a la resta dels elements que conformen la metallengua es fonamenta en el fet que l'impacte d'aquests elements en les desviacions estructurals del diccionari és mínima, pel fet que no concentren cap dels nuclis descriptius de la definició.

Es poden distingir tres casos diferenciats en què es produeix aquesta situació:

1. Elements que no són susceptibles de tenir definició lingüística pròpiament dita.
2. Verbs funcionals que en general serveixen per a l'activació semàntica dels substantius amb què apareixen. Aquests verbs són *ser*, *estar*, *fer* i *tenir*.

3. Elements que formen part de determinades sèries lèxiques sistemàtiques: numerals (noms i adjectius, excepte *primer*), dies de la setmana i mesos de l'any.

El perfil general dels mots del tipus 1, en termes lexicogràfics, correspon al de les entrades hipolexemàtiques que he analitzat a § 2.6. Des del punt de vista de la metallengua de definició, el seu caràcter no és gaire diferent del que desenvolupen en la llengua general: es tracta de mots funcionals necessaris per a l'expressió lingüística, però no dotats del mateix tipus de significat conceptual que els mots lexemàtics pròpiament dits, sinó més aviat d'unes propietats semàntiques que tenen la seva referència en la gramàtica (i més precisament en la sintaxi) de la llengua.

Entre els elements del tipus 1 inclosos en la metallengua de descripció hi ha:

preposicions:¹⁵⁵ *a, amb, de...*

determinants: *el, un, aquell, aquest, cadascun, seu...*

adverbis:¹⁵⁶ *abans, darrere,...*

quantificadors: *més, gairebé...*

conjuncions: *que, i, o,...*

pronoms: *allò, cadascú, algú, ell, el, li, ho, en, hi,...*

Per altra banda, els elements del tipus 2, malgrat que pertanyen a la classe gramatical oberta dels verbs, tenen un significat conceptual difícilment objectivable i que depèn en gran mesura del seu context sintagmàtic.

Finalment, els elements del tipus 3 corresponen a sèries lèxiques que en molts casos són necessàries per a la determinació del significat conceptual de certs mots. La incorporació parcial d'aquestes sèries, a través de la incorporació selectiva d'alguns dels seus elements d'acord amb les necessitats del nivell 0, entraria en conflicte amb la seva

¹⁵⁵ També les contraccions: *al, del i pel*.

¹⁵⁶ Alguns d'aquest grup presenten comportaments preposicionals (*darrere la cortina*) o adjectivals (*el dia abans*).

pròpia naturalesa dins del sistema de la llengua. Cal considerar, doncs, aquestes sèries, com a conjunts d'elements disponibles per a l'expressió del significat en qualsevol nivell de descripció.

La llista completa dels elements complementaris (presentats com a formes gràfiques diferents) que apareixen en el nivell 0 és la següent:

a	del	gairebé	-ne
abans	dels	gens	ni
així	des	ha	no
al	després	haver	només
algú	deu	havia	o
algun	dins	hi	on
alguna	dos	-hi	pel
algunes	dues	hom	pels
alguns	el	i	per
als	els	ja	perquè
altra	ell	'l	qual
altre	ella	l'	qualsevol
altres	elles	la	qualssevol
allò	ells	-la	quan
amb	en	les	quant
amunt	endavant	-les	quatre
aquell	endins	li	que
aquest	enfora	-li	què
aquesta	enllà	-lo	quelcom
aquests	entre	-los	qui
avall	era	lluny	quina
bé	es	llur	res
cada	és	llurs	's
cadascú	està	mateix	s'
cadascun	estan	mateixa	se
cadascuna	estar	mateixes	-se
centena	estigui	mentre	segons
com	fa	menys	sempre
contra	faci	més	sense
d'	fan	mitjançant	ser
damunt	fent	molt	serà
darrere	fer	moltes	seva
davant	feta	'n	seves
de	fora	n'	si

sigui	tal	tercera	unes
sinó	tant	tingui	uns
sobre	té	tingut	vagi
sobretot	tenen	tres	vint-i-quatrena
són	tenia	treure	X
sota	tenir	un	Y
sovint	tercer	una	Z

Un altre dels components de les definicions del nivell 0 que no formen part de les seves unitats de descripció són les variables *X*, *Y* i *Z* que uso per a representar l'extrinsicitat de certs components de la definició. Aquesta no és l'única manera possible de representar aquest tipus d'elements, però l'he escollida perquè és d'interpretació fàcil i permet la utilització d'aquests components alhora com a actants sintàctics i semàntics.

Finalment, i fora de la casuística precedent, cal dir que des del punt de vista d'un diccionari complet es fa evident la necessitat de resoldre la necessària incorporació d'elements que no pertanyen a cap nivell de descripció, sota unes condicions clares i específiques. La necessitat d'aquesta incorporació es fonamenta en certs processos de derivació lexical a partir de noms propis (*Mallorca, mallorquí; Verdaguer, verdaguerià;* etc.). Les necessitats imposades per aquests tipus lèxics en particular, i per altres mots que excepcionalment requeriran la incorporació a la metallengua de definició de determinats noms propis.

Amb freqüència, els diccionaris convencionals incorporen al seus vocabularis definidors elements que corresponen de fet a una altra metallengua en general de tipus científic. Un dels casos més notables d'això en els diccionaris catalans a partir del DGLC és la utilització dels noms científics (zoològics i botànics) per a la caracterització de certs animals i plantes. Encara que aquesta opció metodològica ha estat en certs casos valorada per la seva precisió científica (veg. Riera & Vallès, 1991), crec que, en un sistema de definició com el que proposo, la caracterització dels lexemes ha de seguir un patró descriptiu homogeni i, en aquest sentit, evitar la utilització de sistemes de descripció externs a la lexicografia i fora de l'abast de la distinció semàntica pròpiament

dita. No cal insistir gaire en aquesta qüestió per adonar-se de les múltiples i interessants implicacions de tot tipus (metodològiques, teòriques, descriptives, lingüístiques, etc.) que hi ha darrere d'aquesta qüestió, i que una anàlisi detallada exigiria tractar adequadament. A fi de no desviar-me dels objectius fixats, no aprofundiré més en aquesta qüestió.

8.3 Condicions de descripció i nivells de descripció

Les condicions sobre la descripció lexicogràfica que he formulat a § 6.3 es refereixen a diversos aspectes tècnics, lingüístics, descriptius, etc. que cal tenir en compte en l'elaboració de tota obra lexicogràfica, i més concretament en el seu sistema de formulació de definicions. Aquestes condicions, formulades amb independència de les característiques específiques de cada obra, es compleixen d'una manera particular en cada diccionari i, especialment, en el sistema de definició proposat en aquest treball.

La demostració de la correspondència entre les condicions i el sistema suposa, en primer lloc, l'ajustament de les unes a les especificitats de l'altre i, en segon lloc, la validació de la proposta d'acord amb els principis generals de qualsevol sistema de formulació de definicions. Més concretament, l'adaptació de cadascuna d'aquestes condicions al sistema de definició proposat representa l'anàlisi detallada de la seva integració en el procés de redacció d'un diccionari, i n'assegura la cohesió i la motivació des del punt de vista de l'anàlisi lingüística.

Aquestes condicions, de fet, ja s'han complert íntegrament en el procés de constitució del nivell 0 (§ 7), que per raons pràctiques té una estructura i una presentació de les dades similars a la d'un diccionari. En aquesta secció m'interessa, sobretot, la valoració de la influència que aquestes condicions poden tenir en la redacció d'un diccionari semànticament coherent, especialment en el curs de la seva elaboració.

Un breu repàs a les condicions expressades per a la descripció lexicogràfica pot servir per donar una idea més clara de la implicació en el sistema de definició dels criteris de d'organització de la metallengua. Així, s'arriba a la formulació completa d'aquest sistema de definició i a la integració dels seus aspectes descriptius.

1. Condició de regularitat. Representa un important factor de minimització dels problemes estructurals del diccionari. En el cas concret de les regularitats relacionades amb la morfologia derivativa (noms d'acció i efecte derivats de verbs, noms de qualitat derivats d'adjectius, etc.) un tractament sistemàtic basat en la utilització (explícita o implícita)¹⁵⁷ de definicions formulàries contribueix a sistematitzar l'estructura semàntica del diccionari. En termes d'una descripció per nivells, la descripció del derivat correspondrà sempre a un nivell més alt que el de la seva base lèxica, amb l'única excepció dels derivats presents al nivell 0; així, la descripció de *teatre* ha de ser prèvia a la descripció de *teatral* ('relatiu al *teatre*'), la qual cosa representarà l'establiment de paral·lelismes entre l'anàlisi lingüística (semàntica, morfològica) i la lexicogràfica. El mateix raonament pot ser aplicable a altres tipus de regularitats, que afecten els esquemes de subcategorització dels lexemes (per exemple, les alternances d'ergativitat que es produeixen en els verbs, no tractades específicament en aquest treball però que cal integrar en els mecanismes de redacció d'un diccionari) o bé les alternances de tipus semàntic (com ara les alternances metonímiques del tipus continent-contingut, etc. que tampoc he tractat de manera específica). Un tractament màximament sistemàtic d'aquests fenòmens contribueix decisivament a l'organització coherent dels diccionaris, i potencia les capacitats del sistema de descripció proposat.

¹⁵⁷ Entenc aquí per explícita la regularització consistent en la utilització de definicions formulàries en les accepcions que compleixen la regularitat semàntica que es produeix en cadascun dels processos derivatius. Aquesta tècnica és la utilitzada més correntment, bé que amb freqüents manques de sistematicitat, en els diccionaris convencionals. Per regularització implícita entenc qualsevol mecanisme de formulació basat en la referència al procés derivatiu corresponent, que normalment implica el tractament dels derivats en l'article corresponent a la seva base lèxica, i l'associació general de cadascun d'aquests processos a una definició formulària que es dona en les claus d'interpretació del diccionari. Aquest és el sistema que és aplicat, a grans trets, en el tractament dels casos de derivació més regular en diccionaris com ara el LDOCE.

2. Condició de distinció de constituents. Com ja he mostrat a § 4.2, en tractar dels elements *intrínsecs* i *extrínsecs* de la definició, aquesta condició representa un factor important per a l'evitació de les denominades *definicions impròpies*. Els avantatges descriptius de la distinció de constituents es posen especialment de relleu si s'utilitza de forma similar a com ho he fet al nivell 0, ja que els elements extrínsecs són identificats per variables que, en una descripció lexicogràfica integrada, es restringirien fora de la definició estricta. Si la definició de *perpetrar* consta, per exemple, de la següent paràfrasi 'fer X' i, en un apartat específic dedicat a l'exposició de les restriccions semosintàctiques (o bé dins de la mateixa definició, com un element extrínsec), determinem que $X = \text{una acció punible}$.

3. Condició d'adequació. El sistema de definició proposat s'integra en una aplicació rigorosa d'aquesta condició. Les definicions representen la caracterització semàntica del definit a partir dels elements necessaris per a la seva descripció i suficients per a la seva diferenciació de qualsevol altre significat. Aplicada en aquest sentit, la condició d'adequació implica l'eliminació *de facto* de la informació impròpia i, especialment, dels casos de redundància. Ja he mostrat a § 5.2 l'impacte que l'evitació estricta de la redundància comporta en la reducció de les relacions circulars en l'estructura semàntica del diccionari. En el sistema proposat, l'eliminació de la redundància és una de les condicions d'aplicació que s'integra d'una manera més natural, ja que es relaciona directament amb l'evitació de la recurrència d'informació continguda en nivells anteriors i permet la representació externa de la complexitat semàntica del definit i la correspondència d'aquesta complexitat amb la posició que ocupa cada element de la metallengua.

4. Condició d'homogeneïtat discursiva. La determinació d'uns certs recursos expressius vàlids en la redacció de definicions resulta necessària per a la homogeneïtat de les definicions que formen part de la metallengua. Aquesta condició es manifesta en l'explicitació d'aspectes diversos que van des dels més superficials (l'ús de *ser* o *ésser*, o l'ús de les diverses solucions de pronoms relatius amb preposició (*amb què* / *amb el qual*, per exemple), fins a aspectes més directament relacionats amb el mètode de

descripció (com ara la forma dels elements extrínsecs, o l'expressió dels actants que apareixen en certes definicions). En el sistema de definició proposat, la determinació d'aquests aspectes pot ser variable, d'acord amb el tipus de diccionari a què corresponguin. El grau de detall d'aquesta determinació i les tries concretes que es realitzin sobre les diverses qüestions lexicogramaticals i metodològiques relacionades amb l'estil de la definició s'han de fer en funció de les característiques generals del diccionari i del tipus de públic a qui s'adreça.

5. Condició de síntesi descriptiva. L'evitació en el diccionari de les definicions analítiques equivalents (definicions repetides) és també un dels factors importants de reducció de la redundància en l'estructura global de l'obra. La integració de mecanismes de control en el tractament de les sèries sinonímiques en el sistema de definició proposat constitueix una opció que no només millora la coherència estructural del diccionari, sinó que, en el sistema de definició proposat, evita la proliferació indiscriminada de nivells de descripció. En el cas concret d'una sèrie sinonímica donada, si el nivell de descripció a què correspon el cap de sèrie és i , tots els membres subordinats de la sèrie pertanyeran al nivell $i + 1$. En el cas de les definicions analítiques, l'aplicació rigorosa d'aquest principi es correspon amb el *maximum block principle* de Mel'čuk (1988a) (veg. més enrere § 6.1), i té com a conseqüència directa en el sistema de definició proposat la descripció de cada accepció d'acord amb el seu grau de complexitat semàntica. Un exemple pràctic: si *teatre* ^{i} = “edifici on es representen espectacles”, qualsevol aparició en la metallengua de definició del significat conceptual corresponent a *edifici on es representen espectacles* implica la utilització de *teatre* ^{i} i no pas de cap altra fórmula; per consegüent, totes les definicions on aparegui aquest significat correspondran a elements de la metallengua que pertanyeran a un nivell superior o igual a $i + 1$.

6. Condició d'univocitat. L'aplicació rigorosa d'aquesta condició és una exigència del sistema de definició proposat, i un dels elements caracteritzadors més importants de la metallengua d'un diccionari elaborat a partir dels principis que he exposat en aquest treball. Per aquest motiu, tractaré amb detall sobre l'aplicació d'aquesta condició més

avall (§ 8.5) i de les possibilitats d'integració rigorosa d'aquesta condició en un diccionari general. De moment, em limito a assenyalar la presència d'ambigüitat en els elements de la metallengua com un dels factors més importants d'incertituds i de possibles circularitats en el diccionari. La necessària correspondència entre la metallengua de descripció i els lexemes descrits exigeix una solució satisfactòria per a aquesta qüestió; sense aquesta solució, la coherència interna del diccionari se'n veurà ressentida o simplement quedarà amagada per les ambigüitats superficials. Amb la finalitat de posar les bases d'aquest tractament unívoc dels elements de la metallengua de descripció, he desambiguat (amb una referència numèrica superindexada per referència a la seva descripció), tots els elements del nivell 0 que tenien més d'una descripció pertanyent al mateix nivell. La projecció dels nivells subsegüents que determina el sistema de definició ha de tenir el mateix grau d'univocitat.

7. Condició de projecció coherent de relacions. Els principis a partir dels quals es produeix la projecció de nivells assegurin la projecció coherent de les relacions de tipus semàntic que es produeixen entre els lexemes de la llengua. Les possibilitats de circularitat queden, com he dit, reduïdes al nivell 0, que és el que concentra totes les possibles desviacions descriptives del diccionari. L'aplicació del sistema de descripció evita qualsevol tipus de circularitat fora del conjunt limitat que he assenyalat, per la qual cosa aquesta condició es compleix amb el màxim grau de rigor que permet el mètode de descripció.

8. Condició de transparència. La transparència de la metallengua de descripció queda assegurada, d'una banda, per la manera com s'aplicarà la condició d'univocitat i, de l'altra, per l'evitació de la incorporació a la metallengua de fórmules que tinguin un sentit fraseològic diferent del seu sentit merament combinatori. La metallengua del sistema de definició proposat es compon exclusivament de lexemes que no formen part de cap combinació lèxica locutiva o fraseològica. S'evita, doncs, en el procés de redacció del diccionari, la utilització de fórmules de significat no componencial, ja que (tret del cas que es representin amb alguna marca distintiva) produeixen enunciats de definició no transparents.

9. Condició de limitació del vocabulari definidor. La limitació d'unitats va implícita en la manera com s'ha constituït el nivell 0. Cal insistir, però, en el fet que el sistema proposat no representa l'aplicació d'un conjunt predeterminat i tancat d'elements que conformen la metallengua de descripció. La forma de la metallengua de definició del diccionari i, per tant, el conjunt del seu vocabulari definidor, es determina de fet durant el procés de la redacció de l'obra. No és fins després de completar totalment la descripció dels lexemes que hom arriba a la determinació completa de la metallengua del diccionari. No es tracta, per tant, d'una limitació apriorística del vocabulari en la línia dels diccionaris fets amb vocabularis controlats, sinó de la integració dels elements d'aquest vocabulari amb un sistema de descripció que determina les lleis a partir de les quals els elements descrits en el diccionari poden ser usats com a elements de la metallengua de definició.

10. Condició d'integració. En molts casos, la definició lexicogràfica no pot, per si sola, caracteritzar els múltiples aspectes associats al significat lèxic. L'objectiu principal de la definició lexicogràfica és la caracterització del significat conceptual del definit, per la qual cosa convé el desenvolupament de mecanismes de descripció per a les qüestions relatives al comportament sintagmàtic (col·locacions, patrons d'inserció, etc.) i gramatical (comptabilitat dels noms, gradabilitat dels adjectius, etc.). Només descarregant les definicions de la necessitat de recollir aquest tipus d'informacions es pot arribar a la utilització d'un sistema de definició veritablement eficient.

8.4 Els nivells de descripció en la redacció lexicogràfica

Un cop definit el nivell 0 de la metallengua i el sistema de projecció successiva i organitzada dels nivells de descripció, cal abordar la principal qüestió pendent: la integració pràctica d'aquests components en el diccionari i en el seu procés de redacció.

Aquesta qüestió es pot plantejar a partir de tres interrogants específics que la relacionen amb els aspectes del diccionari amb què està en estreta connexió:

1. És possible limitar el nombre de nivells de descripció a partir de l'aplicació de criteris del sistema de definició?
2. Des del punt de vista pràctic, quines adaptacions estructurals cal fer sobre el model d'un diccionari convencional per a la constitució d'un diccionari descriptivament coherent?
3. Com es pot integrar el sistema de definició descrit en les tasques de redacció del diccionari, entenent que aquest procés confereix al diccionari una estructura en què el conjunt lèxic i el vocabulari definidor estan en creixement continu durant el procés de redacció?

La primera d'aquestes qüestions es fonamenta en el fet que la proliferació indiscriminada i incontrolada de nivells de descripció representaria, en cas de produir-se, una minva de les capacitats descriptives del sistema de definició proposat, ja que les possibilitats de prendre el diccionari com a font per a l'establiment de classes semàntiques serien molt menors, i segurament també se'n ressentiria la seva capacitat informativa en els usos més convencionals. Potencialment, el sistema que he descrit a §7.6 permet l'existència d'un nombre de nivells que formalment està només limitat pel nombre d'accepcions possibles del diccionari.

De fet, Weinreich (1962), Martin (1974, 1977*a*, 1977*b*) o Calzolari (1977) no fan cap referència explícita a aquesta qüestió, ni a les interaccions que es poden produir entre aquest tipus de plantejaments de descripció i el procés de redacció d'un diccionari.

En el marc d'aquest treball, aquesta qüestió ja ha rebut resposta implícita en l'adaptació al sistema de les condicions de descripció que he tractat a § 8.3. Les cinc primeres condicions (les de **regularitat**, **distinció de constituents**, **adequació**, **homogeneïtat discursiva**, i **síntesi descriptiva**) redueixen la possibilitat d'una proliferació excessiva de nivells de descripció per l'aplicació que se'n fa al sistema de descripció proposat.

Pel que fa a la segona qüestió, a § 8.5 tracto els aspectes fonamentals dels canvis que s'hi plantegen.

M'ocuparé, doncs, en aquesta secció dels aspectes vinculats al procés de redacció lexicogràfica que planteja la tercera qüestió.

La projecció coherent dels nivells de descripció que proposo a través de l'aplicació d'aquest sistema no representa un simple ajustament superficial dels mètodes de treball lexicogràfic, sinó que condiciona múltiples aspectes de la redacció, el desenvolupament i el progrés d'aquest treball, ja que el procés de redacció de les definicions és allò que, en darrer terme, va configurant l'estructura semàntica del diccionari. Aquest mètode de treball condiona, de manera destacada, el criteri amb què s'han de seleccionar els mots que han de ser incorporats, successivament, al procés de redacció.

Per raons d'ordre pràctic que tenen més a veure amb factors externs que lexicogràfics, la majoria de diccionaris elaborats fins al moment han estat redactats de la *a* a la *z*, seguint l'ordre alfabètic.¹⁵⁸ La classificació alfabètica dels articles que presenten els diccionaris en la seva forma final no condiona, en un model com el que proposo, ni de fet en cap altre model avançat de diccionari, que la selecció de les entrades a redactar hagi de seguir aquest mateix ordre. En un sistema de descripció organitzat per nivells, la progressió alfabètica dels treballs com a únic factor de progressió de la redacció és especialment inadequada.

A diferència del que succeeix en els diccionaris convencionals, en el sistema de definició proposat, la unitat de treball bàsica —pel que fa al procés de construcció de la metallengua— no és l'article, sinó l'accepció. Allò que conforma la pertinença d'un element a un nivell de descripció no és la compleció de les informacions de l'article a què pertany, sinó la seva descripció semàntica al nivell de l'accepció.

¹⁵⁸ Pascual (1996) assenyalava aquest fet com un dels principals factors d'introducció d'incoherències en els diccionaris d'ús.

Ja he fet notar a § 2.7 que el model de lèxic a partir del qual s'organitza un diccionari és bàsicament enumeratiu, i que pot ser més o menys matisat segons el tractament de determinades qüestions referents al lèxic (mecanismes de derivació, alternança d'esquemes de subcategorització, etc.). Traslladant aquesta observació al procés de redacció d'un diccionari, una conseqüència fonamental que se'n pot extreure és que les relacions estructurals en la descripció semàntica dels lexemes no es produeixen entre entrades sinó entre les accepcions, els sentits, en què cada entrada organitza la seva descripció dins de l'article. El procés de redacció d'un diccionari, que habitualment s'organitza prenent l'entrada aïllada com una unitat bàsica de progressió, pot concebre's de manera radicalment diferent si es considera que les relacions internes en l'estructura semàntica del diccionari es produeixen entre els sentits enumeratius de què es compon cada article.

Es tracta, doncs, de concebre el procés de redacció com l'enumeració dels sentits associats a cada entrada. L'objectiu principal que determina el progrés del treball és, doncs, la determinació de tots els sentits corresponents a cada entrada. A diferència d'allò que sol ser habitual en els diccionaris convencionals, en aquest sistema les relacions internes que s'estableixen en la descripció semàntica del diccionari poden exigir la redacció de sentits que no corresponen a l'entrada que és objecte de redacció.

El procés de redacció d'un article lexicogràfic té una fase prèvia que no es distingeix de les d'un diccionari convencional (veg. Kilgarriff, 1998), i que consisteix bàsicament en l'organització de l'estructura general (divisió en accepcions) de l'article a partir de l'observació de les dades del corpus. Un cop completada aquesta fase, el sistema de descripció proposat imposa certes condicions metodològiques pel que fa a la redacció de les definicions pròpiament dites.

Partint de la idea de la projecció estructurada de nivells de descripció s'estableix que, a l'hora de redactar una definició lexicogràfica d'un diccionari basat en una estructuració de nivells, pot passar que:

1. Tots els sentits que cal incorporar a la definició formen part del nivell 0. En aquest cas acaba el procés, i la descripció de tots els sentits s'incorpora al nivell 1 de la metallengua. En cas contrari, continua el procés.
2. Alguns dels sentits que cal incorporar a la definició formen part d'un nivell >0 . En aquest cas pot passar que:
3. Els elements necessaris ja han estat objecte de descripció i, per tant, ja tenen nivell de descripció assignat. En aquest cas es redacta la definició i s'incorpora el nou element a la metallengua amb el nivell que li correspongui. O bé pot passar que:
4. Els elements necessaris no han estat descrits encara en el procés de redacció. En aquest cas pot passar que:
5. Són descriptibles integralment amb elements de nivell 0. En aquest cas, es fa la descripció, se'ls assigna nivell 1, i es torna al punt 3. O bé pot passar que:
6. Són elements que han de pertànyer a un nivell >1 . En aquest cas se'n fa la descripció corresponent integrant-los en el procés de redacció (seguint els passos descrits per al procés a partir d'1) fins que estan en condicions de complir els requeriments de 3.

Aquesta concepció recursiva del procés de redacció lexicogràfica pot veure's representada en la figura 8.1, que recull de manera sintètica —i simplificada— les principals operacions del procés que he desglossat més amunt.

L'aplicació d'aquest sistema origina *de facto* la progressió estructurada de la descripció lexicogràfica del diccionari (veg. Pascual, 1996); entre les necessitats que suposa hi ha el fet de poder comptar amb un sistema adequat per al tractament rigorós de tot aquest volum de dades. El diccionari es concep, així, com una base de dades on s'incorpora tota la informació referent a la caracterització dels lexemes.

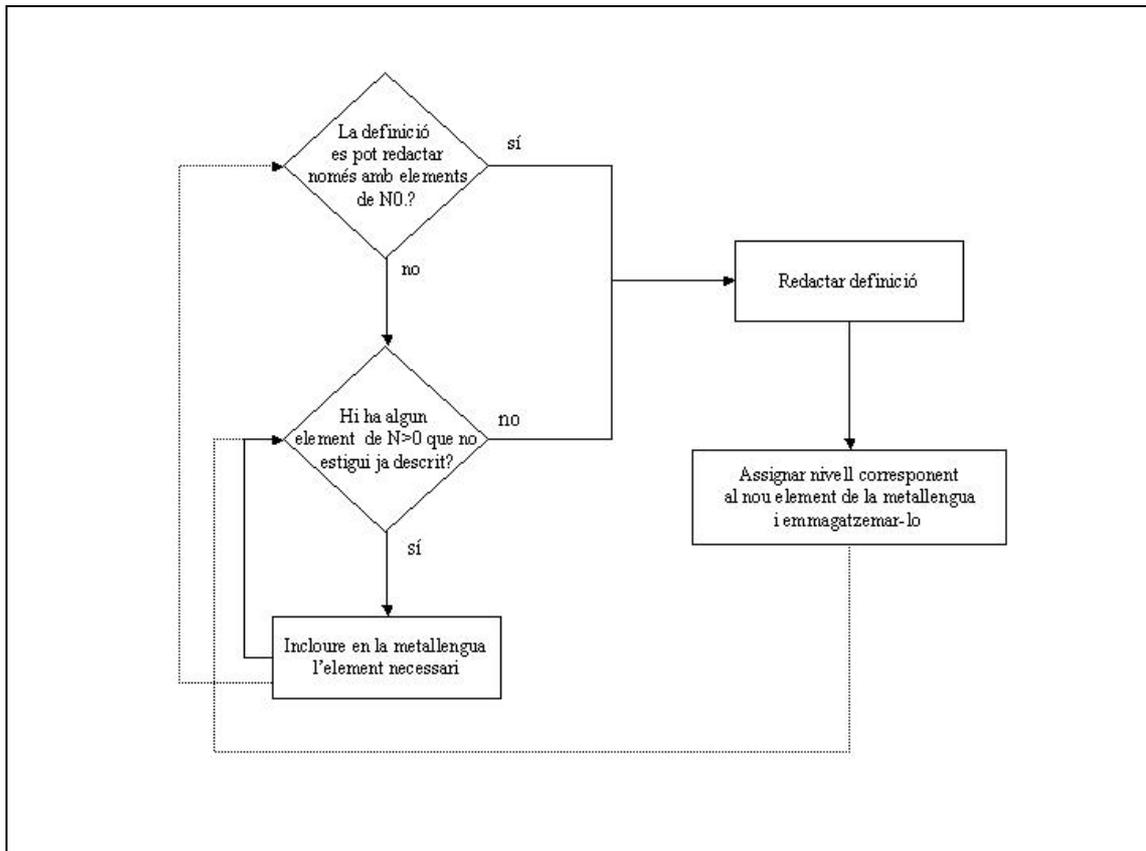


Figura 8.1. Procés de redacció lexicogràfica.

Els aspectes pràctics de l'organització del treball de redacció es poden resoldre a partir del disseny adequat d'aquesta base de dades i del desenvolupament de les eines de redacció específiques. Cal fer esment, per exemple, que aquest sistema implica que durant la redacció d'un article lexicogràfic calgui formular descripcions semàntiques que corresponen a articles diferents (pas 6); una base de dades informatitzada és el recurs òptim per a l'emmagatzemament d'aquest tipus d'informacions, que formaran part dels articles encara no redactats en la seva totalitat. Així, en redactar l'article corresponent a una entrada nova (seleccionada a partir dels criteris de progressió que s'estableixin) el lexicògraf pot trobar-se que a la base de dades ja es predeterminen aquells sentits que hauran estat utilitzats en la descripció semàntica d'una altra acepció corresponent a un altre article. La tasca de redacció d'aquests articles, doncs, es redueix

a la incorporació d'aquells sentits que no han estat utilitzats com a elements de descripció en altres articles.

Cada sentit de cada mot que cal redactar esdevé, així, un node que contreu relacions amb la resta de sentits que integren la seva descripció lexicogràfica. El grau de complexitat major o menor d'aquestes relacions representa la complexitat del significat de cadascun dels sentits d'una entrada.

En un procés de redacció com el que proposo, la millor alternativa és la redacció conjunta dels articles que mantenen algun tipus de vincle derivatiu o semàntic comú. Així, un lexicògraf no redactaria un sol article sinó un conjunt d'articles relacionats, la qual cosa reduiria les possibilitats que el sistema resultés poc operatiu en la seva aplicació pràctica.

En aquest sistema de redacció *en xarxa* cal posar una atenció especial en la selecció dels blocs d'articles que cal redactar conjuntament; atès que el sistema de redacció està orientat a la constitució estructurada de la metallengua de descripció, cal tenir present l'avantatge que pot representar la redacció simultània de les entrades que formen part d'una sèrie sinonímica. La determinació del cap de sèrie i el tractament estructurat de tot el conjunt són operacions que es poden veure notablement afavorides amb aquesta organització del treball lexicogràfic, ja que concentra la creació dels elements de descripció necessaris en la metallengua.

Alhora, convé tenir present la importància de les sèries derivatives, que impliquen en general un alt grau de recurrència interna en les descripcions de cadascun dels seus elements. El fet de comptar, en l'actualitat, amb els recursos lingüístics adients per a la preparació d'aquests treballs (fonamentalment diccionaris informatitzats, que permeten veure els lligams que contremen les seves definicions) facilita de forma decisiva la tasca de selecció d'unitats per a la constitució dels blocs de redacció esmentats.

El principi fonamental que segueix aquest sistema per a l'evitació de la circularitat des del mateix moment de la redacció del diccionari es basa en el fet que l'inventari

definidor disponible és funció de la complexitat semàntica de la definició de cada element de la metallengua, i obliga a descompondre i a explicitar no només el sentit del mot definit, sinó també tots aquells sentits que intervenen en la seva definició. A diferència del *principi de descomposició* que planteja Mel'čuk (1988a) (veg. més enrere § 6.2), aquest sistema explicita clarament el procés compositiu de les unitats des del mateix moment de la redacció de la definició. Alhora, el sistema situa les unitats de descripció en l'estructura general del diccionari. L'inventari definidor s'organitza, així, en nivells successius que conformen un conjunt estructurat d'unitats de metallengua disponibles per a la redacció de definicions. Aquesta organització bloqueja la possibilitat d'establiment de relacions circulars i permet explicitar al màxim la descripció lexicogràfica.

Finalment, cal dir que algunes de les implicacions que comporta un diccionari redactat a partir d'aquests principis no es poden explicitar en tot el seu detall sense el desenvolupament efectiu d'un tal diccionari. Tot i així, es poden avançar algunes d'aquestes implicacions prenent com a base l'impacte que un sistema d'aquest tipus podria tenir en la sistematització d'un diccionari existent, ja que aquesta tasca, per la seva naturalesa, presenta nombrosos punts de contacte amb el conjunt d'operacions que he aplicat per a l'obtenció del nivell 0 de la metallengua de definició.

8.5 Desambiguació de la metallengua i forma del diccionari

Els diccionaris convencionals, fins i tot els basats en tècniques de control de vocabulari, tenen una estructura definida que no pot ser considerada totalment coherent des del punt de vista de la descripció semàntica. Ja he mostrat com certes desviacions es presenten de manera recurrent en les obres lexicogràfiques, i la manera com a partir d'un sistema estructurat aquestes desviacions poden ser reduïdes al nivell descriptiu 0. D'aquestes desviacions, les que es presenten de manera més evident són les que he analitzat a § 5.1 i a § 5.2, és a dir, la circularitat i la redundància, però també les derivades de les

diferències semàntiques amb què s'usen, segons la definició de què es tracta, els elements de la metallengua de definició.

Un diccionari fet amb un sistema com el que es proposa reuneix una sèrie de característiques de manera necessària i unes altres que poden ser formulades com a proposta addicional. De fet, podria acostar-se molt a un diccionari convencional, però alhora és susceptible de presentar aspectes que el particularitzen fins a un rigor metodològic que fins ara no ha estat assolit.

El sistema proposat es concep com altament estructurat, per tant, convé que sigui aplicat amb el màxim rigor per no incórrer en possibles incoherències estructurals. Una estructura organitzada i coherent no pot permetre l'existència d'elements ambigus en les definicions lexicogràfiques. La projecció per nivells obliga a saber no només quins mots cal utilitzar en una definició, sinó a quin nivell pertanyen com a elements de la metallengua de descripció. En cas contrari, l'extensió i la quantitat de descripcions particulars pròpies d'un diccionari poden resultar un element de distorsió definitiu per a tota la seva estructura semàntica.

La construcció d'un diccionari d'aquest tipus implica, doncs, l'etiquetatge, per referència a la unitat que els correspon en la metallengua de descripció, de tots els elements de la definició. Aquesta desambiguació per referència al mateix diccionari representa la integració de la condició d'univocitat en el sistema de descripció proposat i assegura la correcta aplicació del sistema i la coherència final del resultat. Amb els mitjans actuals d'assistència a la redacció lexicogràfica aquesta condició en els termes exposats no és només una possibilitat de millora dels diccionaris, sinó que comença a ser una necessitat per a l'elaboració de diccionaris més avançats que els que fins ara ha produït la lexicografia convencional. Al nivell 0 ja he aplicat aquesta desambiguació d'elements, de forma que tots els que hi surten són utilitzats només en els sentits descrits al mateix nivell, i aquells que tenen més d'un sentit incorporen un superíndex que determina el sentit amb què s'utilitza en cada cas.

Cada unitat de la metallengua, doncs, ha d'estar etiquetada amb dues informacions:

a) El sentit concret amb què s'utilitza i que la diferencia de la resta de les unitats de la metallengua que s'associen a la mateixa entrada.

b) El nivell de descripció a què pertany la unitat de metallengua.

A més de la desambiguació d'elements de la definició, el procés de redacció permet també la incorporació d'informacions de gran interès per a l'explotació automatitzada de la informació lexicogràfica que fan referència als constituents de la definició. La condició de distinció de constituents ja determina la separació entre els elements extrínsecs i els intrínsecs. Hi ha, però, informacions que en general van associades a la part intrínseca de la definició, com ara els *descriptors* (als quals m'he referit a § 4.1) el marcatge dels quals permetria, a partir de l'anàlisi del diccionari, l'obtenció d'estructures d'organització sintacticosemàntica del lèxic. En especial, el marcatge sistemàtic dels *descriptors semàntics* permet la projecció sobre la totalitat del diccionari de totes les relacions de tipus jeràrquic.

El tractament no ambigu de tots els elements de la metallengua de definició no suposa una major complexitat en la consulta del diccionari. Fins i tot, si es considerés així, pot pensar-se en diferents possibilitats d'edició de l'obra que ometin la representació superficial de l'etiquetatge de desambiguació, especialment en edicions convencionals (no electròniques) de l'obra. Però és essencial que el diccionari com a estructura de dades estigui marcat en aquest sentit, tant per als usos humans (la possibilitat de referenciacions internes en les edicions electròniques permeten la vinculació de les unitats relacionades) com per als computacionals (la derivació i l'extracció d'informació semàntica dels diccionaris demana un tractament màximament estructurat de les dades lèxiques).

Cal distingir els aspectes relatius al sistema proposat de la forma final del diccionari. En primer lloc, perquè aquests aspectes no prefiguren necessàriament una sola i única possibilitat de presentació de les dades en l'obra lexicogràfica acabada i posada a disposició del públic; en segon lloc, perquè les diferents possibilitats de publicació amb què es compta avui (edicions convencionals, electròniques per Internet o CD-Rom, etc.)

possibiliten la convivència de diversos formats de presentació final adequats a cada suport de publicació.

Un diccionari publicat en forma de llibre, i constituït a partir dels criteris proposats en aquest treball no té per què distingir-se, des del punt de vista de la seva presentació formal, d'un diccionari corrent. Un diccionari és de fet un text altament estructurat i les innovacions proposades, tot i que fan internament més rica aquesta estructura, no suposen necessàriament la introducció d'elements de descripció que només estan a l'abast d'usuaris especialitzats. Crec que, en la seva representació externa, el sistema de descripció proposat pot resultar perfectament integrable en un diccionari adreçat a un lector mitjà culte, amb un nivell cultural corresponent a l'educació secundària. La representació externa de certes característiques, com ara la informació referent a la desambiguació dels elements que integren les definicions no constitueix una condició necessària per a certs formats del diccionari (la publicació en forma de llibre, per exemple) que podrien introduir elements de dificultat en la lectura de les definicions.

Un diccionari en suport electrònic, però, sí que és susceptible de permetre la presència de les marques de desambiguació de la metallengua de definició, que poden vincular-se directament amb les informacions corresponents (això és, les accepcions) que formen part del mateix diccionari. D'aquesta manera, l'usuari pot recuperar de forma instantània la descripció de cadascun dels elements de la definició, veure el nivell a què pertany, etc. Aquestes utilitats ja són presents en la majoria dels diccionaris publicats actualment en format electrònic; moltes d'aquestes obres permeten, per exemple, la recuperació d'un article a partir de les remissions que s'hi refereixen, i que es troben en altres articles. Simplement es tracta d'aplicar de manera sistemàtica aquestes possibilitats d'accés a la informació.

8.6 Un cas pràctic de redacció d'un article

Finalment, em proposo d'il·lustrar els aspectes pràctics del sistema de definició proposat a través de la simulació d'un article de diccionari complet corresponent al substantiu masculí *teatre*. La polisèmia d'aquest mot, que denota significats diversos però relacionats amb la semàntica de l'activitat teatral (edifici, professió, literatura, etc.) i les recurrències internes que mantenen les diferents accepcions d'aquest mot fan que sigui especialment adequat per a aquesta tasca, que té com a objectiu principal mostrar, d'una banda, la viabilitat del sistema proposat i, d'una altra banda, les millores descriptives i metodològiques que comporta la utilització d'aquest sistema.

Com ja he dit a bastament al llarg d'aquest treball, les relacions de naturalesa circular queden circumscrites al nivell 0 de descripció. Es pot il·lustrar l'existència d'aquestes relacions de naturalesa circular entre els components del nivell 0 a partir del següent exemple, que conté els elements de nivell 0 que utilitzo com a punt de partida per a l'organització de l'article *teatre*:

representar² *v.* (...) [X ~ Y] X Executar un espectacle Y.

representació² *f.* (...) Fet de representar².

espectacle¹ *m.* (...) Representació² oferta al públic.

públic¹ *m.* (...) Conjunt de les persones que van a veure un espectacle.

La representació gràfica d'aquestes relacions circulars es pot veure a la figura 8.2, que mostra clarament la circularitat que, amb diversos graus d'importància, es dona entre aquests quatre components del nivell 0. Poden observar-se recorreguts circulars mínims motivats per la vinculació semàntica dels components (*espectacle* i *públic*), i també recorreguts més extensos en què alguns dels lexemes implicats estan relacionats morfològicament (*representar* → *espectacle* → *representació* → *representar*).

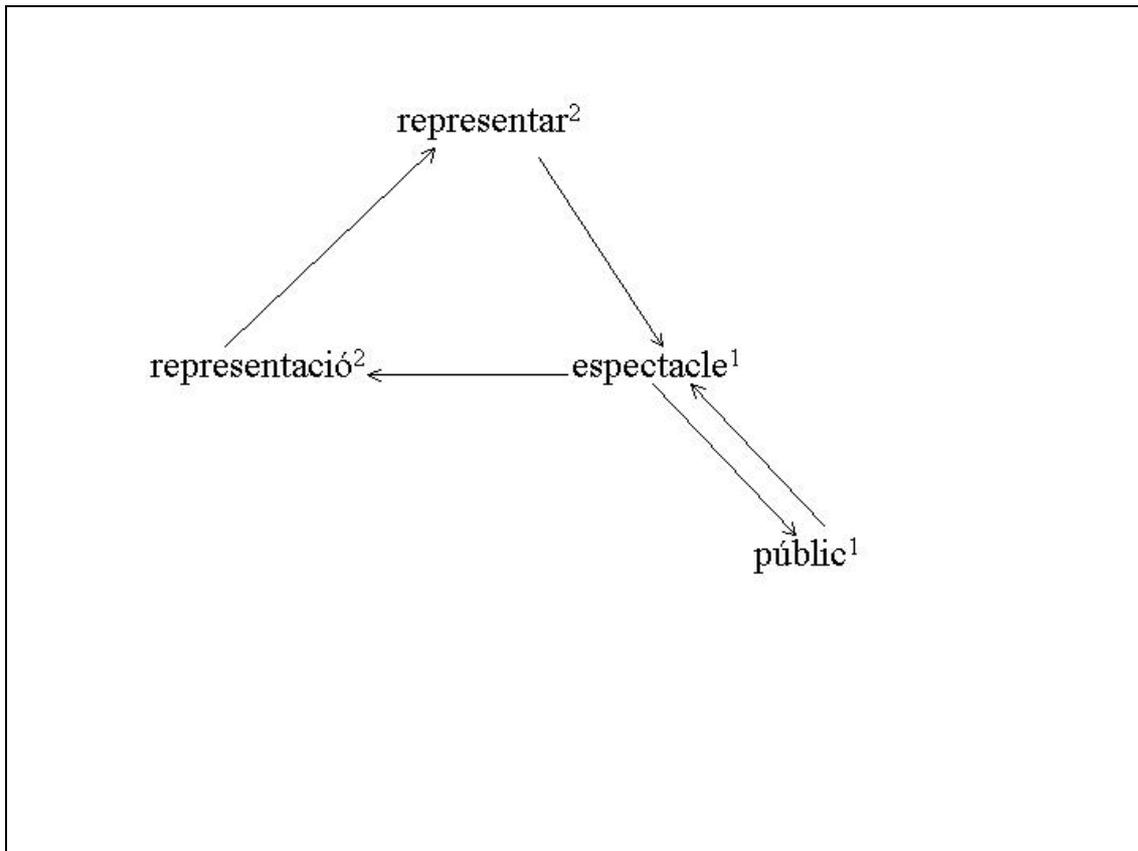


Figura 8.2. Representació de la circularitat entre els elements del nivell 0.

Cal fer notar que l'absència de redundància en aquestes definicions és un important factor de reducció de la circularitat. Si la definició de **representar²** fos 'executar un espectacle davant del públic', la de **representació²** 'Fet de representar² un espectacle', i la de **públic¹** 'Conjunt de persones que van a veure la representació² d'un espectacle', el nombre de relacions circulars que podrien establir-se entre els components del conjunt seria significativament més gran.

Per a la redacció completa de l'article *teatre*, prendré com a referència descriptiva l'article corresponent del DIEC i, d'acord amb la seva divisió de sentits i estructura general de l'article, formularé les definicions corresponents a partir del sistema de descripció proposat. L'article *teatre* en el DIEC és el següent:

teatre *m.* Edifici destinat a la representació d'obres dramàtiques. *La platea, l'escenari, d'un teatre. Les llotges d'un teatre.* | *per ext.* Edifici amb les persones que el dirigeixen. *La direcció del Teatre de Novetats.* | El públic que assisteix a una representació teatral. *Els aplaudiments de tot el teatre.* | Literatura dramàtica. *Aquest escriptor, aquest autor, coneix molt el teatre.* | Professi6 d'actor. *Gent de teatre. Dedicar-se algú al teatre. Aquell actor ja fa temps que deixà el teatre.* | Conjunt d'obres dramàtiques d'un autor, d'una època, d'un poble, etc. *El teatre grec. El teatre de Molière. El teatre del segle XVII.* | **cop de teatre** Cop d'efecte. || Lloc on s'acompleix un esdeveniment, un fet. *Polònia fou el teatre d'aquella guerra.* | **teatre d'operacions** Zona geogràfica, normalment molt àmplia, on es desenvolupa el conjunt d'una campanya bèl·lica.

Una de les primeres operacions que cal fer és trobar quina mena de relacions internes es poden detectar, des del punt de vista del contingut semàntic, entre les diferents accepcions d'aquest article. A partir d'una reformulació bàsica de les definicions, Defineixo el punt d'arribada a partir de les següents accepcions, entre parèntesis indico les referències internes que s'estableixen entre les accepcions del mateix article:

1. Edifici destinat la representació d'obres dramàtiques. (És a dir: d'obres de teatre en la tercera accepció).
- 1b Edifici amb les persones que el dirigeixen. (És a dir: que dirigeixen un teatre en la primera accepció).
2. Públic que assisteix a una representació d'un espectacle teatral. (És a dir: públic que assisteix a un teatre en la primera accepció)
3. Literatura dramàtica. (És a dir: literatura destinada a ser representada en un teatre en la primera accepció).
4. Conjunt d'obres dramàtiques d'X. (És a dir: conjunt d'obres pertanyents al teatre en la tercera accepció).
5. Professi6 d'actor. (És a dir: professi6 de qui intervé en la representació d'obres de teatre en la tercera accepció).
6. Lloc on s'acompleix un esdeveniment.

En segon terme, cal establir l'estratègia de descripci6. L'accepci6 6 és l'única que no presenta cap tipus de lligam intern amb altres accepcions i, a més, reuneix la condici6 de ser formulable íntegrament a partir d'elements inclosos en el nivell 0 de la metallengua de descripci6. De les accepcions restants (1-5) es pot assenyalar que entre

les accepcions 1 i 3 es produeix una recurrència circular¹⁵⁹ que cal desfer, mentre que la resta de relacions (la que es produeix en les accepcions 2, 4 i 5) són subsidiàries d'una de les dues anteriors (la 1b i la 2 integren la 1 en la seva descripció) i la 4 i la 5 integren la 3. L'anàlisi d'aquestes relacions pot resultar de gran utilitat per a la decisió sobre quina és la millor estratègia per enfocar la redacció completa de l'article aplicant els mecanismes de control de coherència proposats.

Així, doncs, convé considerar la recurrència que es produeix entre 1 i 3 com una de les primeres qüestions a resoldre, ja que d'aquesta recurrència depenen la resta de relacions internes de l'article. En aquest cas, cal trobar quina de les dues accepcions pot resultar més caracteritzable amb elements d'un nivell més baix de metallengua de definició. En el cas de **teatre** sembla que aquesta accepció pot ser la 1, que té el descriptor *edifici*. El primer que cal fer és determinar els passos per arribar a la caracterització del descriptor (semàntic i sintàctic). S'hi pot arribar partint de l'element **0construir**¹:

0construir¹ v. (...) [X ~ Y] X 0Formar¹ Y 0ajuntant¹ les 0parts¹ 0constituents¹ d'Y segons una 0ordenació¹.

Aquesta descripció permet d'arribar a la formulació¹⁶⁰ de l'element **1construcció**¹:

¹⁵⁹ En parlar de la circularitat (§ 5.1) no m'he referit explícitament al fet que aquest fenomen pot produir-se entre accepcions del mateix article pel fet que no es tracta de cap tipus específic de circularitat. La circularitat entre accepcions que pertanyen a un mateix article es produeix idènticament als casos que he descrit i tractat, és a dir, entre parafrasis diferents formulades en la metallengua de definició del diccionari. El fet que els diccionaris no tinguin com a norma la utilització d'una metallengua de descripció totalment desambiguada fa que la utilització com a exemples de casos de circularitat que impliquin diversos articles resultin més il·lustratius des del punt de vista d'una presentació didàctica de la casuística d'aquest fenomen.

¹⁶⁰ Atès que en aquesta secció cal combinar elements de la metallengua de diversos nivells de descripció, per a una major claredat en la representació d'aquests elements he fet alguns canvis sobre el format de representació del nivell 0 (veg. Apèndix II). Un subíndex numèric a l'esquerra de cada element de la metallengua indica el nivell a què pertany l'element; mentre que el superíndex de la dreta (usat amb la mateixa funció que en el nivell 0) indica el sentit concret del mot a què s'al·ludeix. A diferència de les dades contingudes al nivell 0, no he deixat sense marcar els valors per defecte en cas de sentits únics, que aquí es consignen tots amb el superíndex 1. Ometo, en canvi, les informacions relatives al paradigma de

${}_1\text{construcció}^1 f. (...) {}_0\text{Cosa}^1 {}_0\text{construïda}^1$.

Que al seu torn permet la formulació de l'element de nivell 2 *edifici*¹:

${}_2\text{edifici}^1 m. (...) {}_1\text{Construcció}^1 {}_0\text{aixecada}^1 \text{ sobre el } {}_0\text{sòl}^1, {}_0\text{destinada}^1 \text{ a un } {}_0\text{ús}^1 {}_0\text{determinat}^1$.

S'arriba així a poder formular ${}_3\text{teatre}^1$ com un element de metallengua de definició pertanyent al nivell 3, recorrent a la utilització de ${}_2\text{edifici}^1$ com a descriptor i a la completió de la definició a través d'elements de nivell 0:

${}_3\text{teatre}^1 m. (...) {}_2\text{Edifici}^1 {}_0\text{destinat}^1 \text{ a fer-hi } {}_0\text{representacions}^2$.

Afegint al component ${}_3\text{teatre}^1$ alguns elements que formen part del nivell 0 de descripció s'obté la descripció de ${}_4\text{teatre}^2$:

${}_4\text{teatre}^2 m. (...) {}_3\text{Teatre}^1 \text{ amb les } {}_0\text{persones}^1 \text{ que el } {}_0\text{dirigeixen}^2$.

El contingut d'aquesta descripció es pot veure representat en l'estructura per nivells que recull la figura 8.3. Es poden distingir clarament els vincles semàntics que s'estableixen per la caracterització del descriptor semàntic de la definició (${}_2\text{edifici}^1 \rightarrow {}_1\text{construcció}^1 \rightarrow {}_0\text{construir}^1$) i els que s'estableixen per raó de la resta de nuclis semàntics de la definició, que es poden fer correspondre a especificacions com ara la finalitat (${}_0\text{representació}^2$).

Les relacions que, al seu torn, contreen els elements de la descripció, per exemple ${}_0\text{representació}^2 \rightarrow {}_0\text{espectacle}^1$, ja mostrada a la figura 8.2 determinen encara nous

flexió, substituït sistemàticament per la indicació (...) i el número d'accepció encapçalant la definició, ja que aquest ja està indicat en el superíndex amb què es marca l'element de la metallengua descrit.

vincles semàntics entre la descripció de ${}_3\text{teatre}^1$ i altres elements de la metallengua (com ara, en aquest cas concret, l'element ${}_0\text{espectacle}^1$).

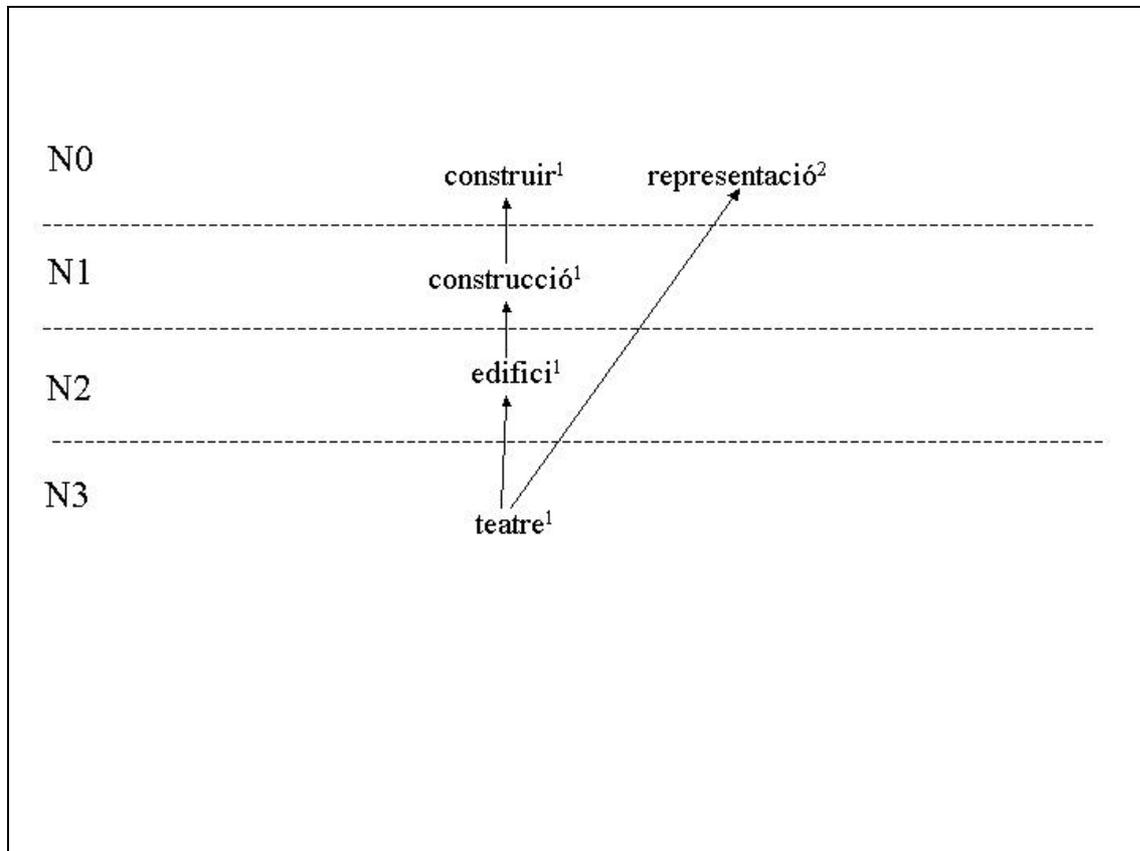


Figura 8.3. Representació per nivells de la descripció de *teatre* (1).

L'element de la metallengua ${}_3\text{teatre}^1$ pot intervenir també en la formulació d'un altre sentit del mateix article, com ara ${}_4\text{teatre}^3$:

${}_4\text{teatre}^3$ m. (...) ${}_0\text{Públic}^1$ d'un ${}_3\text{teatre}^1$.

La representació de la descripció dels elements de la metallengua ${}_4\text{teatre}^2$ i ${}_4\text{teatre}^3$, es recull a la figura 8.4.

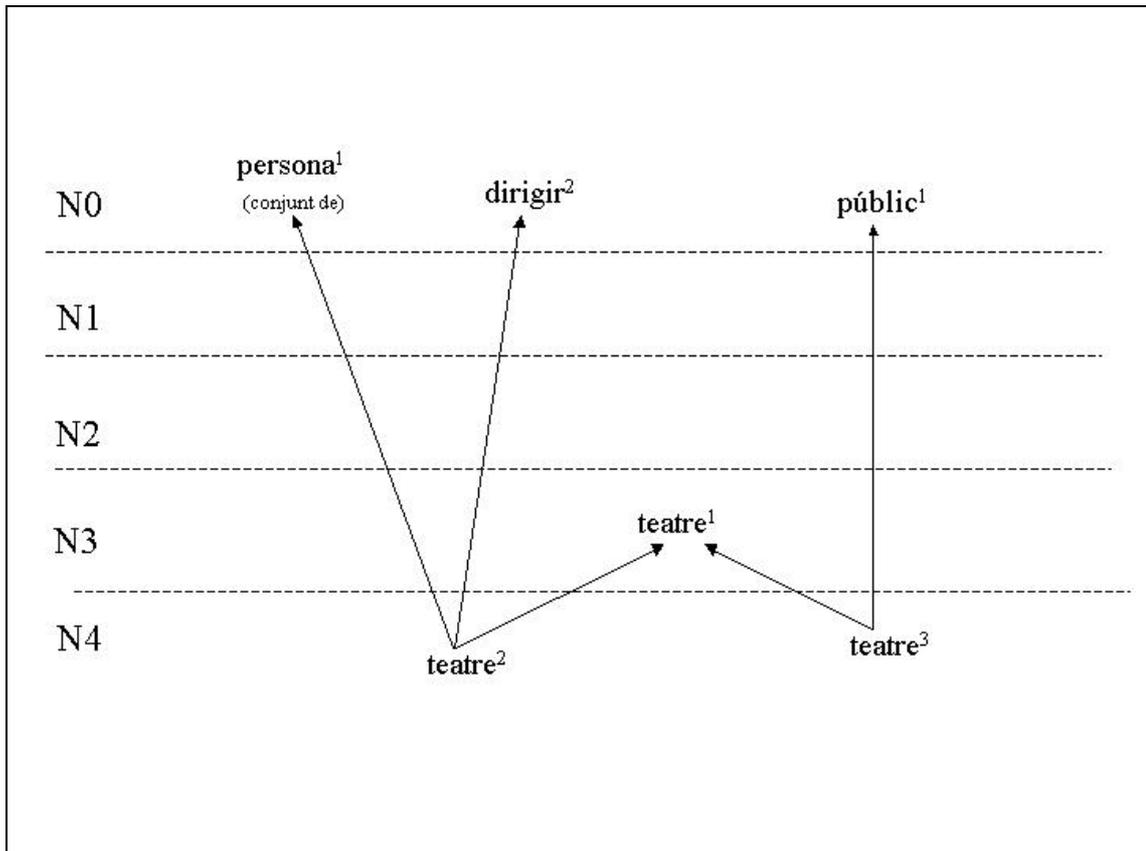


Figura 8.4. Representació per nivells de la descripció de *teatre* (2).

La resta d'accepcions requereix la inclusió prèvia en la metallengua de descripció d'altres elements; l'accepció 'Conjunt d'obres literàries per a ser representades' fa necessària la inclusió en la metallengua dels components que permetin l'expressió de fórmules com ara *obra literària*, amb un grau de complexitat semàntica considerable, bàsicament per raó de la presència de l'adjectiu *literari*, semànticament bastant complex. Es pot arribar a la caracterització de *literari* com a element de la metallengua de definició a partir del següent itinerari:

¹**bellesa**¹ *f.* (...) ₀Conjunt¹ de ₀qualitats¹ que fan que alguna ₀cosa¹ ₀produeixi¹ un ₀sentiment¹ ₀agradable¹ en qui l'₀observa¹.

²**art**¹ *m.* (...) ₀Activitat¹ ₀humana¹ que té per ₀objecte² l'₀expressió¹ de la ¹bellesa¹ sobre un ₀mitjà¹.

³**artístic**¹ *adj.* (...) ₀Pertanyent¹ a l'₂art¹.

4literatura¹ *f.* (...) ₀Activitat¹ ₃artística¹ que té l'₁escriptura¹ com a ₀mitjà¹ d'₀expressió¹.
1escriptura¹ *f.* (...) ₀Fet¹ d'₀escriure¹.
3literari¹ *adj.* (...) ₀Pertanyent¹ a la ₄literatura¹.

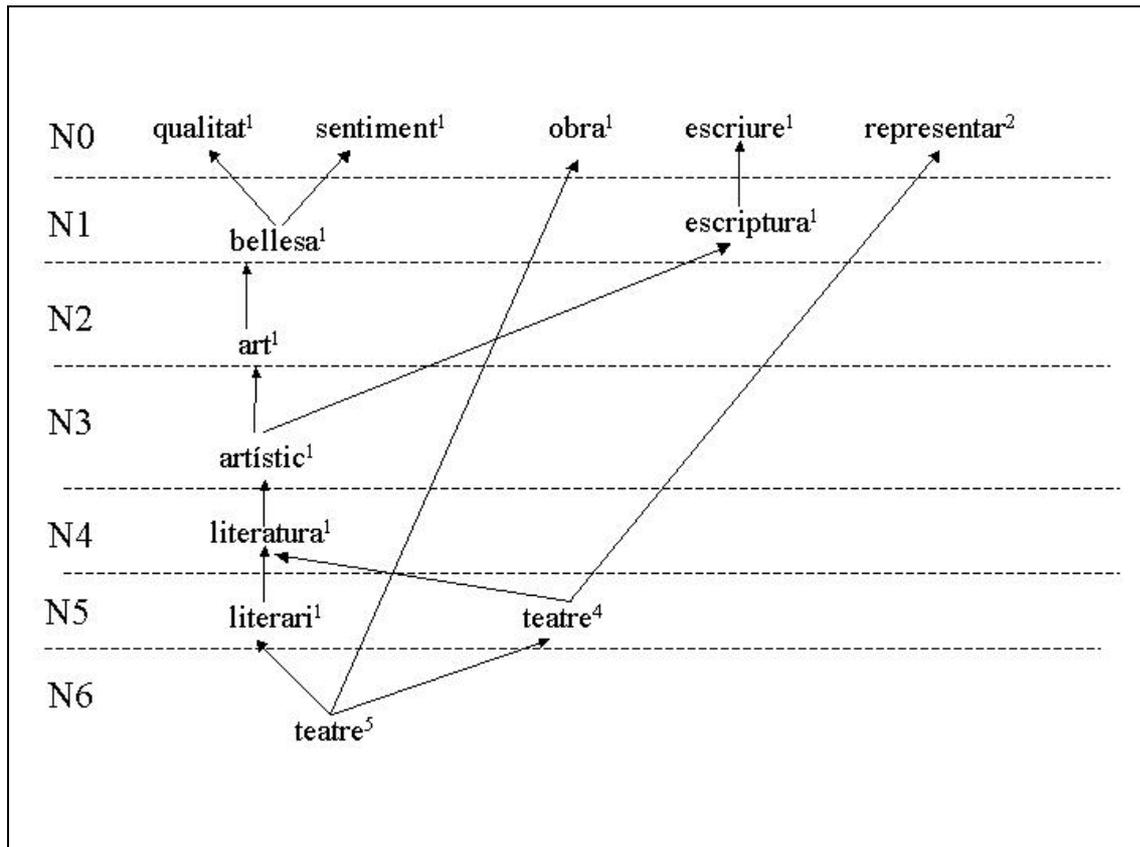


Figura 8.5. Representació per nivells de la descripció de *teatre* (3).

Un cop establerts aquests elements previs de descripció, es pot caracteritzar ₅teatre⁴ i, a partir d'aquest element, caracteritzar ₆teatre⁵:

5teatre⁴ *m.* (...) ₄Literatura¹ ₀destinada¹ a ser ₀representada².
6teatre⁵ *m.* (...) [_~ d'X] Conjunt d'obres de ₅teatre⁴ d'X

L'accepció 'activitat pròpia dels actors' ('professió d'actor', al DIEC) requereix la

descripció prèvia d'*actor* que, al seu torn, requereix la descripció de *personatge*, la descripció del qual integra també el component *obra literària* descrit anteriorment:

personatge¹ *m.* (...) ₀Persona¹ d'una ₀obra¹ ₅literària¹.

actor¹ *m.* (...) ₀Persona¹ que ₀actua¹ en un ₀espectacle¹ com si fos un ₆personatge¹ determinat.

teatre⁶ *m.* (...) ₀Activitat¹ ₀pròpia¹ dels ₇actors¹

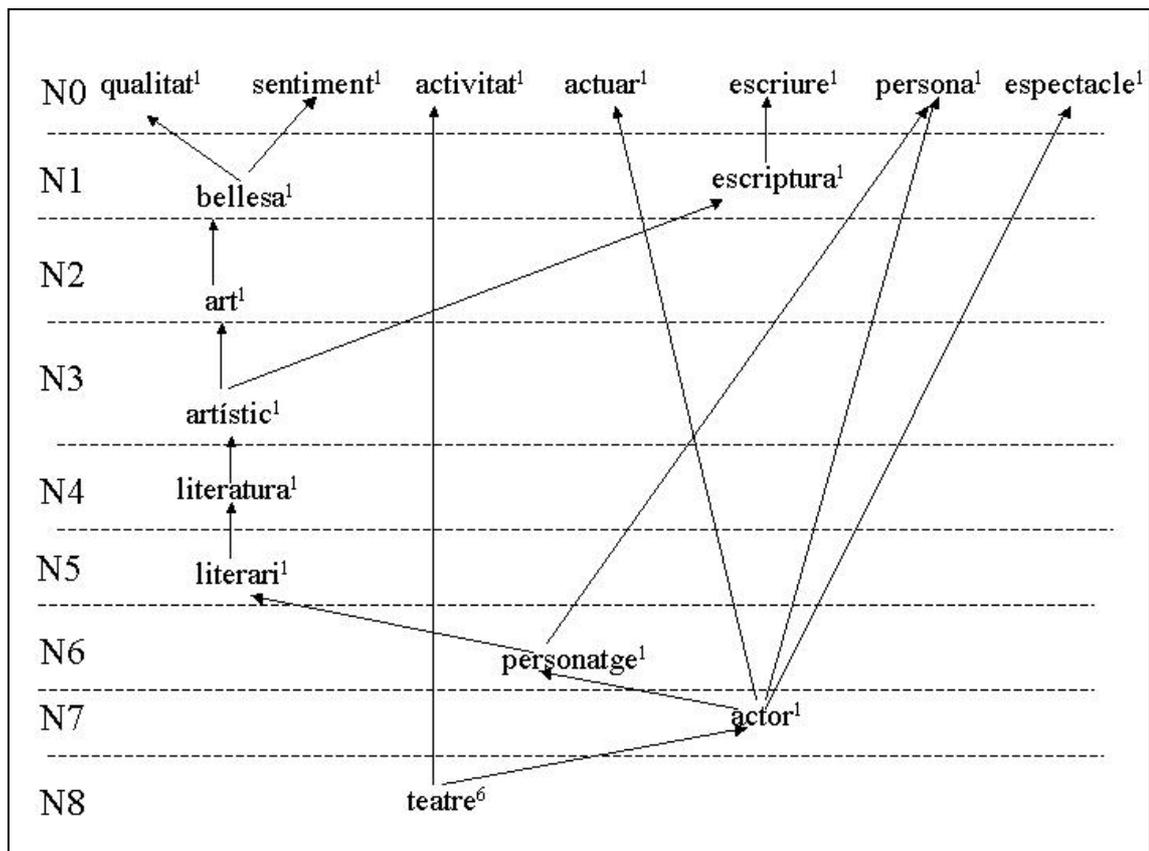


Figura 8.6. Representació per nivells de la descripció de *teatre* (4).

La major complexitat que suposa arribar a les descripcions de ₅teatre⁴ i de ₆teatre⁵, d'una banda, i de ₈teatre⁶, de l'altra, pot veure's representada gràficament a les figures 8.5 i 8.6, en què s'assenyalen els principals vincles semàntics d'aquests elements amb altres elements de la metallengua pertanyents a nivells de descripció inferiors.

Es pot inferir a partir de l'exemple precedent que la redacció d'un diccionari a partir del sistema proposat integra millor, ja des de la mateixa fase de redacció, la complexitat de les relacions entre el significat conceptual dels lexemes, la qual cosa condueix, de fet, a la resolució de la majoria dels problemes de descripció associats a la utilització d'una metallengua no organitzada i que presenta en molts casos recurrències que impliquen la incorporació en el diccionari d'un gran nombre de casos en què es produeixen desviacions estructurals.

El conjunt d'elements de la metallengua del diccionari constituïts a partir d'aquest sistema de descripció és, doncs, el següent:

- 3teatre**¹ *m.* (...) ₂Edifici¹ ₀destinat¹ a fer-hi ₀representacions².
- 4teatre**² *m.* (...) ₃Teatre¹ amb les ₀persones¹ que el ₀dirigeixen².
- 4teatre**³ *m.* (...) ₀Públic¹ d'un ₃teatre¹.
- 5teatre**⁴ *m.* (...) ₄Literatura¹ ₀destinada¹ a ser ₀representada².
- 6teatre**⁵ *m.* (...) [_~ *d'*X] Conjunt d'obres de ₅teatre⁴ d'X.
- 8teatre**⁶ *m.* (...) ₀Activitat¹ ₀pròpia¹ dels ₇actors¹
- 1teatre**⁷ *m.* (...) ₀Lloc¹ on s'₀acompleix¹ un ₀esdeveniment¹.

Organitzada en forma d'article de diccionari, aquesta informació podria tenir la següent forma:

teatre *m.* (...) **1.** Edifici¹ destinat¹ a fer-hi representacions². || **2.** Teatre¹ amb les persones¹ que el dirigeixen². || **3.** Públic¹ d'un teatre¹. || **4.** Literatura¹ destinada¹ a ser representada² || **5.** [_~ *d'*X] Conjunt d'obres de teatre⁴ d'X. || **6.** Activitat¹ pròpia¹ dels actors¹ || **7.** Lloc¹ on s'acompleix¹ un esdeveniment¹.

Fins i tot, aquesta informació es podria representar en el format final del diccionari sense els superíndexs de desambiguació, la qual cosa conferiria a la representació final una forma idèntica a la d'un diccionari convencional:

teatre *m.* (...) **1.** Edifici destinat a fer-hi representacions. || **2.** Teatre amb les persones que el dirigeixen. || **3.** Públic d'un teatre. || **4.** Literatura destinada a ser representada || **5.** [_~ *d'*X] Conjunt d'obres de teatre d'X. || **6.** Activitat pròpia dels actors || **7.** Lloc on s'acompleix un esdeveniment.

A partir d'una descripció sistemàtica i regular dels diferents sentits de l'entrada **teatre** s'arriba, doncs, a una caracterització semàntica interna coherent des del punt de vista de l'estructura semàntica del diccionari, i que pot ser representada externament amb la mateixa forma amb què ho faria un diccionari convencional. L'interès primordial de la sistematicitat en la descripció radica en el fet, però, que *internament*, les dades del diccionari estan organitzades a partir d'un sistema explícit que projecta una estructura semàntica del lèxic màximament coherent des del punt de vista de la descripció lexicogràfica. L'estructura de dades amb què es dóna forma al diccionari (la base de dades lèxica, el diccionari electrònic, o el nom que se li vulgui donar) que hi ha a la base de tot projecte de redacció d'un diccionari modern, conté en aquest cas tota la informació necessària per a garantir en un grau màxim les seves possibilitats de reutilització i, a més, aplica un rigorós control per a l'evitació de les circularitats i de les redundàncies en la metallengua de descripció.

9 Conclusions

El panorama de canvi i d'avenç metodològic en què se situa actualment la lexicografia (veg. § 1.1) demana l'aportació de recerca teòrica i la formulació de propostes per a la sistematització descriptiva. Entre aquestes necessitats, figura en un lloc destacat l'explicitació de sistemes de definició viables que conduixin a diccionaris més coherents des del punt de vista de la seva estructura conceptual.

Al llarg d'aquest treball m'he proposat l'objectiu principal (§ 1.3) d'arribar a la formulació d'un sistema de definició integrat amb una metallengua adequada a la descripció lexicogràfica del català. Per a la consecució d'aquest objectiu he plantejat a l'inici del treball dues línies fonamentals d'actuació: d'una banda, la reflexió i l'anàlisi teòrica sobre certs aspectes de la descripció lexicogràfica (la delimitació de les entrades, les aportacions de la lingüística i dels models teòrics del lèxic, i la naturalesa de la definició lexicogràfica) i, d'altra banda, la demostració d'una sèrie de tres hipòtesis prèvies plantejades a l'inici del treball.

En la primera línia d'actuació, aquest treball fa tres aportacions concretes, de cara mostrar que l'estructura dels diccionaris reflecteix l'estructura del lèxic, tant pel que fa a l'inventari de les seves unitats com a les relacions semàntiques, morfològiques, etc. que estableixen els seus elements entre si. Aquestes aportacions són:

1. L'anàlisi del concepte d'entrada de diccionari i la formulació, sobre bases exclusivament sincròniques, d'uns criteris explícits i inequívocs per a la seva delimitació a partir del concepte lingüístic de lexema. Aquesta formulació ha estat valorada críticament i adaptada a la casuística específica del català.
2. L'exposició dels principals punts de contacte i de divergència entre la idea que la lingüística té del lèxic i la que es desprèn dels diccionaris. Aquesta exposició mostra amb claredat que els diccionaris poden millorar-se a partir dels progressos que ha

experimentat la lingüística teòrica i els models de lèxic que han estat en molts casos implementats en lèxics electrònics. I encara, que és possible la normalització dels aspectes més fonamentals de la lexicografia, de cara a integrar en una mateixa operació descriptiva els processos de constitució dels diccionaris convencionals i dels lèxics electrònics, respectivament.

3. L'estudi aprofundit de les implicacions metodològiques que presenta la descripció lexicogràfica pel que fa a la formulació de definicions. Amb aquest estudi he analitzat els diferents tipus de condicions lingüístiques que planteja l'equivalència entre definit i definició, amb la intenció de formular propostes concretes de representació per a facilitar la verificació d'aquesta relació. La demostració fonamental d'aquesta exposició és que no és possible de desvincular la definició lexicogràfica de les característiques d'inserció sintàctica que presenten els lexemes. L'establiment d'aquesta correspondència entre el *patró* sintàctic del definit i els components (intrínsecs i extrínsecs) que formen part de la definició comporta les condicions d'explicitud que tota descripció lingüística ha de satisfer com a pressupòsit metodològic.

La segona línia d'actuació ha concentrat l'objectiu fonamental del treball, centrat en l'elaboració d'una proposta metodològica original per a la descripció lexicogràfica del català, i m'ha conduït a l'execució dels següents passos:

1. Determinació d'un vocabulari controlat apte per al català.
2. Obtenció d'un conjunt nuclear d'aquest vocabulari que forma la base de la metallengua de definició.
3. Establiment d'un sistema de definició amb els mecanismes de projecció d'aquesta metallengua per a la descripció global del lèxic.

El vocabulari controlat, que no ha estat un objectiu específic d'aquest treball, sinó un punt de partida per a l'obtenció de resultats metodològicament més avançats, ha estat constituït a partir de la informació aportada pel DFDC sobre freqüències de lemes del CTILC. La selecció freqüencial, que satisfà els criteris de constitució dels vocabularis

controlats de l'anglès, ha estat matisada amb un creuament amb la selecció a partir d'una altra informació aportada pel DFDCC: l'*ús*. Amb aquest creuament s'eviten els possibles problemes de selecció estrictament freqüencial, que pot estar excessivament mediatitzada pels criteris de tria de textos, i per l'equilibri cronològic i tipològic dels textos del corpus. L'adequació del vocabulari a les finalitats específicament lexicogràfiques s'ha aconseguit a través d'un segon creuament dels resultats obtinguts amb la llista de lemes més freqüents extreta d'un corpus que conté totes les definicions del DIEC. El resultat final en el procés de determinació del vocabulari controlat consisteix en una llista de 3.398 lemes (veg. Apèndix I.10). El fet que aquesta llista sigui significativament més extensa que la dels vocabularis controlats establerts per a l'anglès (que estan al voltant de les 2.000 unitats) determina la fiabilitat del vocabulari com a eina de descripció lexicogràfica i, sobretot, permet suposar que pot ser utilitzat com un punt de partida en projectes lexicogràfics basats en els mateixos principis que el LDOCE o el CIDE.

A fi d'obtenir una metallengua de definició (o més exactament, la base d'una metallengua de definició) he completat la llista del vocabulari controlat amb els articles del DIEC que es refereixen a cadascuna de les seves unitats. A partir d'una sèrie d'operacions basades en *a)* l'eliminació dels elements del vocabulari innecessaris, *b)* l'associació de cada element del vocabulari als sentits descriptivament necessaris i *c)* la regularització de les definicions, s'estableix un conjunt mínim consistent en una sèrie d'elements de descripció que té la característica de ser *tancada* (és a dir, sense la presència d'elements de descripció externs al mateix conjunt descrit). La integració d'aquest nivell en un conjunt de principis metodològics (un sistema de definició) i la formulació d'una sèrie de principis simples de projecció de l'estructura axiomàtica són els complements necessaris perquè la descripció lexicogràfica implícita en les definicions de diccionari reflecteixi una estructura coherent i, alhora, descriptivament adequada.

Així, es formula un mètode de redacció lexicogràfica que satisfà els requeriments plantejats per l'objectiu principal del treball sobre l'adequació descriptiva i la

coherència estructural del diccionari, ja que no limita *a priori* la metallengua total de definició i, alhora, redueix al mínim del conjunt nuclear les seves desviacions estructurals. Aquest sistema es basa en una estructura de nivells de descripció que s'estableixen en funció dels nivells anteriors i que concentra tots els problemes descriptius del diccionari en el conjunt nuclear de la metallengua (anomenat nivell 0), de forma similar a la hipòtesi de descripció plantejada per Weinreich (1962).

He arribat, doncs, a l'explicitació d'un sistema que satisfà l'objectiu principal del treball i que, per tant:

- Supera els problemes descriptius dels vocabularis controlats, però alhora n'aprofita els avantatges pel que fa l'organització estructurada del diccionari.
- Permet descriure de manera organitzada el significat lèxic i concentrar en un petit grup de mots (mínim, d'acord amb els arguments posats en aquest treball) totes les desviacions estructurals i metodològiques del diccionari.
- Integra per a la seva aplicació la caracterització completa d'un nivell axiomàtic (el nivell 0, en què s'admet la circularitat en les descripcions) que és la base de la projecció estructurada de tot el diccionari.

Aquest sistema, a més, té l'avantatge de ser integrable en el sistema de definicions lexicogràfiques d'un diccionari convencional, ja que utilitza la llengua natural com a eina de descripció bàsica i no s'allunya fonamentalment del sistema de presentació de les dades de la lexicografia convencional.

La formulació del sistema s'ha fet d'una manera completa. Pel que fa a la metallengua, aquest treball només s'ha ocupat de descriure de manera completa la seva base fonamental (el nivell 0). La forma final de la metallengua total d'una obra lexicogràfica elaborada a partir d'aquest sistema pot ser, doncs, variable en funció dels criteris de descripció, de les convencions de representació que s'utilitzin i de l'estructura i el pla general de l'obra. Atès que el nombre de nivells de descripció no està limitat a priori,

podríem preguntar-nos quants nivells de descripció poden necessitar-se en la redacció d'un diccionari perquè la descripció lexicogràfica reflecteixi una certa organització semàntica del lèxic. Aquesta qüestió no té una resposta senzilla ni desvinculable del mateix procés de redacció del diccionari. L'aplicació pràctica d'aquest sistema tindria l'interès de determinar sobre bases empíriques el nombre de nivells i l'estructuració general del lèxic que es desprèn de l'aplicació del sistema proposat. A més podria permetre extreure conclusions d'interès per a la validació del sistema, i del seu grau d'eficiència en l'establiment d'estructures semàntiques i l'anàlisi comparativa d'aquestes estructures amb les que s'han establert des de la semàntica lèxica.

En tot cas, crec que queda demostrat que aquest sistema de descripció (que implica la progressió no lineal de la descripció lexicogràfica i l'estructuració rigorosa de la metallengua de descripció), té la seva base en una concepció integrada de la lexicografia que es justifica tant en la dimensió teòrica com en la pràctica.

Normalitzant de manera adequada la producció de diccionaris convencionals, es tendeix cap a una descripció lexicogràfica que integra les aplicacions humanes i les computacionals. El cost i l'esforç necessari per a la redacció de diccionaris és elevat, de manera que l'aprofitament màxim dels recursos i la concepció multifuncional esdevenen una necessitat de primer ordre. El fet que actualment hi hagi nombrosos punts de contacte entre la lexicografia computacional i la convencional, permet preveure que l'aprofitament d'una sola activitat de descripció integrada pot resultar, en un futur pròxim, una pràctica habitual de constitució de recursos lèxics. Per aconseguir aquest objectiu cal basar-nos en delimitacions rigoroses de les unitats a descriure, i fonamentar la descripció sobre bases empíriques fiables, tot respectant el mètode lexicogràfic basat en l'equivalència parafràstica; a més cal resoldre les incoherències en l'estructura semàntica dels diccionaris convencionals, a través de la proposta d'unes metallengües de descripció basades en les diferents llengües naturals, i integrades en sistemes explícits de formulació de definicions que redueixin al mínim les desviacions habituals de la lexicografia convencional.

En l'actualitat l'estat de desenvolupament de la lexicografia computacional permet de pensar en un plantejament integrat de la descripció lexicogràfica capaç d'abordar la producció de diccionaris i de lèxics electrònics com a subproductes d'un mateix procés d'anàlisi lexicogràfica. Això té una doble justificació: teòrica i pràctica. Teòrica, en el sentit que afegeix sistematicitat descriptiva a la lexicografia per a ús humà, i volum de dades a la lexicografia computacional. Pràctica, en el sentit que els costos associats a la redacció d'un diccionari i a la constitució de lèxics electrònics són prou grans perquè sigui necessari de plantejar-se la producció integrada de les tasques que resulten comunes a totes dues disciplines. Un motiu afegit, encara, és el fet que, cada cop més, els anomenats *recursos lingüístics* són objecte d'una gran interès per part del món de l'enginyeria lingüística, per l'alt valor que tenen com a components de qualsevol aplicació computacional. L'evolució immediata de les tècniques de redacció dels diccionaris haurà de tenir en compte aquestes tendències. Aquest treball té la pretensió que la lexicografia catalana, necessitada de renovacions metodològiques profundes, no quedi desvinculada d'aquest procés general de canvis.

Bibliografia

Obres de consulta *

AARTS, J. & MEIJS, W. (eds.) (1986), *Corpus Linguistics II: New Studies in the Analysis and Exploitation of Computer Corpora*, Amsterdam: Rodopi

AHUMADA, I. (1989), *Aspectos de lexicografía teórica*, Granada: Universidad de Granada.

AITCHISON, J. (1987), *Words in the Mind: An Introduction to the Mental Lexicon*, Oxford: Blackwell.

AKKERMAN, E. (1989), “An independent analysis of the LDOCE grammar coding system”, in Boguraev, B. & Briscoe, T. (eds.) (1989), 65-81.

ALLAN, K. (1986), *Linguistic Meaning* (2 vols.), London – New York: Routledge & Kegan Paul.

ALLÉN, S. (1981), “The Lemma-Lexeme Model of the Swedish Lexical Database”, *Quantitative Linguistics*, 13, 376-387.

ALSHAWI, H. (1989), “Analysing the dictionary definitions”, in Boguraev, B. & Briscoe, T. (eds.) (1989), 153-169.

ALSHAWI, H. & al. (1989) [H. ALSHAWI, B. BOGURAEV, D. CARTER], “Placing the dictionary on-line”, in Boguraev, B. & Briscoe, T. (eds.) (1989), 41-63.

* Aquesta bibliografia recull els títols que he consultat en el procés d'elaboració d'aquest treball, independentment que apareguin o no citats en el text definitiu. Les traduccions i indicacions de reproducció de textos que recullo a continuació d'algunes referències són exclusivament les que he utilitzat per a les indicacions de paginació.

ALSINA, V. (1999), “La funció dels exemples en el *Diccionari General de la llengua catalana*. Aportació d’informació gramatical”, in *Estudis de llengua i literatura catalanes*, XX, 203-209.

ALSINA, V. & DECESARIS, J. (1998), “Morphological Structure and Lexicographic Definitions: The Case of *-ful* and *like*”, in *Euralex’98 Proceedings*, Leiden: University of Leiden, 545-554.

ALSINA, V. & DECESARIS, J. (eds.) (2001), *Estudis de lexicografia*, Barcelona: Universitat Pompeu Fabra.

ALSINA, V & al. (1994) [V. ALSINA, L. SOL, J. SOLER], “Una tipologia de la definició lexicogràfica aplicada al DGLC”, *Caplletra*, 17, 311-324.

ALSTON (1971), “How does one Tell whether a Word has one, several or many Senses?”, in Steinberg, D. & Jakobits, L. A. (eds.) (1971), 35-47.

ALTENBERG, B. (1991), “The London-Lund Corpus: Research and Applications”, in *Using Corpora. Seventh Annual Conference of the UW Centre for the New OED and Text Research*, Oxford, 71-83.

ALVAR EZQUERRA, M. (1976), *Proyecto de lexicografía española*, Planeta: Barcelona.

ALVAR EZQUERRA, M. (1983), *Lexicología y lexicografía*, Salamanca: Almar.

ALVAR EZQUERRA, M. (1991), “Desarrollos actuales en lexicografía automatizada del español”, in Vidal Beneyto, J. (ed.) (1991), 219-240.

ALVAR EZQUERRA, M. (1993), *Lexicografía descriptiva*, Barcelona: Biblograf.

ALVAR EZQUERRA, M. (1994), “La forma de los diccionarios a la luz del signo lingüístico”, in Hernández, H. (coord.) (1994), 3-13.

ALVAR EZQUERRA, M. (1999), “La redacción lexicográfica asistida por ordenador: dificultades y deseos”, in Blecua, J. M. & al. (eds.) (1999), 215-226.

ALVAR EZQUERRA, M. (2000), “Retos en la lexicografía del español”, in Ruhstaller, S. & Prado, J. (eds.) (2000), 37-55.

ALVAR EZQUERRA, M. & al. (1992) [M. ALVAR, M. J. BLANCO, J. M. LÓPEZ, M. F. PÉREZ], “La gramática interna en los diccionarios VOX”, in *Euralex’90. Proceedings*, Barcelona: Biblograf, 427-435.

ANDERSON, S. R. (1992), *A-Morphous Morphology*, Cambridge: Cambridge University Press.

- ANTELMÍ, D. & ROVENTINI, A. (1992), "Semantic Relationships within a Set of Verbal Entries in the Italian Lexical Database", in *Euralex'90. Proceedings*, Barcelona: Biblograf, 247-255.
- APRESJAN, J. (1974), "Regular Polysemy", *Linguistics*, 142, 5-32.
- APRESJAN, J. (1997), "A propósito de las definiciones en la Escuela Semántica de Moscú", in Luque, J. de D. & Pamies, A. (eds.) (1997), 1-8.
- APRESJAN, J. (2000), *Systematic Lexicography*, Oxford: Oxford University Press.
- ARONOFF, J. (1976), *Word Formation in Generative Grammar*, Cambridge: The MIT Press.
- ARTOLA, X. & EVRARD, F. J. (1992), "Dictionnaire intelligent d'aide à la compréhension", in *Euralex'90. Proceedings*, Barcelona: Biblograf, 45-57.
- ASTON, G. & BURNARD, L. (1998), *The BNC Handbook. Exploring the British National Corpus with SARA*, Edinburgh: Edinburgh University Press.
- ATKINS, B. T. S. (1991), "Building a Lexicon. The Contribution of Lexicography", in Boguraev, B. (ed.) (1991), 167-204.
- AYTO, J. R. (1983), "On specifying meaning", in Hartmann, R. R. K. (ed.) (1983), 89-98.
- BADIA, T. (1994), "Lexicografía i models lingüístics: les teories lingüístiques i el lèxic", *Caplletra*, 17, 15-46.
- BARGALLÓ, H. (1999), "La información gramatical en los diccionarios didácticos bilingües del español", in Vila, M. N. & al. (eds.) (1999), 15-42.
- BATTANER, M. P. (1999), "Estructura del vocabulario a través de las definiciones lexicográficas: las entradas *re-* del DRAE 1992", in Blecua, J. M. & al. (eds.) (1999), 227-257.
- BAUER, L. (1983), *English Word-formation*, Cambridge: Cambridge University Press.
- BAYLON, C. & FABRE, P. (1978), *La sémantique*, Paris: Nathan.
- BAYLON, C. & MIGNOT, X. (1995), *Sémantique du langage*, Paris: Nathan.
- BEARD, R. (1988), "On the separation of derivation from morphology: toward a Lexeme-Morpheme-Based Morphology", *Quaderni di Semantica*, 9, 3-59.
- BÉJOINT, H. (2000), *Modern Lexicography*, Oxford: Oxford University Press.

- BENSON, M. & al. (1986) [M. BENSON, E. BENSON, R. ILSON], *Lexicographic Description of English*, Amsterdam: John Benjamins.
- BIBER, D. & al. (1998) [D. BIBER, S. CONRAD, R. REPPEN], *Corpus Linguistics. Investigating Language Structure and Use*, Cambridge: Cambridge University Press.
- BLECUA, J. M. & al. (eds.) (1999) [J. M. BLECUA, G. CLAVERÍA, C. SÁNCHEZ, J. TORRUELLA], *Filología e informática. Nuevas tecnologías en los estudios filológicos*, Barcelona: Universitat Autònoma de Barcelona.
- BLOOMFIELD, L. (1933), *Language*, New York: Holt, Rinehart and Winston.
- BOLINGER, D. (1965), "The Atomization of Meaning", *Language*, 41, 555-573.
- BOGAARDS, P. (1991), "Word Frequency in the Search Strategies of French Dictionary Users", *Lexicographica*, 7, 202-212.
- BOGAARDS, P. (1996), "Dictionaries for Learners of English", *International Journal of Lexicography*, 9, 4, 277-320.
- BOGURAEV, B. (1991), "Building a Lexicon: The Contribution of Computers", in Boguraev, B. (ed.) (1991), 227-260.
- BOGURAEV, B. (ed.) (1991), "Building a Lexicon", *International Journal of Lexicography (Special Issue)*, 4, 3.
- BOGURAEV, B. & BRISCOE, T. (eds.) (1989), *Computational Lexicography for Natural Language Processing*, London: Longman.
- BOGURAEV, B. & al. (1992) [B. BOGURAEV, T. BRISCOE, J. CARROLL, A. COPESTAKE], "Database Models for Computational Lexicography", in *Euralex'90. Proceedings*, Barcelona: Bibliograf, 59-78.
- BOSQUE, I. (1982), "Sobre la teoría de la definición lexicográfica", *Verba. Anuario Galego de Filoloxía*, 9, 105-123.
- BOSQUE, I. (1999), "El sintagma adjetival. Modificadores y complementos del adjetivo. Adjetivo y participio", in Bosque, I. & Demonte, V. (dirs.), 217-310.
- BOSQUE, I. (2000), "Objetos que esconden acciones. Una reflexión sobre la sincategorematicidad", in Cabré, M. T. & Gelpí, C. (eds.) (2000), 15-30.
- BOSQUE, I. & DEMONTE, V. (dirs.) (1999), *Gramática de la lengua española*, Madrid: Real Academia Española - Espasa.
- BOUILLON, P. & BUSA, F. (eds.) (2001), *The Language of Word Meaning*, Cambridge: Cambridge University Press.

- BOULANGER, J. C. (2000), “La nature et le dessin des mots dans les dictionnaires de langue”, in Cabré, M. T. & Gelpí (eds.) (2000), 73-94.
- BRAND, G. (1975), *Die grundlegenden Texte von Ludwig Wittgenstein*, Frankfurt am Main: Suhrkamp Verlag.
- BRUCART, J. M. (1999), “La elipsis”, in Bosque, I. & Demonte, V. (dirs.) (1999), 2787-2863.
- BUITELAAR, P. (1998), “CoreLex: Systematic Polysemy and Underspecification”, PhD. dissertation, Brandeis University.
- BULYGINA, T. V. & SHMELEV, A. D. (1992), “Quantifiers in the Dictionary”, in *Euralex'90. Proceedings*, Barcelona: Bibliograf, 257-263.
- BURTON, T. L. & BURTON, J. (eds.) (1988), *Lexicographical and Linguistic Studies. Essays in Honour of G. W. Turner*, Suffolk: D. S. Brewer.
- BUSA, F. & al. (2001) [F. BUSA, N. CALZOLARI, A. LENCI], “Generative Lexicon and the SIMPLE Model: Developing Semantic Resources for NLP”, in Bouillon, P. & Busa, F. (eds.) (2001), 333-349.
- BUTLER, C. (1985), *Computers in Linguistics*, Oxford: Blackwell.
- BUTLER, C. (ed.) (1992), *Computers and Written Texts*, Oxford: Blackwell.
- CABRÉ, M. T. (1994a), *A l'entorn de la paraula (I). Lexicologia general*, València: Universitat de València.
- CABRÉ, M. T. (1994b), *A l'entorn de la paraula (II). Lexicologia catalana*, València: Universitat de València.
- CABRÉ, M. T. (ed.) (1990), *La lingüística aplicada*, Barcelona: Universitat de Barcelona.
- CABRÉ, M. T. (dir.) (1996), *Cicle de conferències 94-95. Lèxic, corpus i diccionaris*, Barcelona: Universitat Pompeu Fabra.
- CABRÉ, M. T. (dir.) (1997), *Cicle de conferències 95-96. Lèxic, corpus i diccionaris*, Barcelona: Universitat Pompeu Fabra.
- CABRÉ, M. T. (dir.) (1998), *Cicle de conferències 96-97. Lèxic, corpus i diccionaris*, Barcelona: Universitat Pompeu Fabra.
- CABRÉ, M. T. & RIGAU, G. (1986), *Lexicologia i semàntica*, Barcelona: Enciclopèdia Catalana.

CABRÉ, M. T. & GELPÍ, C. (eds.) (2000), *Cicle de conferències 97-98. Lèxic, corpus i diccionaris*, Barcelona: Universitat Pompeu Fabra.

CABRÉ, M. T. & al. (1994) [M. BOZZO, M. T. CABRÉ, A. EGEE, M. JUNCADELLA, C. SANTAMARIA], “El Nou diccionari normatiu de la llengua catalana”, *Caplletra*, 17, 201-218.

CALZOLARI, N. (1977), “An Empirical Approach to Circularity in Dictionary Definitions”, *Cahiers de Lexicologie*, XXXI, 118-128.

CALZOLARI, N. (1988), “The Dictionary and the Thesaurus can be Combined”, in Evens, M. (ed.) (1988), 75-96.

CALZOLARI, N. (1997), “Lexicon and Corpus: a Multi-faceted Interaction”, in Cabré, M. T. (dir.) (1997), 77-89.

CALZOLARI, N. & al. (1987) [N. CALZOLARI, E. PICCHI, A. ZAMPOLLI], “The Use of Computers in Lexicography and Lexicology”, in Cowie, A. (ed.) (1987), 55-77.

CANO, M. A. & al. (eds.) (1999) [M. CANO, J. MARTINES, V. MARTINES, J. J. PONSODA], *Bescanvi i identitat: interculturalitat i construcció de la llengua*, Alacant: Institut Interuniversitari de Filologia Valenciana, Universitat d’Alacant, Ajuntament de la Nucia.

CARR, M. (1997), “Internet Dictionaries and Lexicography”, *International Journal of Lexicography*, 10, 3, 209-230.

CARROLL, J. B. (1970), “An Alternative to Juilland’s Coefficient for Lexical Frequencies, and a Proposal for a Standard Frequency Index (SFI)”, *Computer Studies in the Humanities and Verbal Behaviour*, 3, 61-65.

CARSTAIRS-McCARTHY, A. (1992), *Current Morphology*, London: Routledge.

CASARES, J. (1950), *Introducción a la lexicografía española*, Madrid. [Reimpresió: 1969, Madrid: Consejo Superior de investigaciones Científicas.]

CASAS, M. (1999), *Las relaciones léxicas*, Tübingen: Max Niemeyer.

CASAS, M. & MUÑOZ, M. D. (1992), “La polisemia y la homonimia en el marco de las relaciones léxicas”, in Wotjak, G. (ed.) (1992), 134-158.

CERDÀ, R. (1994), “Lexicografia i intel·ligència artificial”, *Caplletra*, 17, 81-108.

CIFUENTES, J. L. (1992), “Polisemia y lexicografía”, in *Euralex’90. Proceedings*, Barcelona: Bibliograf, 265-272.

CHAURAND, J. & MAZIÈRE, F. (eds.) (1990), *La définition*, Paris: Larousse.

- CHEN, N. J. & CHANG, J. S. (1999), "Integrating Machine Readable Dictionaries and Thesaurus for Conceptual Context Representation of Word Sense", in Viegas, E. (ed.) (1999), 189-225.
- CHIERCHIA, G. & MCCONNELL-GINET, S. (1990), *Meaning and Grammar. An Introduction to Semantics*, Cambridge: MIT Press.
- CHOMSKY, N. (1957), *Syntactic Structures*, The Hague: Mouton.
- CHURCH, K. W. & HANKS, P. (1990), "Word Association Norms, Mutual Information, and Lexicography", *Computational Linguistics*, 16, 22-29.
- CLAVERÍA, G. (2000), "El problema de la homonimia en la lexicografía española", in Ruhstaller, S. & Prado, J. (eds.) (2000), 365-375.
- COLLINOT, A. & MAZIÈRE, F. (1997), *Un prêt à parler: le dictionnaire*, Paris: Presses Universitaires de France.
- COLON, G. (1976), "La lexicografía catalana: realitzacions i esperances", in *Actes del Quart Col·loqui Internacional de Llengua i Literatura Catalanes*, Barcelona: Abadia de Montserrat, 11-35.
- COLON, G. (1981), "Sinonímia i diatopisme", *Randa*, 11, 1981, 45-62.
- COPESTAKE, A. (1992), "The Representation of Lexical Semantic Information", PhD. dissertation, Berkeley: University of California.
- COPESTAKE, A. & BRISCOE, T. (1996), "Semi-productive Polysemy and Sense Extension", in Pustejovsky, J. & Boguraev, B. (eds.) (1996), 15-67.
- CORBIN, D. (1983), "Le monde étrange des dictionnaires (4). La créativité lexicale, le lexicographe et le linguiste", *Lexique*, 2, 25-64.
- CORBIN, D. (1987), *Morphologie dérivationnelle et structuration du lexique*, Tübingen: Max Niemeyer, 2 v.
- CORBIN, D. (1991), "La formation des mots: structures et interprétations", *Lexique*, 10, 7-32.
- CORBIN, P. (1998), "La lexicographie française est-elle en panne?", in Cabré, M. T. (dir.) (1998), 83-112.
- COSERIU, E. (1981), *Principios de semántica estructural*, Madrid: Gredos.
- COSERIU, E. (1987), *Gramática, semántica, universales*, Madrid: Gredos.

COWIE, A. (1992), "Verb Syntax in the Revised *Oxford Advanced Learner's Dictionary: Descriptive and Pedagogical Information*", in *Euralex'90. Proceedings*, Barcelona: Bibliograf, 341-347.

COWIE, A. (1998), *Phraseology. Theory, Analysis and Applications*, Oxford: Oxford University Press.

COWIE, A. (1999), "Learner's dictionaries in a historical and a theoretical perspective", in Herbst, T. & Popp, K. (eds.) (1999), 3-13.

COWIE, A. (ed.) (1987), *The Dictionary and the Language Learner*, Tübingen: Max Niemeyer.

CRUSE, D. A. (1986), *Lexical Semantics*, Cambridge: Cambridge University Press.

CRUSE, D. A. (2000), "Aspects of the Micro-structure of Word Meanings", in Ravin, Y. & Leacock, C. (eds.) (2000), 30-51.

DE MAURO, T. (1967), *Ludwig Wittgenstein. His Place in the Development of Semantics*, Dordrecht: D. Reidel.

DORR, B. J. & JONES, D. (1999), "Acquisition of Semantic Lexicons. Using Word Sense Disambiguation to Improve Precision", in Viegas, E. (ed.) (1999), 70-97.

DODD, W. S. (1994a), "Tendencias en la lexicografía anglosajona: Los diccionarios monolingües para usuarios extranjeros", in Hernández, H. (coord.) (1994), 39-59.

DODD, W. S. (1994b), "La lexicoinformática", in Hernández, H. (coord.) (1994), 121-137.

DOLAN W. & al. (2000) [W. DOLAN, L. VANDERWENDE, S. RICHARDSON], "Polysemy in Broad-Coverage Natural Language Processing Systems", in Ravin, Y. & Leacock, C. (eds.) (2000), 178-204.

DOMÈNECH, J. M. & al. (1993) [J. DOMÈNECH, J. SOLER, A. EGEA], "Estructura y estrategias de una base de datos textual: BDTLC-DCC", *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 13, 201-215.

DOWTY, D. (1979), *Word Meaning and Montague Grammar*, Dordrecht: D. Reidel.

DOWTY, D. (2000), "«The Garden Swarms with Bees» and the Fallacy of «Argument Alternation»", in Ravin, Y. & Leacock, C. (eds.) (2000), 111-128.

DUBOIS, J. & DUBOIS, C. (1971), *Introduction à la lexicographie. Le dictionnaire*, Paris: Larousse.

- DUVAL, A. (1992), "From the Printed Dictionary to the CD-ROM", in *Euralex'90. Proceedings*, Barcelona: Biblograf, 59-78.
- EAGLES (1996), "Synopsis and Comparison of Morphosyntactic Phenomena Encoded in Lexicons and Corpora. A Common Proposal and Applications to European Languages", EAGLES Document EAG-CLWG-MORPHSYN/R, version of May 1996.
- ESCOBEDO, A. (1994), *Estudios de lexicología y lexicografía*, Almería: Universidad de Almería.
- ESPINAL, M. T. (1988), *Significat i interpretació*, Barcelona: Publicacions de l'Abadia de Montserrat.
- ESTEVE, F. & al. (1998) [F. ESTEVE, J. FERRER, L. MARQUET, J. MOLL], "El diccionari de l'Institut. Una aproximació sistemàtica", *Els marges*, 60, 5-96.
- EVENS, M. (ed.) (1988), *Relational Models of the Lexicon*, Cambridge: Cambridge University Press.
- FABER, P. (1997), "La vertiente computacional del Modelo Lexemático-Funcional", in Luque, J. de D. & Pamies, A. (eds.) (1997), 73-87.
- FARINA, D. M. T. CR. (trad.) (1995) "L. V. Scherba: Towards a General Theory of Lexicography". *International Journal of Lexicography*, 8, 4, 314-350.
- FELICES, A. M. (1992), "The Integration Process of the Evaluation Classeme in a Dictionary", in *Euralex'90. Proceedings*, Barcelona: Biblograf, 349-357.
- FELLBAUM, C. (1990), "English Verbs as a Semantic Net", *International Journal of Lexicography*, 3, 4, 278-301.
- FELLBAUM, C. (1998), "A Semantic Network of English Verbs", in Fellbaum (ed.) (1998), 69-104.
- FELLBAUM, C. (2000), "Autotroponomy", in Ravin, Y. & Leacock, C. (eds.) (2000), 52-67.
- FELLBAUM, C. (ed.) (1998), *WordNet. An Electronical Lexical Database*, Cambridge: The MIT Press.
- FERNANDO, C. (1996), *Idioms and idiomaticity*, Oxford: Oxford University Press.
- FILLMORE, C. J. & ATKINS, B. T. S. (2000), "Describing Polysemy: The Case of «Crawl»", in Ravin, Y. & Leacock, C. (eds.) (2000), 91-110.

FIRTH, J. R. (1951): "Modes of meaning", *Essays and Studies*, 4, 118-149 [reproduït a J. R. Firth, *Papers in Linguistics 1934-51*, London: Oxford University Press, 1957, 190-215].

FIRTH, J. R. (1957): "A Synopsis of Linguistic Theory", *Studies in Linguistic Analysis*, Oxford: Philological Society, 1-31 [reproduït a F. R. Palmer (ed.), *Selected Papers of J. R. Firth*, London: Longman, 1968, 168-205].

FODOR, J. A. & LEPORE, E. (2001), "The Emptiness of the Lexicon: Critical Reflections on J. Pustejovsky's «The Generative Lexicon»", in Bouillon, P. & Busa, F. (eds.) (2001), 28-50.

FONTENELLE, T. (1992), "Automatic Extraction of Lexical-Semantic Relations from Dictionary Definitions", in *Euralex '90. Proceedings*, Barcelona: Bibliograf, 89-103.

FONTENELLE, T. (1996), "Ergativity, Collocations and Lexical Functions", in *Euralex '96. Proceedings*, Göteborg: Göteborg University.

FONTENELLE, T. (1997a), *Turning a Bilingual Dictionary into a Lexical-Semantic Database*, Tübingen: Max Niemeyer.

FONTENELLE, T. (1997b), "Using a Bilingual Dictionary to Create Semantic Networks", *International Journal of Lexicography*, 10, 4, 275-303.

FONTENELLE, T. (1998), "Discovering significant lexical functions in Dictionary Entries", in Cowie, A. (ed.) (1998), 189-207.

FORGAS, E. (coord.) (1996), *Léxico y diccionarios*, Tarragona: Universitat Rovira i Virgili.

FUCHS, C. (1982), *La paraphrase*, Paris: Presses Universitaires de France.

GARRIDO, J. (1992), "Lexical Entry Format and Contextual Information", in *Euralex '90. Proceedings*, Barcelona: Bibliograf, 273-280.

GECKELER, H. (1971), *Strukturelle Semantik und Worterfeldtheorie*, München: Wilhelm Fink.

GEERAERTS, D. (1985), "Les donées stéréotypiques, prototypiques et encyclopédiques dans le dictionnaire", *Cahiers de lexicologie*, XLVI, 27-43.

GEERAERTS, D. (1987), "Types of Semantic Information in Dictionaries", in Ilson, R. (ed.) (1987), 1-10.

GELBUKH, A. & SIDOROV, G. D. (2002), "Selección automática del vocabulario definidor en un diccionario explicativo", *Procesamiento del lenguaje natural*, 29, 55-62.

- GEORGACAS, D. J. (1976), "The Present State of Lexicography and Zgusta's «Manual of Lexicography»", *Orbis*, XXV, 2, 359-400.
- GERMAIN, C. (1981), *La sémantique fonctionnelle*, Paris: Presses Universitaires de France.
- GILLETTE, M. & WIT, E.-J. C. (1998), *What is Linguistic Redundancy?*, document electrònic: <http://www.stats.gla.ac.uk/~ernst/redundan.html>.
- GIURESCU, A. (1975), *Les mots composés dans les langues romanes*, The Hague: Mouton.
- GODDARD, C. (2000), "Polysemy: A Problem of Definition", in Ravin, Y. & Leacock, C. (eds.) (2000), 129-151.
- GOLD, D. L. (1980), "The Dictionary and Lexical Structure", *Babel*, 3, 152-160.
- GOLD, D. L. (1983), "Lexicographical metalanguage", *Babel*, 29, 134-151.
- GONZÁLEZ CALVO, J. M. (1982), "Consideraciones sobre la palabra como unidad lingüística", *Revista Española de Lingüística*, 12-2, 375-410.
- GONZÁLEZ REY, M. (1998), "Estudio de la idiomática en las unidades fraseológicas", in Wotjak, G. (ed.) (1998), 57-73.
- GOÑI, J. M. (1998), "Arquitectura para representación del conocimiento léxico en sistemas de procesamiento de lenguaje natural", Tesis doctoral, Madrid: Universidad Politécnica de Madrid.
- GRAY, J. C. (1986), "Creating the Electronic New Oxford English Dictionary", *Computers and the Humanities*, 20, 45-49.
- GREENBERG, J. H. (ed.) (1966), *Universals of Language*, Cambridge: MIT Press.
- GRISHMAN, R. (1986), *Computational linguistics*, Cambridge: Cambridge University Press.
- GROSS, M. (1991), "La elaboración de diccionarios electrónicos del francés", in Vidal Beneyto, J. (ed.) (1991), 185-202.
- GRUBER, T. R. (1993): "A Translation Approach to Portable Ontologies", *Knowledge Acquisition*, 5, 2, 199-220.
- GUILBERT, L. (1975), *La créativité lexicale*, Paris: Larousse.
- GUIRAUD, P. (1955), *La sémantique*, Paris: Presses Universitaires de France.

- GUTIÉRREZ, J. (1994), “Gramática y diccionario”, in *Actas del Congreso de la Lengua Española (Sevilla, 1992)*, Madrid: Instituto Cervantes, 637-656.
- GUTIÉRREZ, J. (1996), “Enciclopedia y diccionario”, in Forgas, E. (coord.) (1996), 133-159.
- GUTIÉRREZ, J. (1999), “Notas a propósito de la ejemplificación y la sinonimia en los diccionarios para extranjeros”, in Vila, M. N. & al. (eds.) (1999), 77-95.
- HAAN, P. (1986), “A Corpus Investigation into the Behaviour of Prepositional Verbs”, in Aarts, J. & Meijs, W. (eds.) (1986), 121-135.
- HABERT, B. & al. (1997) [B. HABERT, A. NAZARENKO, A. SALEM], *Les linguistiques de corpus*, Paris: Armand Colin.
- HAENSCH, G. (1982a), “Tipología de las obras lexicográficas”, in Haensch, G. & al. (1982), 95-187.
- HAENSCH, G. (1982b), “Aspectos prácticos de la elaboración de diccionarios”, in Haensch, G. & al. (1982), 395-534.
- HAENSCH, G. (1996), “Misèria i esplendor dels diccionaris”, in Cabré, M. T. (dir.) (1996), 35-46.
- HAENSCH, G. & al. (1982) [G. HAENSCH, L. WOLF, S. ETTINGER, R. WERNER], *La lexicografia*, Gredos, Madrid.
- HANKS, P. (1987), “Definitions and Explanations”, in Sinclair, J. (ed.) (1987), 116-136.
- HARTMANN, R. R. K. (ed.) (1983), *Lexicography: Principles and Practice*, London: Academic Press.
- HAUSMANN, F. J. & al. (1989) [F. J. HAUSMANN, O. REICHMANN, H. E. WIEGAND, L. ZGUSTA], *Wörterbücher. Ein internationales Handbuch zur Lexicographie*, I, Berlin: Walter de Gruyter.
- HEID, U. (1992), “Monolingual, Bilingual, «Interlingual» Description. Some Remarks on a New Method for the Production of Bilingual Dictionaries”, in *Euralex'90. Proceedings*, Barcelona: Biblograf, 167-184.
- HEID, U. (1996), “Using Lexical Functions for the Extraction of Collocations from Dictionaries and Corpora”, in Wanner, L. (ed.) (1996), 115-146.
- HERBST, T. (1986), “Defining with a Controlled Defining Vocabulary in Foreign Learners' Dictionaries”, *Lexicographica*, 2, 101-119.

- HERBST, T. (1996), "On the Way to the Perfect Learners' Dictionary: a first Comparison of OALD5, LDOCE3, COBUILD2 and CIDE", *International Journal of Lexicography*, 9, 4, 321-357.
- HERBST, T. & POPP, K. (eds.) (1999), *The Perfect Learner's Dictionary (?)*, Tübingen: Max Niemeyer.
- HERNÁNDEZ, H. (1992), "Los diccionarios de uso del último decenio (1980-1990): estudio crítico", in *Euralex '90. Proceedings*, Barcelona: Biblograf, 474-481.
- HERNÁNDEZ, H. (1994a), "Los diccionarios enciclopédicos", in Hernández, H. (coord.) (1994), 61-70.
- HERNÁNDEZ, H. (1994b), "El diccionario entre la semántica y las necesidades de los usuarios", in Hernández, H. (coord.) (1994), 107-118.
- HERNÁNDEZ, H. (2000), "Últimas tendencias en la lexicografía didáctica", in Ruhstaller, S. & Prado, J. (eds.) (2000), 101-115.
- HERNÁNDEZ, H. (coord.) (1994), *Aspectos de lexicografía contemporánea*, Barcelona: Biblograf.
- HERRERO, J. L. (2000), "El método lexicográfico de Gaston Gross", in Ruhstaller, S. & Prado, J. (eds.) (2000), 457-472.
- HJELMSLEV, L. (1928), *Principes de grammaire générale*, Kobenhavn: Munksgaard.
- HJELMSLEV, L. (1943), "Omkring Sprogteoriens Grundlæggelse", *Festskrift udg. Kobenhavns Universitet*, 1-113. [Versió espanyola: Díaz de Liaño, J. L. (trad.), *Prolegómenos a una teoría del lenguaje*, Madrid: Gredos, 1969].
- HOCKETT, C. F. (1958), *A Course in Modern Linguistics*, New York: Macmillan.
- HOCKEY, S. & IDE, N. (eds.) (1994), *Research in Humanities Computing 2*, Oxford: Oxford University Press.
- HOUSEHOLDER, F. W. & SAPORTA, S. (eds.) (1967), *Problems in Lexicography*, Bloomington: Indiana University.
- HURK, I. & MEIJS, W. (1986), "The dictionary as corpus: analyzing LDOCE's definition language", in Aarts, J. & Meijs, W. (eds.) (1986), 99-125.
- IBRAHIM, A. H. (ed.) (1989), *Lexiques*, Paris: Hachette.
- IDE, N. (1998), "Corpus Encoding Standard: SGML Guidelines for Encoding Linguistic Corpora", in *Proceedings of the First International Conference on Language Resources and Evaluation*, European Language Resources Association, Granada, 463-473.

IDE, N. & VÉRONIS, J. (1994), "Refining Taxonomies Extracted from Machine Readable Dictionaries", in Hockey, S. & Ide, N. (eds.) (1994), 145-170.

IDE, N. & VÉRONIS, J. (1996), *Corpus Encoding Standard. Document CES 1*. Document electrònic, <http://ilc.pi.cnr.it/EAGLES/>.

ILSON, R. (1986), *Lexicography. An Emerging International Profession*, Manchester: Manchester University Press.

ILSON, R. (1987), "Towards a Taxonomy of Dictionary Definitions", in Ilson, R. (ed.) (1987), 61-73.

ILSON, R. (ed.) (1985), *Dictionaries, Lexicography and Language Learning*, Oxford: Pergamon Press.

ILSON, R. (ed.) (1987), *A Spectrum of Lexicography*, Amsterdam: John Benjamins.

IMBS, P. (1960), "Au seuil de la lexicographie", *Cahiers de lexicologie*, 2, 3-17.

IVIR, V. (1988), "Collocations in Dictionaries. Monolingual and Bilingual", in Burton, T. L. & Burton, J. (eds.) (1988), 43-50.

JACKENDOFF, R. (1985), *Semantics and Cognition*, Cambridge: MIT Press.

JACKSON, H. (1988), *Words and their Meaning*, London: Longman.

JACKSON, H. (1990), *Grammar and Meaning: A semantic Approach to English Grammar*, Longman: London - New York.

JANSEN, J. & al. (1987) [J. JANSEN, J. P. MERGEAL, J. VANANDROYE], "Controlling LDOCE's controlled vocabulary", in Cowie, A. (ed.) (1987), 78-94.

JIMÉNEZ, E. (2000), "Aspectos de lexicografía y traducción automática", in Ruhstaller, S. & Prado, J. (eds.) (2000), 485-496.

JUILLAND, A & CHANG-RODRIGUEZ, E. (1964), *Frequency Dictionary of Spanish Words*, The Hague: Mouton.

KATZ, J. J. (1972), *Semantic Theory*, New York: Harper and Row.

KATZ, J. J. & FODOR, J. A. (1963), "The Structure of a Semantic Theory", *Language*, 39, 2, 170-210.

KEMPSON, R. M. (1977), *Semantic Theory*, Cambridge: Cambridge University Press.

KILGARRIFF, A. (1998), "The Hard Parts of Lexicography", *International Journal of Lexicography*, 11, 1, 51-54.

- KIRKPATRICK, B. (1985), "A Lexicographical Dilemma: Monolingual Dictionaries for the Native Speaker and for the Learner", in Ilson, R. (ed.) (1985), 7-13.
- KLEIBER, G. (1990), *La sémantique du prototype*, Paris: Presses Universitaires de France.
- KLEIBER, G. (1999), *Problèmes de sémantique. La polysémie en question*, Paris: Presses Universitaires du Septentrion.
- KRISHNAMURTHY, R. (1997), "Keeping Good Company: Collocations, Corpus, and Dictionaries", in Cabré, M. T. (dir.) (1997), 31-56.
- KROMAN, H. & al. (1984) [H. KROMAN, T. RIIBER, P. ROSBACH], "Active and passive bilingual dictionaries", in *LEXeter'83 Proceedings. Papers from the International Conference on Lexicography*, Tübingen, Max Niemeyer, 207-215.
- LABROPOULOU, P. & al. (2000) [P. GAVRILIDOU, P. LABROPOULOU, E. MANTZARI, H. PAPAGEORGIOU], "Automatic Generation of Dictionary Definitions from a Computational Lexicon", in *LREC 2000. Second International Conference on Language Resources and Evaluation. Proceedings*, Athens: European Language Resources Association, 247-253.
- LANDAU, S. I. (1984), *Dictionaries. The Art and Craft of Lexicography*, Cambridge: Cambridge University Press.
- LANGACKER, R. W. (1984), "Active Zones", *Proceedings of the Annual Meeting of the Berkeley Linguistic Society*, 10, 172-188.
- LARA, L. F. (1992), "El discurso del diccionario", in Wotjak, G. (ed.) (1992), 1-12.
- LARA, L. F. (1997), *Teoría del diccionario monolingüe*, México: Colegio de México.
- LEECH, G. (1971), *Semantics*, London: Penguin.
- LEES, R. B. (1960), *The Grammar of English Nominalizations*, The Hague: Mouton.
- LEHMANN, A. M. (1989), "Les représentations idéologiques dans le discours du dictionnaire", in Ibrahim, A. H. (ed.) (1989), 106-112.
- LEHMANN, A. M. (1990), "De définition à définition. L'interprétation dans le dictionnaire par le jeu des renvois", in Chaurand, J. & Mazière, F. (eds.) (1990), 208-224..
- LEVIN, B. (1991), "Building a Lexicon: The Contribution of Linguistics", in Boguraev, B. (ed.) (1991), 167-204.

LEVIN, B. & PINKER, S. (eds.) (1992), *Lexical and Conceptual Semantics*, Oxford: Blackwell.

LEWANDOWSKA, B. (1988), "Universal Concepts and Language Specific Meaning", in *ZüriLEX'86 Proceedings*, Tübingen: Francke, 17-26.

LEWICKA, H. & BOGACKI, K. (dirs.) (1983), *Dictionnaire sémantique et syntaxique des verbes français*, Warszawa: Państwowe Wydawnictwo Naukowe.

LORENTE, M. (1994), "Aspectes de lexicografia: representació i interpretació gramatical", Tesi Doctoral, Barcelona: Universitat de Barcelona.

LORENTE, M. (1999), "Els diccionaris de referència. L'aportació de la lexicografia actual a la norma", in Cano, M. A. & al. (eds.) (1999), 117-133.

LUQUE, J. DE D. & MANJÓN, F. J. (1997), "Los diccionarios explicativo-combinatorios de la Escuela Semántica de Moscú-Montreal, in Luque, J. de D. & Pamies, A. (eds.) (1997), 55-71.

LUQUE, J. DE D. & PAMIES, A. (eds.) (1997), *Problemas de lexicología i lexicografía*, Granada: Universidad de Granada.

LYONS, J. (1968), *Introduction to Theoretical Linguistics*, Cambridge: Cambridge University Press.

LYONS, J. (1977), *Semantics*, Cambridge: Cambridge University Press.

LYONS, J. (1981a), *Language, Meaning and Context*, London: Fontana/Collins.

LYONS, J. (1981b), *Language and Linguistics*, Cambridge: Cambridge University Press.

LYONS, J. (1995), *Linguistic Semantics*, Cambridge: Cambridge University Press.

MALKIEL, Y. (1967), "A Typological Classification of Dictionaries on the Basis of Distinctive Features". in Householder, F. & Saporta, S. (eds.) (1967), 3-24.

MARCUS, S. (1970), "Définitions logiques et définitions lexicographiques", *Langages*, 19, 87-91.

MARÍN, R. (1999), "L'anàlisi aspectual: alguns assoliments", *Llengua i Literatura. Revista anual de la Societat Catalana de Llengua i Literatura*, 10, 91-124.

MARTÍ, M. A. & al. (1991) [A. AGENO, S. CARDOCE, I. CASTELLÓN, M. A. MARTÍ, G. RIGAU, H. RODRÍGUEZ, M. TAULÉ, F. VERDEJO], "Análisis de definiciones del diccionario VOX", *Procesamiento del lenguaje natural*, 11, 139-154.

- MARTÍ, M. A. & al. (1992) [A. AGENO, I. CASTELLÓN, M. A. MARTÍ, F. RIBAS, G. RIGAU, H. RODRÍGUEZ, M. TAULÉ, F. VERDEJO], “Adquisición de conocimiento léxico: Acquilex”, *Procesamiento del lenguaje natural*, 12, 43-50.
- MARTÍ, M. A. (coord.) (2001), *Les tecnologies del llenguatge*, Barcelona: Universitat Oberta de Catalunya.
- MARTIN, R. (1972), “Esquisse d’une analyse formelle de la polysémie”, *Travaux de Linguistique et de Littérature*, X, 1, 125-136.
- MARTIN, R. (1974), “Syntaxe de la définition lexicographique: étude quantitative des définissants dans le *Dictionnaire fondamental de la langue française*”, in David, J. & Martin, R. (eds.), *Statistique et linguistique*, Paris: Klincksieck.
- MARTIN, R. (1977a), “Essai d’une typologie des définitions dans le dictionnaire de la langue”, *Travaux de linguistique et de littérature*, XV, 1, 362-378.
- MARTIN, R. (1977b), “Reflexions sur la structure logique du dictionnaire”, in *XIV Congresso Internazionale di Lingüistica e Filologia Romanza. Napoli, 15-20 Aprile 1974. Atti*, IV, Napoli - Amsterdam, 57-69.
- MARTIN, R. (1979), “La polysémie verbale. Esquisse d’une typologie formelle”, *Travaux de Linguistique et de Littérature*, XVII, 1, 251-261.
- MARTIN, R. (1983), *Pour une logique du sens*, Paris: Presses Universitaires de France.
- MARTIN, R. (1989), “L’exemple lexicographique dans le dictionnaire monolingue”, in Hausmann, F. J. & al. (1989), 599-607.
- MARTÍN MINGORANCE, L. (1994a), “La lexicografía onomasiológica”, in Hernández, H. (coord.) (1994), 15-27.
- MARTÍN MINGORANCE, L. (1994b), “La lexicografía de valencias”, in Hernández, H. (coord.) (1994), 29-37.
- MARTINET, A. (1960), *Éléments de linguistique générale*, Paris: Armand Colin.
- MARTÍNEZ DE SOUSA, J. (1995), *Diccionario de lexicografía práctica*, Barcelona: Biblograf.
- MASCARÓ, J. (1986), *Morfología*, Barcelona: Enciclopèdia Catalana.
- MATTHEWS, P. H. (1974), *Morphology. An Introduction to the Theory of Word-Structure*, Cambridge: Cambridge University Press.
- MAZIÈRE, F. (1989), “Une analyse de la définition: formes, historicité et idéologie”, in Ibrahim, A. H. (ed.), 97-105.

- MCENERY, T. & WILSON, A (1996), *Corpus Linguistics*, Edinburgh: Edinburgh University Press.
- MEDEROS, H. (1994), “A propósito de la definición lexicográfica”, in Hernández, H. (coord.) (1994), 95-106.
- MEIJS, W. (ed.) (1987), *Corpus Linguistics and Beyond*, Amsterdam: Rodopi.
- MEIJS, W. (1992), “The expanding lexical universe: extracting taxonomies from machine-readable dictionaries”, in *Euralex'90. Proceedings*, Bibliograf, Barcelona, 1992, 119-128.
- MEL'ČUK, I. (1984), “Un nouveau type de dictionnaire: Le Dictionnaire explicatif et combinatoire du français contemporain”, in Mel'čuk & al. (1984), 3-16.
- MEL'ČUK, I. (1988a), “Semantic Description of Lexical Units in an Explanatory Combinatorial Dictionary: Basic Principles and Heuristic Criteria”, *International Journal of Lexicography*, 1, 3, 165-188.
- MEL'ČUK, I. (1988b), “Paraphrase et lexique dans la théorie linguistique Sens-Text”, *Lexique*, 6, 13-54.
- MEL'ČUK, I. (1993), *Cours de morphologie générale. Introduction et première partie: le mot*, Montréal: Université de Montréal.
- MEL'ČUK, I. (1996), “Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon”, in Wanner, L. (ed.) (1996), 37-102.
- MEL'ČUK, I. (1998), “Collocations and Lexical Functions”, in Cowie, A. (1998), 23-53.
- MEL'ČUK, I. & ZHOLKOVSKY, A. (1988), “The explanatory combinatorial dictionary”, in Evens, M. (ed.) (1988), 41-74.
- MEL'ČUK, I. & al. (1984) [I. MEL'ČUK, N. ARBATCHEWSKY-JUMARIE, L. ELNITSKY, L. IORDANSKAJA, A. LESSARD.], *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques I*, Montréal: Presses de l'Université de Montréal.
- MEL'ČUK, I. & al. (1988) [I. MEL'ČUK, N. ARBATCHEWSKY-JUMARIE, L. DAGENAI, L. ELNITSKY, L. IORDANSKAJA, M.-N. LEFEBVRE, S. MANTHA.], *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques II*, Montréal: Presses de l'Université de Montréal.
- MEL'ČUK, I. & al. (1992) [I. MEL'ČUK, N. ARBATCHEWSKY-JUMARIE, L. IORDANSKAJA, S. MANTHA] *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques III*, Montréal: Presses de l'Université de Montréal.

- MEL'ČUK, I. & al. (1995) [I. MEL'ČUK, A. CLAS, A. POLGUÈRE], *Introduction à la lexicologie explicative et combinatoire*, Duculot: Louvain-la-Neuve.
- MEL'ČUK, I. & al. (1999) [I. MEL'ČUK, N. ARBATCHEWSKY-JUMARIE, L. IORDANSKAJA, S. MANTHA, A. POLGUÈRE], *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques IV*, Montréal: Presses de l'Université de Montréal.
- METTINGER, A. (1994), *Aspects of Semantic Opposition in English*, Oxford: Clarendon Press.
- MICHIELS, A. & NOËL, J. (1984), "The Pro's and Con's of a Controlled Defining Vocabulary in a Learner's Dictionary", in *LEXeter'83 Proceedings*, Tübingen: Max Niemeyer, 385-394.
- MIGUEL, E. (1999), "El aspecto léxico", in Bosque, I. & Demonte, V. (dirs.), 2977-3056.
- MILLÁN, J. A. (1999), "Estaciones filológicas", in Blecua, J. M. & al. (eds.) (1999), 143-164.
- MILLER, G. A. (1998a), "Introduction", in Fellbaum, C. (ed.) (1998), 1-19.
- MILLER, G. A. (1998b), "Nouns in WordNet", in Fellbaum, C. (ed.) (1998), 23-46.
- MILLER, G. A. (1990), "Nouns in WordNet: A Lexical Inheritance System", *International Journal of Lexicography*, 3, 4, 245-264.
- MILLER, G. A. & al. (1990), "Introduction to WordNet: An On-line Lexical Database", *International Journal of Lexicography*, 3, 4, 235-244.
- MILLER, G. A. & LEACOCK, C. (2000), "Lexical Representations for Sentence Processing", in Ravin, Y. & Leacock, C. (eds.) (2000), 152-160.
- MILLER, K. J. (1998), *Modifiers in WordNet*, in Fellbaum, C. (ed.) (1998), 47-67.
- MINSKY, M. (ed.) (1968), *Semantic Information Processing*, Cambridge: MIT Press.
- MOON, R. (1987), "The Analysis of Meaning", in Sinclair, J. (ed.) (1987), 86-103.
- MORALES, C. (2000), "La lexicografía del siglo XXI", in Ruhstaller, S. & Prado, J. (eds.) (2000), 505-514.
- MORKOVKIN, V. V. (1992), "Fundamentos teóricos de la lexicografía docente contemporánea", in *Euralex'90. Proceedings*, Barcelona: Biblograf, 359-368.
- MULLER, C. (1965), "Fréquence, dispersion et usage", *Cahiers de Lexicologie*, 7, 33-42.

- MULLER, C. (1968), *Initiation à la statistique linguistique*, Paris: Larousse.
- MULLER, C. (1977), *Principes et méthodes de statistique lexicale*, Paris: Hachette.
- MÜLLER, D. (1992), “¿Pertenece la metonimia al diccionario?”, in Wotjak, G. (ed.) (1992), 78-89.
- NAKHIMOVSKY, A. (1990), “Word Meaning and Syntactic Structure”, in Steele, J. (ed.) (1990), 3-17.
- NAKAMOTO, K. (1998), “From which Perspective does the Definer Define the Definendum: Anthropocentric or Referent-Based?”, *International Journal of Lexicography*, 11, 3, 205-218.
- NEUBAUER, F. (1987), “How to Define a Defining Vocabulary”, in Ilson, R. (ed.) (1987), 49-59.
- NEUBAUER, F. (1989), “Vocabulary Control in the Definitions and Examples of Monolingual Dictionaries”, in Hausmann, F. J. & al. (1989), 899-905.
- NIDA, E. A. (1997), “The Molecular Level of Lexical Semantics”, *International Journal of Lexicography*, 10, 4, 265-274.
- OOI, V. B. Y. (1998), *Computer Corpus Lexicography*, Edinburgh: Edindurgh University Press.
- PALMER, F. (1976), *Semantics*, Cambridge: Cambridge University Press.
- PASCUAL, J. A. (1996), “La coherencia en los diccionarios de uso”, in Forgas, E. (coord.) (1996), 167-198.
- PASCUAL, J. A. (1997), “La caracterización de los arcaísmos en un diccionario de uso”, in Cabré, M. T. (dir.) (1997), 9-30.
- PICOCHÉ, J. (1977), *Précis de lexicologie française*, Paris: Nathan.
- PICOCHÉ, J. (1998), “Les champs actanciels et leur intérêt pédagogique”, in Wotjak, G. (coord.) (1998), 183-208.
- PINEDA, M. A. (2000), “Lexicografía computacional y análisis morfológico automático”, in Ruhstaller, S. & Prado, J. (eds.) (2000), 515-525.
- PIOTROWSKI, T. (1988), “Defining Natural-kind Words”, in *ZüriLEX'86 Proceedings*, Tübingen: Francke, 55-62.
- POLGUÈRE, A. (2001), *Notions de base en lexicologie*, Montréal: Université de Montréal.

- PORTO DAPENA, J. A. (2002), *Manual de técnica lexicográfica*, Madrid: Arco.
- POTTIER, B. (1964), “Vers une sémantique moderne”, *Travaux de Linguistique et de Littérature*, II, 1, 107-137.
- POTTIER, B. (1965), “La définition sémantique dans les dictionnaires”, *Travaux de Linguistique et de Littérature*, III, 1, 33-39.
- POTTIER, B. (1974), *Linguistique générale. Théorie et description*, Paris: Klincksieck.
- POTTIER, B. (1978), “Organisation sémantique de l'article de dictionnaire”, *Bulletin de la Société de Linguistique de Paris*, 68, 1, 339-366.
- POTTS, T. C. (1994), *Structures and Categories for the Representation of Meaning*, Cambridge: Cambridge University Press.
- PUSTEJOVSKY, J. (1992), “The Syntax of Event Structure”, in Levin, B. & Pinker, S. (eds.) (1992), 47-81.
- PUSTEJOVSKY, J. (1995a), *The Generative Lexicon*, Cambridge: MIT Press.
- PUSTEJOVSKY, J. (1995b), “Linguistic Constraints on Type Coercion”, in Saint-Dizier, P. & Viegas, E. (eds.) (1995), 71-97.
- PUSTEJOVSKY, J. (2000), “Lexical Shadowing and Argument Closure”, in Ravin, Y. & Leacock, C. (eds.) (2000), 68-90.
- PUSTEJOVSKY, J. (2001), “Generativity and Explanation in Semantics. A Reply to Fodor and Lepore”, in Bouillon, P. & Busa, F. (eds.) (2001), 51-74.
- PUSTEJOVSKY, J. & BOGURAEV, B. (eds.) (1996), *Lexical Semantics. The Problem of Polysemy*, Oxford: Clarendon Press.
- QUEMADA, B. (1968), *Les dictionnaires du français moderne, 1539-1863*, Paris: Didier.
- QUEMADA, B. (1990), “La nouvelle lexicographie”, in Cabré, M. T. (ed.) (1990), 55-78.
- QUILLIAN, M. R. (1968), *Semantic Memory*, in Minsky, M. (ed.) (1968), 227-270.
- RAFEL, J. (1988), “Sobre la normalització dels diccionaris catalans: el tractament de les sèries sinonímiques”, in *Estudis de llengua i literatura catalanes*, XVI, 57-112.
- RAFEL, J. (1989), “Sobre la normalització dels diccionaris catalans: els elements intrínsecs i extrínsecs en les definicions lexicogràfiques”, in *La Corona d'Aragó i les llengües romàniques. Miscel·lània d'homenatge per a Germà Colon*, Tübingen: Gunter Narr.

- RAFEL, J. (1994): “Un corpus general de referència de la llengua catalana”, *Caplletra*, 17, 219-250.
- RAFEL, J. (1996), “Introducció”, in Rafel, J. (dir.), *Diccionari de freqüències. 1 Llengua No literària*, Barcelona: Institut d’Estudis Catalans, VII-LIX.
- RAFEL, J. (1998), “Le réel et le virtuel dans le lexique. Le rôle des corpus textuels”, in *Atti del XXI Congresso Internazionale di Lingüistica e Filologia Romanza. Sezione 3, Lessicologia e semantica delle lingue romanze*, Tübingen: Max Niemeyer, 729-740.
- RAFEL, J. (2002), “Lexicografia”, in Climent, S. (coord.), *Lexicografia i terminologia*, Barcelona: Universitat Oberta de Catalunya.
- RAFEL, J. & SOLER, J. (2001), “El processament de corpus. La lingüística empírica”, in Martí, M. A. (coord.) (2001), p. 27-59.
- RAVIN, Y. & LEACOCK, C. (2000), “Polysemy: an Overview”, in Ravin, Y. & Leacock, C. (eds.) (2000), 1-29.
- RAVIN, Y. & LEACOCK, C. (eds.) (2000), *Polysemy. Theoretical and Computational Approaches*, Oxford: Oxford University Press.
- REINHEIMER-RÎPEANU, S. (1974), *Les dérivés parasynthétiques dans les langues romanes*, Mouton: The Hague-Paris.
- REY, A. (1968), “Les bases théoriques de la description lexicographique du français: tendances actuelles”, *Travaux de Linguistique et de Littérature*, VI, 1, 55-72.
- REY, A. (1970a), *La lexicologie. Lectures*, Paris: Klincksieck.
- REY, A. (1970b), “Typologie génétique des dictionnaires”, *Langages*, 19, 48-68.
- REY, A. (1977), *Le lexique, images et modèles. Du dictionnaire à la lexicologie*, Paris: Armand Colin.
- REY, A. (2000), “Defining Definition”, in Sager, J. C. (ed.) (2000), 1-14.
- REY-DEBOVE, J. (1967), “La définition lexicographique, bases d’une typologie formelle”, *Travaux de Linguistique et de Littérature*, V, 1, 141-159.
- REY-DEBOVE, J. (1969), “Les relations entre le signe et la chose dans le discours métalinguistique: être, s’appeler, désigner et se dire”, *Travaux de Linguistique et de Littérature*, VII, 1, 113-129.
- REY-DEBOVE, J. (1971), *Étude linguistique et sémiotique des dictionnaires français contemporains*, The Hague: Mouton.

- REY-DEBOVE, J. (1978), *Le métalanguage, étude linguistique du discours sur le language*, Paris: Robert.
- RIERA, C. & VALLÈS, J. (1991), *Un epistolari fonamental per a la lexicografia científica catalana (1928-1953)*, Barcelona: Institut d'Estudis Catalans.
- RIZO, A. J. (1992), "A Proposal for a Valency Lexicon of English Catenative Verbs", in *Euralex '90. Proceedings*, Barcelona: Biblograf, 381-390.
- ROBINSON, R. (1954), *Definition*, Oxford: Clarendon Press.
- RODRÍGUEZ, A. (2000), "El régimen preposicional en los diccionarios", in Ruhstaller, S. & Prado, J. (eds.) (2000), 485-496.
- ROJO, G. (1993), "La base de datos sintácticos del español actual", *Español actual*, 59, 15-20.
- ROJO, G. (1997), "Gramática y diccionario", in *IV Jornadas de metodología y didáctica de la lengua española: sintaxis*, Cáceres: Universidad de Extremadura.
- RUHSTALLER, S. & PRADO, J. (eds.) (2000), *Tendencias en la investigación lexicográfica del español. El diccionario como objeto de estudio lingüístico y didáctico*, Huelva: Universidad de Huelva.
- RUIMY, N. & al. (2001) [N. RUIMY, E. GOLA, M. MONACHINI], "Lexicography Informs Lexical Semantics", in Bouillon, P. & Busa, F. (eds.) (2001), 350-362.
- RUIZ GURILLO, L. (1998) "Una clasificación no discreta de las unidades fraseológicas", in Wojtak, G. (ed.) (1998), 15-37.
- RUNDELL, M. (1999), "Dictionary Use in Production", *International Journal of Lexicography*, 12-1, 35-53.
- SACRISTÁN, M. (1964), *Introducción a la lógica y al análisis formal*, Barcelona: Ariel.
- SAGER, J. C. (ed.) (2000), *Essays on Definition*, Amsterdam: John Benjamins.
- SAINT-DIZIER, P. & VIEGAS, E. (eds.) (1995), *Computational Lexical Semantics*, Cambridge: Cambridge University Press.
- SÁNCHEZ, A. & al. (1991) [A. SÁNCHEZ, E. TORREJÓN, M. A. DE PINEDA], "Informatización de diccionarios convencionales: un sistema de consulta para el 'Diccionario ideológico de la lengua española' de J. Casares", *Procesamiento del Lenguaje Natural*, 11, 101-125.
- SARFATI, G. (1995), *Dire, agir, définir. Dictionnaires et langage ordinaire*, Paris: L'Harmattan.

- SCALISE, S. (1984), *Generative Morphology*, Dordrecht: Foris.
- SCHACK RASMUSSEN, L. (1992), “Los aspectos metalexigráficos de un diccionario semántico. Reflexiones en torno a un diccionario semántico de los verbos españoles”, in Wotjak, G. (ed.) (1992), 108-124.
- SCHOLFIELD, P. (1999), “Dictionary Use in Reception”, *International Journal of Lexicography*, 12-1, 13-34.
- SCHROTEN, J. (1992), “Estructura argumental y polisemia”, in *Euralex'90. Proceedings*, Barcelona: Bibliograf, 233-243.
- SECO, M. (1978), “Problemas formales de la definición lexicográfica”, in *Estudios ofrecidos a Emilio Alarcos Llorach, II*, Oviedo, 217-39 [reproduït a Seco, M. (1987), 15-34].
- SECO, M. (1979), “El ‘contorno’ en la definición lexicográfica”, in *Homenaje a Samuel Gili Gaya (in memoriam)*, Barcelona, 183-191 [reproduït a Seco, M. (1987), 35-45].
- SECO, M. (1987) *Estudios de lexicografía española*, Madrid: Paraninfo.
- SECO, M. (1988), “El problema de la diacronía en los diccionarios generales”, *Revista de Dialectología y Tradiciones Populares*, xliii, 559-567 [reproduït a Seco, M. (2003), 70-80].
- SECO, M. (2000), “Los pilares de un diccionario moderno”, *Saber / Leer*, 4-5, [reproduït a Seco, M. (2003), 81-90].
- SECO M. (1997), “El diccionario sincrónico del español”, in Cabré, M. T. (dir.) (1997), 133-149.
- SECO, M. (2002), “¿Para quién hacemos los diccionarios?”, in *Pulchre, bene, recte. Estudios en homenaje al Prof. Fernando González Ollé*, Pamplona: Universidad de Navarra, 1333-1347 [reproduït a Seco, M. (2003), 91-105].
- SECO, M. (2003) *Estudios de lexicografía española*, (2a. ed. augmentada), Madrid: Gredos.
- SINCLAIR, J. (1991), *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.
- SINCLAIR, J. (1996a), “The Search for Units of Meaning”, in Cabré, M. T. (dir.) (1996), 97-107.
- SINCLAIR, J. (1996b), *EAGLES Preliminary Recommendations on Corpus Typology*, EAGLES Document EAG-TCWG-CTYP/P (<http://ilc.pi.cnr.it/EAGLES>).

- SINCLAIR, J. (1996c), "An International Project in Multilingual Lexicography", *International Journal of Lexicography*, 9, 3, 179-196.
- SINCLAIR, J. (ed.) (1987), *Looking Up. An account of the COBUILD Project in lexical computing and the development of the Collins COBUILD English Language Dictionary*, Collins: London - Glasgow.
- SINCLAIR, J. & BALL, J. (1996), *EAGLES Preliminary Recommendations on Text Typology*, EAGLES Document EAG-TCWG-TTYP/P (<http://ilc.pi.cnr.it/EAGLES>).
- SINGLETON, D. (2000), *Language and the Lexicon*, London: Arnold.
- SLAGTER, P. J. (2000), "Diccionaris bilingües y monolingües para extranjeros. Apuntes de un usuario", in Cabré, M. T. & Gelpí, C. (eds.) (2000), 31-72.
- SMITH, N. & WILSON, D. (1979), *Modern Linguistics. The Results of Chomsky's Revolution*, London: Penguin.
- SOL, L. M. (1992a), "El DGLC i la formalització lexicogràfica dels verbs", Tesi de llicenciatura, Barcelona: Universitat de Barcelona.
- SOL, L. M. (1992b), "El caràcter aspectual i la definició lexicogràfica", *Anuari de Filologia*, XV, Secció C, 3, 75-93.
- SOLÀ, J. (1994), "La lexicografia catalana i la normativa", *Caplletra*, 17, 273-290.
- SOLÀ, J. (ed.) (1992), *Aspectes de la lexicografia catalana moderna*, Barcelona: Empúries.
- SOLER, J. (1992a), "Circularitat lexicogràfica i primitius semàntics", *Anuari de Filologia*, XV, Secció C, 3, 75-93.
- SOLER, J. (1992b), "Jerarquies lèxiques i descripció lexicogràfica", Tesi de llicenciatura, Barcelona: Universitat de Barcelona.
- SOLER, J. (2000), "Producing LRs in Parallel with Lexicographic Description: the DCC Project", in *LREC 2000. Second International Conference on Language Resources and Evaluation. Proceedings*, Athens: European Language Resources Association, 341-346.
- SOLER, J. (2001), "Elements extrínsecs i definicions adjectivals", in Alsina, V. & DeCesaris, J. (eds.) (2001), 49-67.
- SPERBERG-MCQUEEN, C. M. & BURNARD, L., (eds.) (1994), *Guidelines for Electronic Text Encoding and Interchange*, Oxford: Text Encoding Initiative.
- STEELE, J. (ed.) (1990), *Meaning-Text Theory. Linguistics, Lexicography, and Implications*, Ottawa: University of Ottawa Press.

- STEINBERG, D. & JACKOBITS, L. A. (eds.) (1971), *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, Cambridge: Cambridge University Press.
- STERKENBURG, P. (1989), "Towards a Basic Vocabulary of Dutch", *Lexicographica*, 5, 159-167.
- STEVENSON, M. & WILKS, Y. (2000), "Large Vocabulary Word Sense Disambiguation", in Ravin, Y. & Leacock, C. (eds.) (2000), 161-177.
- STOCK, P. (1988), "The Structure and Function of Definitions", in *ZüriLEX'86 Proceedings*, Tübingen: Francke, 81-89.
- SVENSÉN, B. (1993), *Practical Lexicography. Principles and Methods of Dictionary-Making*, Oxford: Oxford University Press.
- SWANEPOEL, P. H. (1992), "Linguistic Motivation and its Lexicographical Application", in *Euralex '90. Proceedings*, Barcelona: Biblograf, 291-314.
- SZYMANEK, B. (1988), *Categories and Categorization in Morphology*, Lublin: Redackja Wydawnictw Katolickiego Uniwersytetu Lubelskiego.
- SZYMANEK, B. (1989), *Introduction to Morphological Analysis*, Warsaw: Panstwowe Wydawnictwo Naukowe.
- TAZAWA, K. (1998), "Proposta per a un diccionari japonès-català per a catalanoparlants", Tesi doctoral, Barcelona: Universitat de Barcelona.
- TELIJA, V. N. (1992), "Lexicographic Description of Words and Collocations: Feature-Functional Model", in *Euralex '90. Proceedings*, Barcelona: Biblograf, 315-320.
- TOMASZCZYK, J. & LEWANDOWSKA, B. (1988), "Universal Concepts and Language Specific Meaning", in *ZüriLEX '86 Proceedings*, Tübingen: Francke, 17-26.
- TRUJILLO, R. (1994), "El diccionario frente a la semántica", in Hernández, H. (coord.) (1994), 73-93.
- ULLMANN, S. (1962), *Semantics*, Oxford: Blackwell. [Versió espanyola: *Semántica. Introducción a la ciencia del significado*, Madrid: Aguilar, 1967.]
- VENDLER, Z. (1967), *Linguistics in Philosophy*, Ithaca: Cornell University Press.
- VENY, J. (1996), "Gènesi, mètode, estructura i valoració del *Diccionari català-valencià-balear*", in Cabré, M. T. (dir.) (1996), 23-33.
- VIDAL BENEYTO, J. (ed.) (1991), *Las industrias de la lengua*, Madrid: Fundación Germán Sánchez Rupérez.

- VIEGAS, E. (ed.) (1999), *Breadth and Depth of Semantic Lexicons*, Dordrecht: Kluwer.
- VILA, M. N. & al. (eds.) (1999) [M. A. CALERO, R. M. MATEU, M. CASANOVAS, J. L. ORDUÑA, M. N. VILA], *Así son los diccionarios*, Lleida: Universitat de Lleida.
- VOSSEN, P. & al. (1989) [P. VOSSEN, W. MEIJS, M. BROEDER], "Meaning and structure in dictionary definitions", in Boguraev, B. & Briscoe, T. (eds.) (1989), 171-190.
- VOSSEN, P. (ed.) (1998), *EuroWordNet. A Multilingual Database with Lexical Semantic Networks*, Kluwer: Dordrecht.
- WANNER, L. (ed.) (1996), *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam: John Benjamins.
- WEINREICH, U. (1962), "Lexicographic Definition in Descriptive Semantics", *International Journal of American Linguistics*, 28-2. [Reproduït a: Householder, F. & Saporta, S. (eds.) (1967), 25-44.]
- WEINREICH, U. (1963), "Lexicology", in Sebeok, T. A. (ed.), *Current Trends in Linguistics*, vol. I, The Hague: Mouton, 60-93.
- WEINREICH, U. (1964), "Webster's Third: A critique of its Semantics", *International Journal of American Linguistics*, 30, 405-409.
- WEINREICH, U. (1966), "Explorations in Semantic Theory", in Sebeok, T. A. (ed.), *Current Trends in Linguistics, III*, The Hague - Paris: Mouton, 395-477.
- WERNER, R. (1982a), "Léxico y teoría general del lenguaje", in Haensch, G. & al. (1982), 21-94.
- WERNER, R. (1982b), "La definición lexicográfica", in Haensch, G. & al. (1982), 259-328.
- WEST, M. (1953), *A General Service List of English Words with Semantic Frequencies*, London: Longmans.
- WHEELER, M. W. (1977), "Problemes lèxics per al generativista: els mots catalans a doble arrel", in *Actes del IV Col·loqui internacional de llengua i literatura catalanes (Basilea)*, Barcelona: Abadia de Montserrat, 235-250.
- WIEGAND, H. E. (1999a), *Semantics and Lexicography*, Tübingen: Max Niemeyer.
- WIEGAND, H. E. (1999b), "A New Theory of the So-called Lexicographic Definition", in Wiegand, H. E. (1999a), 153-201.
- WIEGAND, H. E. (1999c), "Elements of a Theory towards a So-called Lexicographic Definition", in Wiegand, H. E. (1999a), 203-281.

- WIERZBICKA, A. (1972), *Semantic Primitives*, Frankfurt: Athenaeum.
- WIERZBICKA, A. (1985), *Lexicography and Conceptual Analysis*, Ann Arbor: Karoma.
- WIERZBICKA, A. (1987), *English Speech Act Verbs*, Sidney: Academic Press.
- WIERZBICKA, A. (1996), *Semantics. Primes and Universals*, Oxford: Oxford University Press.
- WIERZBICKA, A. (2000), “Semantics, Emotions, and the Meaning of the Gospels”, in Cabré, M. T. & Gelpí, C. (eds.) (2000), 103-121.
- WILKS, Y. & al. (1989) [D. FASS, C. GUO, J. E. McDONALD, T. PLATE, B. M. SLATOR, Y. WILKS], “A tractable machine dictionary as a resource for computational semantics”, in Boguraev, B. & Briscoe, T. (eds.) (1989), 193-228.
- WILKS, Y. (2001), “The ‘fodor’-FODOR Fallacy Bites Back”, in Bouillon, P. & Busa, F. (eds.) (2001), 75-85.
- WOTJAK, G. (1971), *Untersuchungen zur Struktur der Bedeutung. Ein Beitrag zu Gegenstand und Methode der modernen Bedeutungsforschung unter Besonderer Berücksichtigung der semantischen Konstituentenanalyse*, Berlin: Akademie.
- WOTJAK, G. (1992), “Estructuras en el léxico”, in Wotjak, G. (ed.) (1992), 38-62.
- WOTJAK, G. (1998), “¿Cómo tratar las unidades fraseológicas (UF) en el diccionario?”, in Wotjak, G. (ed.) (1998), 307-321.
- WOTJAK, G. (ed.) (1992), *Estudios de lexicografía y metalexicografía del español actual*, Tübingen: Max Niemeyer.
- WOTJAK, G. (ed.) (1998), *Estudios de fraseología y fraseografía del español actual*, Frankfurt am Main: Vervuert Iberoamericana.
- WOTJAK, G. (coord.) (1998), *Teoría del campo y semántica léxica*, Frankfurt am Main: Peter Lang.
- ZAMPOLLI, A. (1991a), “Los bancos de datos léxicos: Bases multifuncionales de datos léxicos”, in Vidal Beneyto, J. (ed.) (1991), 127-146.
- ZAMPOLLI, A. (1991b), “Hacia bases multifuncionales de datos léxicos”, in Vidal Beneyto, J. (ed.) (1991), 185-202.
- ZGUSTA, L. (1971), *Manual of Lexicography*, The Hague: Mouton.
- ZGUSTA, L. (1989), “The Role of Dictionaries in the Genesis and Development of the Standard”, in Hausmann, F. J. & al. (1989), 70-79.

Sigles i referències de diccionaris

BBI = BENSON, M. & al. [M. BENSON, E. BENSON, R. ILSON], *The BBI Combinatory Dictionary of English, A Guide to Word Combinations*, Amsterdam: John Benjamins, 1986.

CIDE = *Cambridge International Dictionary of English*, London: Cambridge University Press, 1995.

COBUILD = *Collins COBUILD English Language Dictionary*, London [etc.]: Collins, 1988.

DCVB = ALCOVER, A. M. & MOLL, F. de B., *Diccionari català-valencià-balear*, Palma de Mallorca: Editorial Moll, 1988, 10 v.

DEA = SECO, M. & al. [O. ANDRÉS, G. RAMOS, M. SECO], *Diccionario abreviado del español actual*, Madrid: Aguilar, 2000.

DGLC = FABRA, P., *Diccionari general de la llengua catalana*, Barcelona: Llibreria Catalònia, 1932.

DFDCC = RAFEL, J. (dir.), *Diccionari de freqüències*, Barcelona: Institut d'Estudis Catalans, 1998, 3 v.

DIDAC = *Didac: diccionari del català*, Barcelona: Enciclopèdia Catalana, 1995.

DIEC = INSTITUT D'ESTUDIS CATALANS, *Diccionari de la llengua catalana*, Barcelona [etc.]: Edicions 62 [etc.], 1995.

DLC = *Diccionari de la llengua catalana*, 2a ed. Barcelona: Enciclopèdia Catalana, 1985.

DMLC = INSTITUT D'ESTUDIS CATALANS, *Diccionari manual de la llengua catalana*, Barcelona: Institut d'Estudis Catalans, 2000.

DOP = BRUGUERA, J., *Diccionari ortogràfic i de pronúncia*, Barcelona: Enciclopèdia catalana, 1990.

Bibliografia

- DPR = ROBERT, P., *Le Petit Robert 1: dictionnaire alphabétique et analogique de la langue française*, Rédaction dirigée par A. Rey et J. Rey-Debove. Nouvelle éd. revue, corrigée et mise à jour. Paris: Le Robert, 1990.
- DRAE = REAL ACADEMIA ESPAÑOLA, *Diccionario de la lengua española*, 21^a ed. Madrid: Espasa Calpe, 1992, 2 v.
- DSLE = *Diccionario Salamanca de la lengua española*, Dir. Juan Gutiérrez Cuadrado. Madrid: Santillana; Salamanca: Universidad de Salamanca; Barcelona: Círculo de Lectores, 1996.
- DUE = MOLINER, M., *Diccionario de uso del español*, Madrid: Gredos, 1990.
- GD62 = LÓPEZ DEL CASTILLO, L. (dir.), *Gran diccionari 62 de la llengua catalana*, Barcelona: Edicions 62, 2000.
- GDLC = *Gran diccionari de la llengua catalana*, Barcelona: Enciclopèdia Catalana, 1998.
- LDOCE = LONGMAN: *Dictionary of Contemporary English*, 3rd ed. Essex: Longman, 1995.
- NMED = WEST, M. P. & ENDICOTT, J. G., *The New Method English Dictionary*, London: Longmans, 1935.
- OALD = HORNBY, A. S., *Oxford advanced learner's dictionary of current english*, 4th ed. Oxford: Oxford University Press, 1989.
- Primer Diccionari = *Primer diccionari*, Vic: Eumo, 1998.
- VLI = ZINGARELLI, N., *Lo Zingarelli: vocabolario della lingua italiana*, 12a ed. Bologna: Zanichelli, 1997.

Altres sigles utilitzades en la redacció del treball

ACQUILEX = Projecte patrocinat per la Comissió Europea (ESPRIT BRA-3030), desenvolupat entre 1989 i 1992). Tingué com a objectiu fonamental el processament de la informació de diccionaris existents per a l'extracció d'informació de cara a la constitució de bases de dades lèxiques i de bases de coneixement lèxic.

CTILC = Corpus Textual Informatitzat de la Llengua Catalana, constituït en el marc del projecte DCC. És un corpus de referència del català de 52 milions de paraules, i actualment s'usa per a les tasques de redacció del DDLC.

DCC = Diccionari del Català Contemporani, programa de recerca que es desenvolupa des de 1985 a l'Institut d'Estudis Catalans (IEC). En la seva fase de constitució de recursos lingüístics va constituir el CTILC. Actualment el DCC està en fase de redacció del DDLC.

DDLC = Diccionari descriptiu de la llengua catalana, en curs d'elaboració en el marc del programa DCC de l'IEC.

DEC = Diccionari explicatiu i combinatori, tipus de diccionari desenvolupat en el marc teòric de la MTT.

EAGLES = Projecte patrocinat per la Comissió Europea (programa LRE1-6110), desenvolupat entre 1993 i 1998. Tingué com a finalitats principals el desenvolupament d'estàndards en l'àmbit del disseny, la creació i l'avaluació dels recursos lingüístics en format electrònic, a través de la reunió de diversos grups d'experts.

EFL = *English as a Foreign Language*, denominació genèrica que s'aplica als diccionaris (com el LDOCE, el COBUILD, el CIDE, l'OALD,...) per a aprenents estrangers de l'anglès.

EuroWordNet = Projecte patrocinat per la Comissió Europea, desenvolupat entre 1996 i 1998. Tingué com a finalitat principal la constitució de lèxics electrònics (basats en el model de WordNet) per a diverses llengües europees, de cara a l'explotació multilingüe.

FL = Funció lèxica: fórmula amb què la MTT descriu les diferents relacions paradigmàtiques i sintagmàtiques que contreen els lexemes.

GENELEX = Projecte patrocinat per la Comissió Europea (programa EUREKA E!-524) desenvolupat entre 1990 i 1994). Tingué com a finalitat principal el

desenvolupament d'un model de lèxic electrònic apte per a diverses aplicacions funcionals.

MTT = *Meaning Text Theory*, formulada a Moscou l'any 1965, per Alexander Zholkovsky i Igor Mel'čuk; més tard s'hi afegí Juri Apresjan. La teoria s'orienta bàsicament a una lingüística basada en la descripció del lèxic, i els seus esforços es dirigeixen principalment a la producció d'obres lexicogràfiques denominades genèricament DEC.

MULTILEX = Projecte patrocinat per la Comissió Europea (programa ESPRIT 5304) desenvolupat entre 1990 i 1993. Tingué com a finalitat principal l'elaboració d'eines i l'establiment d'estàndards per a la constitució de lèxics electrònics.

PAROLE = Projecte patrocinat per la Comissió Europea (programa LE2-4017) desenvolupat entre 1996 i 1998. Tingué com a finalitat principal la constitució de corpus i lèxics electrònics (basats en el model de GENELEX) per a 12 llengües europees.

SGML = Sistema genèric de codificació (Standard Generalized Mark-up Language); és un esquema de marcatge autodeclaratiu definit per la norma ISO 8879. Fou adoptat pel consorci TEI com a base per al desenvolupament dels seus estàndards de codificació textual. Posteriorment, aquests estàndards s'han adaptat a XML.

SIMPLE = Projecte patrocinat per la Comissió Europea (programa LE4-8346) desenvolupat entre 1998 i 2000. Fou la continuació de PAROLE, i tingué com a finalitat l'afegiment d'informació semàntica als lèxics electrònics constituïts en el projecte PAROLE.

TEI = *Text Encoding Initiative*, consorci internacional que té com a objectiu la provisió d'estàndards de codificació per a diferents tipus de textos, amb la finalitat de facilitar-ne l'intercanvi i el desenvolupament d'eines de tractament.

WordNet = Projecte de constitució d'un lèxic electrònic que es desenvolupa a la Universitat de Princeton d'ençà de 1985. Ha originat un model computacional *de facto* per a lèxics orientats a les relacions semàntiques que ha estat la gènesi dels lèxics per a diverses llengües europees constituïts pel projecte EuroWordNet.

XML = Sistema genèric de codificació (eXtended Markup Language) desenvolupat a partir de SGML, però específicament orientat a la codificació d'estructures de dades.