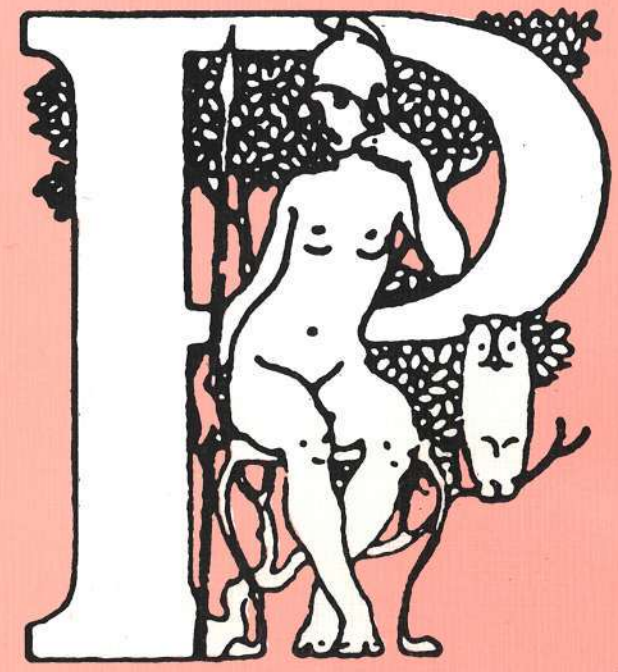
# CAPLLETRA 17

REVISTA INTERNACIONAL DE FILOLOGIA



*Tardor 1994*

**Redacció:**
Institut Universitari de Filologia Valenciana
Avda. Blasco Ibáñez, 28
46010 València
Telèfon 96-386 42 55
**Edita i publica:**
Publicacions de l'Abadia de Montserrat
Institut de Filologia Valenciana
**Subscripcions i correspondència:**
Ausiàs Marc, 92-98 – 08013 Barcelona
**Disseny:**
Manuel Granell
Joan Josep Tornero

**ISSN:**
0214 - 8188
**Dipòsit legal:**
B - 25.992 - 1990
**Fotocomposició:**
Servei de Publicacions
Universitat de València
**Imprès per:**
Novagràfik, S.A. - Barcelona
(febrer 1995)

La publicació d'aquest volum ha estat possible gràcies a la col·laboració del Departament de Filologia Catalana de la Universitat de València i a un ajut a revistes d'investigació de la Universitat de València.

*Caplletra no es responsabilitza del contingut de les col·laboracions que s'hi publiquen.*

Volum monogràfic sobre
***Lexicografia***

coordinat per
*M. Teresa Cabré*

*Tardor, 1994*

**Director:**
Antoni Ferrando

**Consell de Redacció:**
Consell de l'Institut Universitari de Filologia Valenciana

**Comitè Científic:**
E. Casanova, A. Ferrando, A. G. Hauf, M. Pérez Saldanya,
V. Salvador, V. Simbor

**Consell Assessor:**

Antoni Badia i Margarit
Jordi Castellanos
Germà Colón
Joseph Gulsoy
Joan Solà
Giuseppe Sansone
Arthur Terry
Joan Veny
Curt Wittlin
M. Claire Zimmermann

# NORMES D'EDICIÓ DE LA REVISTA *CAPLLETRA*

Els articles originals, els exemplars de recensió, les comandes de subscripció i la correspondència en general s'han de trametre a la redacció de la revista (Institut Universitari de Filologia Valenciana, Av. Blasco Ibáñez, 28, 46010, València).

Els articles originals s'enviaran mecanoscrits a dos espais, preferentment en format DINA4. Si és possible, a més de la versió mecanoscrita s'inclourà el text informatitzat, preferiblement en els formats d'ordinadors Macintosh o en WordPerfect 5.0.

Els articles es presentaran preferentment en català i els autors n'inclouran un resum de 8/10 ratlles en català o anglés.

Les notes crítiques apareixeran a peu de pàgina i es reservaran a explicacions o aclariments complementaris. Les referències bibliogràfiques s'inclouran dins del cos de l'article i seguiran el sistema AUTOR-ANY-PÀGINA. L'any i les pàgines de la citació o referència apareixeran sempre entre parèntesis, separats per dos punts. Només quan la referència abaste tota una obra s'hi ometrà explicitar les pàgines a què es remet. Entre l'últim cognom de l'autor citat i l'any de publicació del llibre o article a què es fa referència no s'intercalarà cap signe :

... com ja s'ha observat (Sanchis Guarner 1980: 144)
... com observa Sanchis Guarner ( 1980: 144 )
... "Dels documents del Principat desaparegué plenament l'article *es* ja en el segle XIV" (Sanchis Guarner 1980: 144)
... Aquesta era la idea recurrent del llibre (Sanchis Guarner 1980)

Totes les referències que apareixen en el text es repetiran al final del treball dins d'un apèndix de REFERÈNCIES BIBLIOGRÀFIQUES. Hi aniran ordenades alfabèticament i s'ajustaran a les convencions de la bibliografia científica internacional :

- articles ( títol entre cometes i nom de la revista en cursiva, numeració completa, lloc i any d'edició, pàgines):
  PICALLO, C. (1990) "Elements anafòrics i localitat", *Caplletra* 8, València, IUFV, pp. 41-53.
- llibres (títol en cursiva, lloc d'edició i casa editorial):
  SANCHIS GUARNER, M. (1980) *Aproximació a la història de la llengua catalana*, Barcelona, Salvat.
- capítol o parts de llibres col·lectius o miscel·lanis (autor, títol, dins, primer autor, nom de l'editor o editors de l'aplec, ed. o eds., títol del recull, any, pàgines):
  ANDERLE, A. (1991) "La crisi dels col·lectivismes" dins San Martín, A. ed., *Públic / Privat: un debat obert*, València, Universitat de València/Ajuntament de Gandia.

Els noms dels autors hi figuraran en versaletes, primer el cognom o cognoms; separat per una coma, el nom de fonts abreviat amb la lletra inicial.

Les contribucions acceptades i compostes es retornaran als autors per a la correcció de proves d'impremta. Les esmenes s'efectuaran seguint els signes convencionals del codi habitual de correctors.

INTRODUCCIÓ. M. Teresa Cabré ..... 9

**PART I: LEXICOGRAFIA GENERAL I METODOLOGIA LEXICOGRÀFICA** ..... 13

LEXICOGRAFIA I MODELS LINGÜÍSTICS: LES TEORIES LINGÜÍSTIQUES I EL LÈXIC. Toni Badia ..... 15

LEXICOGRAFIA I VARIACIÓ FUNCIONAL. Lluís Payrató ..... 47

LEXICOGRAFIA I CULTURA: ANTROPOLOGIA I SISTEMA SOCIAL(NOMS, RENOMS I NOMS DE CASA AL PIRINEU ARAGONÈS). Joan J. Pujadas/Dolors Comas ..... 59

LEXICOGRAFIA I INTEL·LIGÈNCIA ARTIFICIAL. Ramon Cerdà ..... 81

L'ÚS DELS DICCIONARIS ESCOLARS. M. Paz Battaner ..... 109

LA PROFESSIÓ DE LEXICÒGRAF. Emili Pascual ..... 125

**PART II: LEXICOGRAFIA CATALANA** ..... 137

LA LEXICOGRAFIA CATALANA ACTUAL. M. Teresa Cabré ..... 139

UN TAST DELS VOCABULARIS BILINGÜES VALENCIANS ACTUALS. Jordi Colomina ..... 157

LA LEXICOGRAFIA CATALANA MODERNA A LES ILLES BALEARS. Jaume Corbera ..... 179

LA LEXICOGRAFIA A L'ALGUER. Jaume Corbera ..... 195

EL NOU DICCIONARI NORMATIU. M. Teresa Cabré, Maria Bozzo, Àngels Egea, Marta Juncadella, Carolina Santamaria ..... 201

EL CORPUS TEXTUAL DE LA LLENGUA CATALANA. Joaquim Rafel ..... 219

LA TERMINOLOGIA CATALANA ACTUAL. Carles Duarte ..... 251

LA LEXICOGRAFIA BILINGÜE. Jordi Bruguera ..... 267

LA LEXICOGRAFIA GENERAL I LA NORMATIVA. Joan Solà ..... 273

LEXICOGRAFIA CATALANA I CORRECCIÓ: LA FUNCIÓ DELS DICCIONARIS EN LA REVISIÓ DE TEXTOS. Mercè Lorente/Carles Tebé ..... 291

UNA TIPOLOGIA DE LA DEFINICIÓ LEXICOGRÀFICA APLICADA AL DGLC. Victòria Alsina/Lluís Sol/Joan Soler ..... 311

**PART III: ENTREVISTES** ..... 325

ENTREVISTA A ALAIN REY, DIRECTOR DELS DICCIONARIS LE ROBERT ..... 327

ENTREVISTA A JOHN SINCLAIR, EDITOR EN CAP DEL COLLINS COBUILD ENGLISH LANGUAGE DICTIONARY ..... 331

ENTREVISTA A DANIELLE CORBIN I PIERRE CORBIN, DIRECTORS DEL DIPLOMA EUROPEU DE LEXICOGRAFIA. Universitat de Lille ..... 337

# INTRODUCCIÓ

La lexicografia és avui dia una matèria plenament renovada. Ho diuen els lingüistes, ho diuen els editors de diccionaris, ho diuen fins i tot els lexicògrafs tradicionals. Fer un diccionari significa avui dia reaprendre les regles d'un joc antic, i això tant des del punt de vista dels supòsits con del de les tecnologies.

La vella polèmica sobre una sèrie d'activitats derivades de l'aplicació lingüística, de si es tracta de teories o de pràctiques, de si són una disciplina o un art, de si requereixen un estudi de base o, simplement, una pràctica continuada, sembla finalment força superada. I s'ha superat per la via de considerar que aquestes matèries que, com la lexicografia, tenen (com qualsevol altra matèria, diria jo) un vessant de base teòrica i unes possibilitats aplicades, són matèries interdisciplinàries, amb una base teòrica inevitable (els seus fonaments lingüístics i psicològics), amb un marc disciplinari establert (la intersecció de les ciències del llenguatge i d'altres disciplines o vessants de disciplines) i amb unes aplicacions destinades a resoldre necessitats de tipus expressiu o comunicatiu, monolingües o plurilingües, en situacions naturals de comunicació (la comunicació humana) o artificials (la comunicació persona-màquina o màquina-màquina).

No hi ha dubte que la lexicografia serveix inevitablement a totes aquestes necessitats, perquè la construcció d'un diccionari no té cap mena de valor en ella mateixa si no és concebuda en relació a una necessitat comunicativa. La recopilació de mots, formants o frases fetes d'una llengua, a més de donar a l'especialista una idea de com funciona aquesta llengua en el seu vessant lèxic, pot reflectir també una manera d'entendre el món, o una manera de simular com els parlants tenen els coneixements lèxics emmagatzemats a la ment, serveixen també per fixar unes determinades formes lèxiques de referència o per delimitar els vocabulari que un programa informàtic és capaç de manipular, entendre i generar. Vet aquí la multifuncionalitat dels diccionaris, i vet aquí la multicaracterització de les paraules.

Però en la renovació de la lexicografia (o millor dit, en la seva reconceptualització) hi han pesat també, al costat de l'evolució del pensament lingüístic, l'aparició de les tecnologies de la comunicació i la informació. Fer avui dia un diccionari requereix una racionalització tecnològica que permeti obtenir molts més resultats que no s'obtenien abans d'un fitxer de paperetes, per ben concebudes i sistemàtiques que fossin. La possibilitat d'extreure dels textos les paraules més usuals, la possibilitat de definir-les a partir dels seus usos (i no inspirant-se en altres diccionaris), la possibilitat d'il·lustrar el seu funcionament amb exemples reals, de delimitar la seva categorització a partir de materials fiables, de controlar les remissions, de verificar que totes les unitats que s'usen en un diccionari són alhora entrades d'aquest diccionari; en suma i encara més enllà, la possibilitat que d'uns materials ben constituïts, ben organitzats, ben il·lustrats, se'n puguin extreure molts diccionaris diferents, dóna a l'activitat lexicogràfica una volada impressionant.

La llengua catalana participa sense cap mena de dubte d'aquesta evolució. Fins i tot diria que és als primers llocs en uns determinats aspectes. El volum que aquí presentem pretén donar una visió de conjunt de la lexicografia catalana avui, sense perdre de vista el marc general en què s'insereix.

Amb aquest propòsit hem dividit aquesta obra en tres parts. La primera part presenta algunes qüestions generals sobre els diccionaris, la seva teoria (T. Badia), les relacions que s'estableixen entre els diccionaris i la varietat funcional del llenguatge (Ll. Payrató), entre els diccionaris i els valors culturals (D. Comas i J. Pujadas), entre els diccionaris i l'aprenentatge (M. P. Battaner), entre els diccionaris i la intel·ligència artificial (R. Cerdà). Aquesta primera part es tanca amb una reflexió d'E. Pascual sobre la professió mateixa, feta des d'una llarga experiència de lexicògraf.

La segona part és una recopilació de treballs sobre la lexicografia catalana, que s'inicia amb una panoràmica general (M. T. Cabré), amb una presentació de conjunt de la lexicografia balear (J. Corbera), valenciana (J. Colomina) i algueresa (J. Corbera); i amb una síntesi de l'estat actual de la lexicografia bilingüe (J. Bruguera), i de la terminologia, entesa com a lexicografia especialitzada (C. Duarte). S'inclou a continuació la presentació de dos projectes d'importància rellevant per a la llengua catalana: el *Nou Diccionari Normatiu de la Llengua Catalana* i el *Corpus Textual del Català Contemporani*, ambdós projectes de l'Institut d'Estudis Catalans. Per cloure aquest apartat,

es presenten articles sobre aspectes específics dels diccionaris, centrats en preocupacions i problemes reals: la connexió dels diccionaris i la normativa (J. Solà), la relació dels diccionaris amb la correcció (M. Lorente i C. Tebé), i alguns problemes a l'entorn de la definició lexicogràfica (V. Alsina, Ll. Sol i J. Soler).

La darrera part d'aquest monogràfic inclou tres entrevistes amb especialistes de primera línia sobre aspectes diferents de la lexicografia: la professió (A. Rey), la formació (D. Corbin i P. Corbin) i la necessitat de disposar de corpus per a l'elaboració de diccionaris (J. Sinclair).

En conjunt aquest volum dóna una idea general de la lexicografia que esperem que serà útil com a informació i com a reflexió en general, i que serà estimulant per tal com mostra que la lexicografia catalana desperta un interès important i té avui dia un grau de vitalitat envejable. Aquest ha estat el nostre propòsit des de l'inici de la concepció d'aquest volum de la revista *Caplletra*.*

M. Teresa Cabré Castellví
*Barcelona, gener de 1994*

* Vull agrair de cor a Rosa Estopà l'ajut que m'ha prestat en la confecció d'aquest volum. Sense la seva valuosa participació, aquest monogràfic hauria trigat molt més temps a aparèixer.

# Part I:

# LEXICOGRAFIA GENERAL I METODOLOGIA LEXICOGRÀFICA

TONI BADIA

# LEXICOGRAFIA I MODELS LINGÜÍSTICS: LES TEORIES LINGÜÍSTIQUES I EL LÈXIC

L'objectiu fonamental d'aquest treball és posar de manifest com la teoria lingüística, qualsevol teoria lingüística, comporta una visió particular sobre el lèxic i, per tant, proposa unes tasques lexicogràfiques determinades. Per a això, ens aturarem en alguns dels models teòrics de la lingüística d'aquest segle i n'observarem la visió del lèxic. Els models escollits no són, evidentment, els únics que ens haurien pogut servir: la proliferació de models lingüístics teòrics és una de les característiques de l'evolució d'aquesta ciència durant aquest segle. Ara bé, la seva successió representa prou adequadament els grans moments pels quals ha passat la lingüística en la seva història moderna. Així doncs, després d'uns mots breus sobre el que hem convingut d'anomenar gramàtica tradicional, començarem el nostre recorregut en l'estructuralisme, per passar després a la gran família de teories generativistes; primer observarem tres moments de l'evolució del pensament de Chomsky (la teoria prestàndard, la teoria estàndard i la teoria de la recció i el lligam), i després ens aturarem a dues altres teories generatives sorgides dels ambients més directament influïts per la lingüística computacional (la gramàtica lèxico-funcional i la gramàtica d'estructura sintagmàtica orientada al nucli).

*Caplletra 17* (Tardor 1994), pp. 15-46

## 1. INTRODUCCIÓ

Si acceptem que els plantejaments científics tenen lloc en la lingüística i que el seu objectiu fonamental és aconseguir una explicació del comportament lingüístic humà, haurem de convenir que allò que pertoca és una caracterització global d'aquest comportament. De fet, les teories lingüístiques tenen com a objectiu l'oferir aquesta descripció global, en què cada component troba el seu lloc en relació als altres. D'aquesta manera, en cada una d'elles, el lèxic ocupa el lloc que li correspon, al costat dels altres elements que la conformen. Per tant, no és estrany que en cada una el lèxic hi tingui un paper diferent: no es tracta d'una posició absoluta, sinó relativa al conjunt de la teoria, o del sistema que aquesta proposa. Així, un dels criteris usats en la classificació de les teories lingüístiques és el pes relatiu que atorguen al lèxic en el sistema global que proposen (Droste i Joseph, 1991:19).

No hi ha dubte que qualsevol proposta de sistema lingüístic ha de concedir un lloc al lèxic, al diccionari. Ara bé, aquest pot ser simplement el mínim necessari (com en les primeres versions de la Gramàtica generativa transformacional –vegeu, més avall, §6), o, en canvi, pot constituir una peça essencial en el sistema proposat (com en la Gramàtica lèxico-funcional –vegeu §8). Aquestes dues teories són probablement dos extrems en la dicotomia entre lèxic i gramàtica. Ara bé, serveixen per a mostrar d'una manera força gràfica com hi ha una relació inversa entre la quantitat de feina atorgada al lèxic i la que correspon a la resta de la gramàtica. Així, en els primers estadis de la gramàtica generativa les regles gramaticals ocupen molts dels espais que en la gramàtica lèxico-funcional ocupa el lèxic. A la vegada, també mostren (en dues versions distintes) com les propietats del lèxic són una conseqüència de les característiques de la teoria lingüística en general. D'aquesta manera, una teoria en la qual tots els elements sistemàtics són formulats en regles exteriors al lèxic no permet que aquest tingui cap altre paper que el d'un veritable calaix de sastre (on es recullen precisament els aspectes que no han pogut ser sotmesos a la sistematicitat). En canvi, una teoria en la qual es preveuen dos nivells de sistematicitat diferents, l'un exterior i l'altre interior al lèxic, està imposant unes condicions al lèxic quant a la seva estructura i al seu enllaç amb els altres elements de la gramàtica.

En el paràgraf anterior hem partit d'una visió teòrica del llenguatge, en la qual ocupa un lloc principal la distribució de la feina entre els diversos components de la teoria. És una característica general de les aproximacions teòriques al llenguatge de proposar un model explicatiu en el qual hi ha diverses estructures independents però interrelacionades (anomenades mòduls). En aquesta visió, el lèxic és un d'aquests mòduls, enfront dels altres components de la gramàtica. Ara bé, és només entrat aquest segle que comencen a sorgir visions teòriques del llenguatge en què es posa a la pràctica una estructuració així. Però, fins i tot ara, a les acaballes del segle XX, el que solem anomenar *gramàtica tradicional*, que té les seves arrels en una mala popularització

de la gramàtica clàssica i medieval, encara perdura en les nostres aules i publicacions. Així doncs, encara que probablement no és massa correcte, sembla procedent de fer una petita menció a la situació del lèxic en la gramàtica tradicional.

## 2. LA GRAMÀTICA TRADICIONAL

Entenem per gramàtica tradicional aquelles visions del llenguatge sense sistematitzar, en les quals es juxtaposen explicacions d'ordre diferent de manera gens rigorosa, barrejant, per exemple, descripcions de tipus semàntic amb característiques morfològiques. Com que, a més, hi manca una perspectiva global del fenomen lingüístic, aquests diferents elements no troben el seu lloc en un marc teòric general i són usats, simplement, en funció de la conveniència del redactor de la gramàtica. Per altra banda, la preocupació fonamental d'aquestes aproximacions gramaticals és determinar els usos lingüístics correctes (enfront dels barbarismes, localismes, vulgarismes...), de manera que es converteixen a la pràctica en manuals prescriptius del bon ús de la llengua. Com que gairebé només interessa establir una normativa, a la quals els parlants (de fet, els escrivents) d'una llengua s'haurien de sotmetre, en aquests manuals, els estudis lingüístics hi tenen una importància relativa.

Aquestes dues característiques que acabem de mencionar (la manca de perspectiva global i la finalitat únicament normativista) fan que molt sovint la gramàtica tradicional caigui en plantejaments absolutament casuístics: es tracta de caracteritzar els usos correctes dels mots i expressions i de distingir-los dels incorrectes, i de fer-ho de manera que pugui ser utilitzada per persones no entrenades especialment per a l'anàlisi lingüística. El resultat n'és un model gramatical no sistemàtic, centrat en els aspectes més fàcilment delimitables de la llengua.

No és estrany que en un model d'aquestes característiques el tema de la relació entre el lèxic i la gramàtica quedi molt difús. En primer lloc, perquè la mateixa noció de gramàtica no hi és gaire clara i, en segon lloc, perquè es fonamenta sempre en els aspectes més superficials de la llengua. Així, el centre d'atenció fonamental és la paraula. Des d'aquesta perspectiva, els aspectes que poden ser tractats més fàcilment són l'ortografia, la morfologia i la semàntica lèxica, que correspondrien, les primeres, a la forma de les paraules i, la segona, al seu significat. I aquesta és precisament la distribució de la feina entre les gramàtiques i els diccionaris: les primeres determinen les normes d'ortografia i les formes morfològiques de les paraules, i els segons fixen els seus significats. Com que les relacions entre les paraules en les emissions lingüístiques no són estudiades de manera sistemàtica, els pocs elements sintàctics o estructurals que s'hi tracten ho són en contextos de paraula. La paraula, de fet, queda convertida en la unitat d'anàlisi lingüística, de manera que els estudis al voltant del lèxic són els que ocupen la major part de les pàgines de les gramàtiques tradicionals. El punt d'enllaç

entre la informació proporcionada per les gramàtiques i els diccionaris és ben dèbil: la breu categorització dels mots en el diccionari segueix els criteris classificatoris proposats per les gramàtiques.

La tasca de la lexicografia, doncs, queda limitada a la recopilació de les paraules de la llengua, a la seva fixació i a la determinació dels seus usos (sentits) correctes, i a la seva categorització d'acord amb els pocs principis fixats per la gramàtica.

## 3. L'ESTRUCTURALISME

Des dels primers intents d'elaboració d'una teoria del llenguatge, apareix la necessitat de buscar una estructura que superi el marc estricte de les paraules. L'observació que hi ha aspectes de la llengua (tant sintàctics, com semàntics) que no poden ser explicats des d'una perspectiva lèxica únicament ha dut els lingüistes a plantejar la necessitat d'una estructura per sobre i més enllà de les paraules. La història de la lingüística en aquest segle és, de fet, la història de la sistematització de les estructures lingüístiques, amb la finalitat de poder descriure i predir els fets lingüístics amb prou objectivitat i amb prou caràcter explicatiu. En aquest procés, el primer estadi és el representat per l'estructuralisme. Amb ell, podríem dir que neix la lingüística com a ciència, ja que és quan, per primer cop, es disposa d'un aparell teòric rigorós per a començar a emprendre seriosament l'anàlisi lingüística.

És clau en el plantejament estructuralista la caracterització de les relacions que els diversos elements de la llengua poden contreure entre ells, les quals són agrupades en dos blocs, el de les relacions sintagmàtiques i el de les paradigmàtiques (Saussure, 1915). Com ja és prou conegut, les primeres són les que estableixen entre ells els elements en la cadena fònica, mentre que les segones són les que tenen amb altres elements del sistema (que podrien ocupar el seu lloc en la cadena fònica, que s'hi oposen, que s'hi complementen, etc.).

La caracterització d'aquestes relacions es fa a partir de dues operacions fonamentals: la segmentació i la classificació (Bierwisch, 1966). El procés de segmentació s'aplica a les cadenes fòniques de manera que s'hi descobreixen els elements que les conformen; com és obvi, els resultats d'aquest procés es poden trobar en nivells diferents de descripció lingüística (fonemes i síl·labes, morfemes, paraules, sintagmes, etc.). Aquests elements són després classificats segons les seves propietats formals i funcionals. En aquest procés de classificació hi juguen un paper principal les posicions relatives que ocupen els diversos elements en l'estructura abstracta formada per tots ells, en el sistema que és la llengua. Naturalment els criteris rellevants per a la classificació són determinats en gran part pel nivell de descripció que es pretén aconseguir: així, si hom pretén obtenir una descripció de tipus fonològic, es tindran en compte factors d'ordre fònic, mentre que per a una descripció morfosintàctica, s'hauran de tenir en compte sobretot factors de tipus sintàctic.

Així doncs, el valor relatiu d'un element en el sistema és el resultat del conjunt de relacions sintagmàtiques i paradigmàtiques que contreu. A més, la mateixa porció física de la cadena fònica pot ser subjecte de diferents tipus de relacions, segons el nivell de descripció que es tingui en compte. Sens dubte, des del punt de vista estructuralista, qualsevol fragment de la cadena fònica (simple o complex) pot ser un element subjecte a diversos tipus de relacions. Ara bé, també és cert que la paraula hi ocupa una posició central, ja que és a partir d'aquesta noció que es caracteritzen completament els conceptes més adequats de *morfema*, *lexema*, *sintagma*, etc. En gran part, és el tractament semàntic el que provoca aquest fet. Vegem-ho amb una mica de detall, ja que és a la base de la concepció estructuralista del diccionari.

El signe lingüístic té un vessant formal i un altre de semàntic, està format per elements físics (els fonemes) que representen elements no físics (els conceptes, en la visió estructuralista majoritària). Aquesta visió del signe lingüístic determina fortament el pes que tindran les paraules en la descripció general de la llengua. Així, trobem en el pla del contingut el mateix biaix que hi ha en el pla de l'expressió, ja que en ambdós es dóna preferència a la descripció dels elements i a la seva classificació: així com les relacions entre els fonemes en la cadena fònica no són explicades a fons (Bierwisch, 1966), el tractament semàntic estructuralista es redueix, gairebé sempre, a una semàntica de la paraula. Que aquest paral·lelisme va ser entès d'una manera quasi literal ho prova el fet que es va intentar aplicar de seguida a la semàntica de les paraules el mètode de la descomposició en trets essencials i universals que tant d'èxit havia tingut en fonologia. És cert que molt aviat apareix la divisió entre lexema i morfema, de manera que resulta possible de distingir, en una paraula, entre els factors pròpiament de contingut i els que determinen l'enllaç de les paraules entre elles. Però aquests elements són ja les unitats mínimes en què es pot descompondre la cadena fònica des del punt de vista del contingut; qualsevol pas posterior en l'anàlisi es fa per un procés de descomposició en trets que no és composicional en el sentit modern del terme.[1] El camí que hauria pogut obrir la porta de l'estudi de les relacions semàntiques horitzontals hauria estat l'estudi integrat de les relacions sintàctiques entre els elements de les cadenes fòniques. De totes maneres, des de la perspectiva estructuralista era pràcticament impossible arribar a la generalitat necessària per a aconseguir-ho, ja que es tractaven els diversos sistemes morfemàtics com a estructures aïllades les unes de les altres.

Fins i tot, la mateixa noció d'arbitrarietat del signe lingüístic sembla confirmar aquest pes important que tenen les paraules en l'anàlisi lingüística: si comparem l'estructura fonològica d'una paraula amb el seu contingut, la seva arbitrarietat salta a la vista (en la majoria dels casos, no hi ha res que relacioni l'una amb l'altre); en canvi, si el que comparem són cadenes fonològiques de més d'una paraula, l'explicació de l'arbitrarietat es fa molt més complexa, ja que cal fer referència a les diverses estructures que hi intervenen i a la manera com s'entrellacen les unes amb les altres. Des d'aquesta perspectiva, doncs, resulta també natural de centrar la descripció lingüística en la paraula.

(1) Entenem que la divisió d'un tot és composicional quan la interpretació del tot és una funció de la interpretació de les parts: és a dir, quan per a interpretar el tot n'hi ha prou amb conèixer la interpretació de les parts i la manera com aquestes es combinen per a formar el tot.

Una altra característica fonamental de l'estructuralisme, que en determina profundament la visió sobre el lèxic, és l'establiment de la relació entre la forma i el contingut del signe a un nivell absolutament superficial: són els diversos elements de la cadena fònica (tan segmentats com calgui) els que són portadors de significat, de manera que tots els elements de contingut de les expressions lingüístiques han de ser relacionats de manera directa amb elements de la cadena fònica (siguin morfemes o lexemes; de fet, l'única variació admesa és la del morfema zero). Des del punt de vista de la descripció gramatical, això comporta problemes pràctics en el tractament d'alguns fenòmens (Tusón, 1980), com els constituents discontinus, la sinonímia sintàctica (tal com apareix, per exemple, en oracions actives transitives i les seves corresponents passives), o l'ambigüitat sintàctica (com la que es mostra en expressions com *llapis de fusta de colors*). Des del punt de vista de l'organització de la gramàtica, però, té conseqüències extraordinàriament importants. En no haver de donar compte de realitats lingüístiques abstractes, la descripció lingüística es redueix a la caracterització dels elements en què es descompon la cadena fònica (elements que resulten sempre del procés de segmentació lineal que hem apuntat abans) i a l'establiment de les possibilitats de combinació dels uns amb els altres. De manera natural, doncs, els morfemes i els lexemes es converteixen en el centre de l'anàlisi estructuralista. Les relacions horitzontals (sintagmàtiques) que contreuen entre ells serveixen fonamentalment al lingüista per a descobrir les relacions verticals (paradigmàtiques) amb els altres elements del sistema; d'aquesta manera la classificació (taxonomia) dels elements es converteix en una activitat bàsica. No hi ha dubte que una perspectiva així potencia el paper del lèxic (del diccionari) i de la morfologia, en detriment del de la resta de la gramàtica: la màxima activitat descriptiva lingüística es concentra en la caracterització dels diversos elements lèxics (lexemes o morfemes).

Per una banda, els morfemes són caracteritzats en descriure el (sub)sistema del qual formen part, de manera que, per exemple, els morfemes de plural són descrits en el subsistema del nombre (o en els diversos subsistemes del nombre), i els morfemes passius són descrits en el subsistema de la veu verbal.[2] Aquesta caracterització, és clar, es fonamenta en propietats formals, i és només a través de la seva posició en el subsistema corresponent, que cada morfema aporta un significat determinat: així, els morfemes de plural en català aporten un significat diferent del que tenen en grec clàssic a causa que en català l'oposició és entre dos únics valors (singular i plural), mentre que en grec clàssic ho és entre tres valors (singular, dual i plural), de manera que el valor de nombre de l'expressió catalana *els rius que limiten Barcelona* hauria de ser expressat en grec clàssic en dual, i no en plural.

Per l'altra banda, els lexemes són caracteritzats per les relacions de tipus semàntic que contreuen entre ells, entre les quals podem destacar dos grans grups de relacions, que Lyons (1977) anomena *camps semàntics* i *relacions de sentit*. De fet, l'estudi d'aquestes relacions constituirà la tasca central de la semàntica estructuralista. Els

(2) Notem de passada que això ens mostra una limitació de l'aparell estructuralista d'anàlisi. La passivitat d'una oració no és, simplement, una característica del seu verb, sinó una propietat general de tota la construcció, en la qual juguen un paper, a més de la morfologia verbal, els règims preposicionals i/o de cas, les relacions semàntiques entre el verb i els seus complements, etc.

camps semàntics conformen els subsistemes en què s'agrupen els lexemes, de manera que diverses de les relacions semàntiques estudiades, en ser-los aplicades, permeten d'establir-ne l'estructura. Indubtablement, les relacions paradigmàtiques que es posen de manifest d'aquesta manera constitueixen el cos de l'anomenada semàntica lèxica, i gran part d'aquesta informació coincideix amb la continguda en els diccionaris que coneixem. Així, de la mateixa manera que els aspectes pròpiament sintàctics són tinguts menys en compte en la tradició estructuralista que els morfosintàctics, els aspectes de semàntica general (o oracional) gairebé no hi són presents.

Totes aquestes opcions tendeixen a afavorir el tractament de la paraula (del lexema) com a nucli de la descripció lingüística, en detriment dels aspectes més generals (tant sintàctics i semàntics, com morfològics). En certa manera, aquesta visió no posa en qüestió la pràctica lexicogràfica tradicional. Això és així perquè, en el fons, manté la posició central de la paraula i no elabora d'una manera prou sistemàtica els aspectes generals de la llengua. Podríem dir, doncs, que l'estructuralisme no proposa encara una visió del llenguatge en la qual tots els ingredients juguin el seu paper conjuntadament, de manera que no ofereix un model (en el sentit modern del terme) del llenguatge humà.

Per a nosaltres, una derivació interessant de la visió estructuralista la constitueixen les primeres versions de la gramàtica de dependència (Tesnière, 1959), ja que se n'ha derivat una de les pràctiques lexicogràfiques més vinculades a una visió teòrica del llenguatge, la dels diccionaris de valències. Tesnière titula el seu llibre *Éléments de syntaxe structurale* i el comença afirmant que l'objecte de la sintaxi estructural és la *phrase*, el sintagma, així com que la noció de sintagma és lògicament anterior a la de mot (Tesnière, 1959:11,25). A la vegada, descriu les relacions entre els diversos elements de les oracions a partir de la noció de dependència, que no ha ser identificada amb la relació de precedència en la cadena fònica; així, com assenyala Baez San José (1988), hi ha una certa similitud entre aquest plantejament i el de la primera gramàtica generativa, en el sentit que les dues pressuposen una realitat abstracta per sobre de la cadena fònica, que en mostra la veritable estructura.

Tot i aquest reconeixement d'una estructura abstracta, la sintaxi de Tesnière està estrictament centrada en la paraula; una simple fullejada del seu llibre permet d'adonar-se que els fets sintàctics són explicats al voltant de les paraules, de les categories majors, sense postular processos sintàctics generals i independents dels lexemes als quals afecten. Així ens trobem davant d'una sistematització de les relacions que els elements superficials de la cadena fònica contreuen entre ells, que revela una taxonomia més adequada que les motivades per criteris estrictament superficials, però que, en definitiva, encara està lluny de poder oferir explicacions generalitzadores de les diferents realitzacions lexemàtiques. En els diccionaris de valències inspirats en la teoria de dependències de Tesnière, hi podem trobar nombrosos exemples d'aquest fet.

Considerem les diverses entrades lèxiques corresponents a l'adjectiu alemany *alt*

('vell'), consignades en el diccionari de Sommerfeldt i Schreiber (1974). Hi trobem nou entrades diferents, la majoria de les quals tenen la mateixa realització sintàctica i, per tant, es diferencien només per la caracterització semàntica dels seus arguments. A més, hi apareix com un fenomen casual un fet important de l'alemany, com és que (molts d)els seus adjectius atributius poden tenir una realització predicativa si no són usats en sentit valoratiu. A la vegada, la relació existent entre l'adjectiu *alt* usat absolutament, com a (1), i amb especificadors, com a (2), queda totalment amagada, quan es tracta d'un fenomen habitual en les llengües germàniques pel qual s'admeten especificadors quantitatius per als adjectius que poden expressar una mesura en una escala graduada (altres adjectius que admeten aquest tipus d'especificador són, per exemple, *hoch* ('alt'), *lang* ('llarg') o *schwer* ('pesat')).

(1) a. Die Frau ist alt ('La dona és vella')
b. die alte Frau ('la dona vella')

(2) a. Die Frau ist fünfzig Jahre alt (lit.: 'La dona és vella cinquanta anys')
b. die fünfzig Jahre alte Frau (lit.: 'la dona vella de cinquanta anys')

L'existència d'entrades diferents i paral·leles per a una mateixa cadena fònica no suposa que tinguin cap relació determinada. Des d'un punt de vista teòric, es podria tractar de simples cadenes fòniques homònimes, de manera que el seu tractament en el diccionari hauria estat el mateix: aparèixer com a entrades distintes sota la mateixa etiqueta de cadena fònica.

Aquest exemple ens mostra el problema principal de la visió estructuralista quant al disseny d'un model gramatical: els fets descoberts tenen un límit en la possibilitat de ser sistematitzats, límit que és imposat per la impossibilitat d'expressar relacions abstractes entre diverses realitzacions oracionals o sintagmàtiques. Expressions complexes que els parlants d'una llengua tenim com a pròximes o relacionades (sigui com a sinònimes, sigui mitjançant altres tipus de relació –implicació, conseqüència, de contrari, de contradicció, etc.) no poden ser relacionades a causa de la manca d'abstracció de la cadena fònica. És només en les propostes generatives que aquesta limitació podrà ser superada. Passem, doncs, a observar el paper que s'ha atorgat al lèxic en les diverses teories generatives que s'han anat produint des de les seves primeres manifestacions per part de Chomsky els anys cinquanta.

## 4. EL GENERATIVISME

Els plantejaments generativistes són els primers que han permès el desenvolupament d'una visió del lèxic integrada dins la teoria lingüística general. Això ha estat així a causa dels plantejaments inicials del generativisme: no es tracta de descriure

els fenòmens lingüístics com fets donats, exteriors als parlants, sinó com a produccions lingüístiques d'aquests. En lloc de partir de les dades lingüístiques tal com es poden trobar recollides en els corpus lingüístics (en una visió fonamentalment estàtica del llenguatge), els generativistes parteixen de les expressions lingüístiques tal com són produïdes pels parlants. D'aquesta manera, els objectius bàsics de la lingüística canvien: en lloc de pretendre posar de manifest les relacions estructurals entre els diversos elements presents en els textos (o manifestacions lingüístiques en general), la lingüística pretén explicar la capacitat humana de fer aquestes manifestacions; es tracta, doncs, de produir un model teòric de com el parlant d'una llengua és capaç de produir i de comprendre textos. Conseqüentment, la posició relativa de la lingüística respecte d'altres ciències també canvia: si Saussure la veia com una branca d'una ciència dels signes, d'abast més general, que ell anomenava *semiologia*, Chomsky la veu com una branca de la ciència cognitiva.

Des d'aquesta nova perspectiva, es poden explicar també les relacions estructurals entre els elements lingüístics, però això es fa en un marc global més ampli. Certament, la pretensió dels generativistes és que l'explicació dels fenòmens lingüístics que en resulta és molt més adequada; segons les nocions establertes per Chomsky el generativisme no pretén aportar ja una mera descripció (elaborada fonamentalment a partir d'una taxonomia), sinó una veritable explicació dels fenòmens lingüístics en oferir una visió general del fet lingüístic humà. Ens trobem, doncs, amb el primer intent seriós d'oferir una visió global dels diversos factors que intervenen en l'activitat lingüística humana; conseqüentment, és en aquest marc on comencen a aparèixer visions diferents sobre com s'ha de relacionar el lèxic amb els altres components de la gramàtica.

Tot i que vist des de l'exterior apareix com un tot monolític, el corrent generativista, des de ben al començament de la seva història, ha estat molt ric en tendències i molt ràpid en la seva evolució. En general, la història de la gramàtica generativa ha estat molt lligada a la figura del seu màxim exponent, N. Chomsky, sigui perquè es tracta de models desenvolupats per ell (o per seguidors immediats seus), sigui perquè es proposen models alternatius als seus. Aquesta visió de la lingüística neix a partir dels primers treballs de Chomsky, en els quals incorpora els principis i la metodologia que havien contribuït amb èxit al desenvolupament dels llenguatges artificials durant els últims anys del segle XIX i la primera part d'aquest segle (treballs de Frege, Russell i Whitehead, Tarski i Carnap, entre molts altres). Com a conseqüència, es concep la capacitat lingüística com una teoria formal, amb els seus diversos components, que és capaç de generar (és a dir, de produir) la totalitat de les oracions gramaticals de la llengua en qüestió. És clar que en una aproximació així a la teoria lingüística, el lloc que hi ocuparà el lèxic resultarà fàcilment determinat, ja que tots els components hi tenen especificat el seu paper. Així doncs, des de la nostra perspectiva repassarem alguns d'aquests models i hi observarem el paper que correspon al lèxic. D'aquesta manera, podrem determinar les tasques que les diverses teories lingüístiques assignen al lèxic i, per tant, el paper que en cada cas hi pot tenir la lexicografia.

## 5. EL PRIMER MODEL GENERATIVISTA: EL PRESTÀNDARD

La primera versió àmpliament coneguda de la gramàtica generativa és la que va presentar Chomsky el 1957 al seu llibre *Syntactic Structures*. En ella, el lèxic no hi té un paper diferent del de la sintaxi, és a dir, no constitueix un component diferenciat de la gramàtica. La gramàtica (la sintaxi), doncs, es compon només de dos mòduls diferenciats: les regles d'estructura de frase i les regles transformacionals.[3] Així, tant les relacions de constituença com la introducció dels elements lèxics es produeixen per regles del mateix tipus, sintagmàtiques o de reescriptura de frase. Ho veiem en el conjunt de regles que li serveix d'exemple de l'anàlisi de constituents (Chomsky, 1957:26):[4]

| | | | | |
|---|---|---|---|---|
| (3) | (i) | Oració → SN + SV | (iv) | T → *el*, *la*, etc. |
| | (ii) | SN → T + N | (v) | N → *noi*, *nena*, *pilota*, etc. |
| | (iii) | SV → Verb + SN | (vi) | Verb → *agafa*, *xuta*, etc. |

En aquesta llista de regles no hi ha distinció (ni teòrica, ni pràctica) entre les de (3,i-iii) i les restants; ens trobem sempre amb regles que reescriuen un símbol per un altre (o uns altres): el símbol *SN* és reescrit per *T* i *N*, i *N* és reescrit per un dels símbols de la llista *noi*, *nena*, etc. És clar que, fins i tot en una teoria fortament sintàctica com aquesta, els elements lèxics hi han de tenir un lloc, ja que altrament ens trobaríem amb estructures vàlides (per a la llengua en qüestió), que podrien ser representades per diagrames vàlids (és a dir, generats per les regles corresponents), però que no podrien servir per a la caracterització de les seves oracions gramaticals; dit simplement, generaríem estructures, però no oracions; per tant, un llenguatge no podria ser definit per la llista de les seves oracions (és a dir, de les generades pel conjunt de regles en consideració), sinó per la llista de les seves estructures. Així doncs, la presència de regles com les de (3,iii-vi) és totalment justificada. El problema rau en què molt aviat apareixen problemes que difícilment poden ser resolts en aquest marc formal tan estret. El tractament mateix de les variacions morfològiques de les paraules hi pot ser resolt d'una manera relativament simple només perquè es tracta de donar compte de l'anglès, que és una llengua amb molt poques variacions flexives.[5] A la vegada, de seguida apareixen dificultats justament per la mala caracterització dels elements lèxics en aspectes sintàctics (i, fins i tot, semàntics): els verbs, per exemple, han de ser classificats i tractats de manera diferent segons els tipus sintàctics dels complements que poden tenir (transitius, intransitius, preposicionals, etc.), i també segons els trets lèxics dels seus complements (així, *la pedra beu* és inacceptable perquè el verb *beure* demana com a subjecte sintagmes nominals amb el tret lèxic *+animat*).

(3) En aquesta obra, es preveu la incorporació d'un mòdul addicional que donés compte de l'estructura fonològica de les oracions (això es fa a partir de la interpretació fonològica de l'estructura sintàctica resultant de les transformacions); a la vegada, no es nega la possibilitat d'incloure-hi un mòdul semàntic, interpretatiu també, tot i que s'argumenta a bastament que l'estructura sintàctica de les oracions és independent de la seva interpretació semàntica.

(4) En aquest cas, com en tots aquells en què ha estat possible, hem adaptat les propostes originals per a l'anglès, de manera que resultin adequades per al català.

(5) Per a adonar-se de la complexitat que un tractament d'aquest tipus comportaria per al català només cal intentar adaptar les regles de l'auxiliar verbal (Chomsky, 1957:39).

## 6. EL MODEL ESTÀNDARD

Fonamentalment per les raons que acabem d'adduir, en la segona versió de la gramàtica generativa posada a punt per Chomsky (1965), el lèxic entra a formar part, com un component més, de l'estructura general de la gramàtica. De fet és aquesta versió la que estableix els fonaments de la concepció generativa del llenguatge, ja que, al costat de la sintaxi, amb els seus dos nivells (d'estructura superficial i profunda), inclou també la fonologia i la semàntica, com a mòduls addicionals, interpretatius de les estructures sintàctiques superficials i profundes, respectivament.

La major part dels canvis que s'introdueixen a aquest model respecte al prestàndard es relacionen amb el lèxic i la semàntica. És natural que ben aviat es vagi considerar necessària la introducció de la semàntica en el model, ja que altrament es fa difícil de mantenir la idea que es tracta d'un model de la competència lingüística dels parlants, atès que hi ha fenòmens que formen part de la capacitat lingüística i que no poden ser explicats si no és des d'una perspectiva semàntica (per exemple, la sinonímia entre diverses expressions lingüístiques, moltes de les ambigüitats pròpies de les expressions, moltes de les restriccions de selecció, etc.). La semàntica, doncs, s'introdueix en la teoria lingüística perquè sense ella queden sense explicar aspectes importants del comportament lingüístic dels parlants; de totes maneres, es fa tot intentant que no envaeixi el camp de la gramàtica formalitzada, que es reserva a la sintaxi. Es pren aquesta actitud perquè, en el fons, hi ha el convenciment que el llenguatge natural en la seva totalitat no pot ésser formalitzat (almenys en l'estadi actual de desenvolupament de la lingüística). Com a conseqüència, Chomsky intenta introduir en la sintaxi tants elements com pot de la semàntica, i deixa per a aquesta tot allò que no pot ser rigorosament formalitzat.

En la visió de la teoria estàndard, el lèxic apareix com un component més del model lingüístic general. Ara bé, en el lèxic és on s'introdueixen la majoria de les característiques de tipus semàntic que s'incorporen a la gramàtica; així, la informació que fa possible l'adopció del mecanisme de les restriccions de selecció prové del lèxic. I aquesta informació, juntament amb la resta de la informació lèxica, és considerada idiosincràtica (Chomsky, 1965:87,136,142): el lèxic és el lloc on es consignen els aspectes irregulars (és a dir, que no poden ser sotmesos a les regles) del llenguatge, entre els quals hi ha els trets semàntics que caracteritzen les entrades lèxiques. Vegem el contingut d'unes entrades lèxiques, segons queden especificades per Chomsky (1965:107):

(4) a. (pilota, [+N, +Det _, +Comptable, -Abstracte...])
b. (noi, [+N, +Det _, +Comptable, +Animat, +Humà...])
c. (xuta, [+V, + _ SN, +[+Animat] Auxiliar _ Det [–Animat], +Eliminació-objecte...])

En elles, s'hi recullen els aspectes que Chomsky mateix indica (1965:87) que formen part del contingut idiosincràtic de les entrades lèxiques: els aspectes de l'estructura fonètica que no són predictibles per regla general (en les entrades anteriors, aquesta informació està representada per la forma ortogràfica), les propietats rellevants per al funcionament de les regles transformacionals (a l'entrada de *xuta*, el tret *+Eliminació-objecte* indica que pot ser objecte de la transformació d'eliminació de l'objecte), les propietats que són rellevants per a la caracterització semàntica de la paraula (a les entrades anteriors, els trets pròpiament semàntics no hi són indicats: haurien de ser especificats en el lloc dels punts suspensius),[6] i els trets que indiquen en quines posicions pot aparèixer el mot en qüestió (la part sensitiva al context de les entrades lèxiques: a les dels noms de (4) s'indica que han d'aparèixer després d'un determinant (*+Det _*), mentre que a la del verb *xuta* hi ha, per una banda, informació sobre la seva subcategorització (*+ _ SN*) i, per l'altra, sobre els contextos en què pot aparèixer, segons les seves restriccions de selecció (*+[+Animat] Auxiliar _ Det [–Animat]*)).

En el lèxic, doncs, es consigna la informació que no pot ser tractada de manera general. Així, per exemple, si hi hagués una regularitat en el fet que els verbs transitius, o sia, els que subcategoritzen per un objecte directe, poguessin aparèixer sense aquest complement (com *menjar*, per exemple), no caldria indicar aquesta informació en l'entrada lèxica de cada verb transitiu, ja que una regla general (en aquest cas transformacional) contindria la informació corresponent. Ara bé, Chomsky mateix es preocupa per possibles simplificacions del lèxic. A la secció *L'estructura del lèxic* (1965:164ss.), discuteix algunes possibilitats de simplificar les entrades lèxiques, a base de compactar la seva informació. Proposa dos tipus de mecanismes: unes convencions d'interpretació de les gramàtiques i unes regles de redundància sintàctica: les primeres són universals i les altres, específiques de cada llengua.

Les dues convencions proposades per Chomsky són les següents. En primer lloc, un mecanisme per a no haver d'especificar tots els trets possibles en relació a la subcategorització i a les restriccions de selecció: els primers apareixen explícitament a la regla només si són positius, i els segons, només si són negatius; els altres trets (amb valors negatius o positius, respectivament) són omplerts automàticament si no hi ha conflicte amb els trets formulats explícitament. I en segon lloc, un mecanisme per a evitar la duplicació d'informació en cadenes de trets com *[+Animat, +–Humà]* o *[+N, +–Comú]*, en què la presència d'un tret a la cadena implica necessàriament la presència, amb els valors indicats, dels trets que el precedeixen, així, la presència, en una entrada lèxica del tret *–Humà* indica que el tret *+Animat* és pertinent; i la indicació que un element lèxic és *+Comú* implica que es tracta d'un nom (és a dir, d'un element lèxic marcat *+N*).

Per altra banda, les regles de redundància sintàctica, específiques de cada llengua, serveixen per a simplificar les entrades lèxiques en casos com el de les implicacions entre regles de subcategorització estricta. Així, en català (com en anglès) hi ha verbs

(6) Probablement molts estaríem d'acord amb què trets com *humà*, *abstracte* o fins i tot, *comptable* són ja trets semàntics; Chomsky, però, considera que els trets que són usats en alguna regla sintàctica són realment trets sintàctics i no semàntics (1965: 142).

transitius que poden tenir un complement adverbial de manera (per exemple, *netejar*, *llegir*..., com a (5a)), mentre que n'hi ha d'altres que no l'admeten (com *témer*, *valer*..., com es mostra a (5b)).

(5) a. Va llegir el llibre (amb molta atenció / despreocupadament)
b. Tem les feres (*amb molta atenció / *despreocupadament)

Hi ha, doncs, una implicació en el sentit que si un verb transitiu pot portar un complement adverbial de manera, a més del sintagma nominal objecte, també pot aparèixer només amb l'objecte (o sia, sense el complement adverbial); en canvi, la implicació en sentit contrari no és correcta (per al català o l'anglès), ja que no és cert que si un verb subcategoritza per un complement d'objecte també pot portar un complement adverbial de manera (com *témer*). Per tant, podem establir una regla de redundància sintàctica com la de (6).

(6) [+ _ SN^Compl_Manera] → [+ _ SN]

Així, l'entrada lèxica del verb *netejar* es pot formular com s'indica, parcialment, a (7).

(7) (netejar, [+ _ SN^Compl_Manera, ...])

Una entrada lèxica així, per aplicació de la regla de redundància (6) serà convertida en la de (8a) (també especificada parcialment, només) o fins i tot la de (8b), si ens atenim a la convenció anterior sobre valors negatius en les regles de subcategorització estricta.

(8) a. (netejar, [+ ^ SN_Compl_Manera, + _ SN, ...])
b. (netejar, [+ ^ SN_Compl_Manera, + _ SN, – _ Compl_Manera, – _, ...])

Així doncs, Chomsky preveu ja la possibilitat d'introduir alguna sistematicitat en la informació continguda en el lèxic, que ha de ser complementada per les idees desenvolupades en els dos apartats dedicats, respectivament, a la morfologia flexiva i a la derivacional. De totes maneres, a partir dels plantejaments d'aquest capítol, resulta bastant clar que l'objectiu d'aquesta sistematicitat és més de simplificar la manera de codificar la informació de les entrades lèxiques, que d'establir generalitzacions gramaticals de valor sintàctic; conseqüentment no hi ha cap contradicció entre les afirmacions generals sobre el caràcter idiosincràtic de la informació continguda en el lèxic i aquestes propostes sobre la possibilitat de simplificar el lèxic. Així i tot, és interessant de remarcar que des dels primers moments de l'establiment formal de la teoria generativa es pren en consideració el paper que hi pot jugar el lèxic i que hi apareixen en

germen els mecanismes de l'herència que hauran de tenir un paper tant rellevant en els plantejaments lexicològics més recents (veure, més avall, §9).[7] Ens trobem, per tant, amb una teoria lingüística que fa un plantejament teòric general de com explicar el fenomen de la llengua; per a això, proposa una sèrie de components de la gramàtica, entre els quals hi ha el lèxic. És, doncs, una teoria que aporta una caracterització clarament definida del lèxic, de manera que les tasques lexicogràfiques en aquest marc poden quedar força delimitades en relació a la informació morfosintàctica (és clar que pel que fa a la semàntica la teoria estàndard no aporta cap idea interessant).

Quasi immediatament després de la publicació de Chomsky (1965) sorgeixen els primers intents de desenvolupar una teoria seriosa del lèxic. Aquests esforços adopten dues direccions diferents: per una banda, hi ha el desenvolupament d'una teoria de les relacions entre paraules (o lexemes), de manera que s'afronten els aspectes relacionats amb la formació de paraules (derivació i composició); per l'altra, apareixen intents de millorar el tractament de la informació semàntica dels elements lèxics i, conseqüentment, d'estructurar més clarament la totalitat de la informació que ha d'aportar el lèxic. Un bon exemple del primer moviment el constitueix Botha (1968), l'objectiu fonamental del qual és precisament justificar la necessitat de tractar de manera sistemàtica la informació lèxica (superant, doncs, la visió que el lèxic conté només informació de tipus idiosincràtic); en aquest estudi, aquesta justificació es fonamenta sobretot en la necessitat que té la teoria lingüística d'explicar les relacions existents entre diverses paraules (a través dels processos de derivació i, sobretot, de composició). Un altre exemple, aquest des de la lexicografia, el constitueix Corbin (1981), on s'analitzen les pràctiques lexicogràfiques d'alguns diccionaris francesos, des del marc teòric aportat per la teoria estàndard.

El segon tipus de reacció pot ser exemplificat amb la posició defensada per Fillmore (1971), on estudia precisament els diversos tipus d'informació que ha de contenir el lèxic, entre els quals inclou el rol que cada argument juga en la situació indicada per les paraules predicatives, les pressuposicions (o condicions) que fan que la paraula sigui usada adequadament, el tipus de relació conceptual o morfològica que la paraula té amb altres elements del lèxic, i el seu significat. Tot i que es tracta fonamentalment d'una proposta programàtica, ens permet de fer-nos una idea de com semblava en aquell moment que els aspectes semàntics i pragmàtics podrien ser incorporats en les entrades lèxiques. La proposta sembla un avenç important en la direcció d'incorporar nous aspectes a la part tractada (és a dir, formalitzada) del llenguatge. Ara bé, en moltes ocasions (i això va ocórrer en els moments històrics posteriors a la publicació de Chomsky (1965)), la introducció de la semàntica (i, és clar, de la pragmàtica) ha comportat una pèrdua de rigorositat en el plantejament explicatiu i formal; serà només fins a la generalització de les tècniques semàntiques desenvolupades a la lògica i a les ciències de la computació que aquest risc podrà ser superat. De totes maneres, des del punt de vista del lexicògraf, un plantejament així apunta en una direcció molt atractiva, ja que

(7) Chomsky se'ns revela, un cop més, com un geni, que és capaç d'intuir, molts anys abans de disposar de l'aparell formal adequat, la rellevança que tindran els mecanismes d'organització del lèxic.

sense oblidar els aspectes sintàctics fonamentals (tal com han estat recollits en la versió estàndard de la teoria generativa) s'indica un camí per a la introducció de la informació semàntica, que tot diccionari ha de contenir, lligada d'alguna manera amb la informació pròpiament sintàctica.

La tumultuosa història de la gramàtica generativa els últims anys 60 (Newmeyer, 1986) va centrar-se en la relació entre la sintaxi i la semàntica (específicament, en si les estructures semàntiques eren també generades –és a dir, si hi havia regles que les produïen– o eren simplement el resultat de la interpretació de les estructures sintàctiques),[8] i també en la de la sintaxi i la morfologia. En part les dues qüestions anaven lligades, en tant que la descomposició de les paraules (tasca fonamentalment morfològica) podia anar directament a la interpretació semàntica o podia donar lloc a regles del component de base. Més que la polèmica en si, el que ens interessa aquí és de fer constar el seu resultat, en la que fou l'opció que ha perdurat posteriorment. El corrent central de la gramàtica generativa, és a dir el liderat per Chomsky, va acabar adoptant el principi exposat a Chomsky (1970) que es coneix amb el nom d'*hipòtesi lexicalista* i que ha donat peu a la majoria dels estudis en morfologia dins d'aquest corrent central de la teoria generativa.

La hipòtesi lexicalista, com suggereix el seu nom, concedeix al lèxic una importància més gran, ja que tracta com a lèxiques unes relacions que s'havien considerat gramaticals. Com menciona Newmeyer (1986:107), la funció essencial de la hipòtesi va ser prohibir les transformacions que provoquessin un canvi de categoria; d'aquesta manera, s'evitava que, per exemple, un verb o un adjectiu fos transformat en un nom i es restringia fortament l'abast del camp d'aplicació de les regles transformacionals. Així, la relació entre *Jaume ha arribat* i *l'arribada de Jaume* ha de ser formulada dins el lèxic i no per mitjà de regles transformacionals en la gramàtica. Naturalment, com Chomsky (1970:22) indica, una posició com aquesta només pot ser presa quan el lèxic ha estat separat del component de base i quan les entrades lèxiques contenen informació contextual (a través de les subcategoritzacions i de les restriccions de selecció); hi pot haver entrades lèxiques de categories diferents que mantinguin les mateixes condicions contextuals.

Com és natural, una posició així va obrir les portes a tota una sèrie d'estudis morfològics en el marc de la teoria generativa. Si fins ara, en general els generativistes no distingien entre els fenòmens morfològics i els sintàctics, l'acceptació d'aquesta hipòtesi comporta una distinció molt clara entre les relacions lèxiques i les gramaticals, i entre els seus constructes: els de les primeres no poden ser afectats per les regles de la gramàtica. D'aquesta manera s'engega la primera gran onada de treballs morfològics de la gramàtica generativa: es tracta de precisar la naturalesa, els límits i les condicions d'aplicació de les regles lèxiques, centrats sempre en els mecanismes de la derivació i la composició.

(8) D'aquí els noms de *semàntica generativa* i *semàntica interpretativa*: el primer designa una teoria que va tenir els seus moments de màxim esplendor aquests anys, i el segon és el lema dels generativistes que es van mantenir en l'ortodòxia i van continuar defensant l'autonomia de la sintaxi.

## 7. LA TEORIA DE LA RECCIÓ I EL LLIGAM

El plantejament teòric actual del corrent generativista majoritari és el conegut sota el nom de *Teoria de la recció i el lligam*, normalment abreujat sota les sigles angleses GB (de *Government and Binding*). En aquesta nova versió, han desaparegut les regles de reescriptura com a element central en la construcció de les representacions sintàctiques; el seu lloc és ocupat per una sèrie de principis generals, que són els que determinen la gramaticalitat de les oracions. L'estructura general de la teoria és encara molt similar a la de la teoria estàndard, amb un doble nivell sintàctic (superficial i profund), i dos mòduls interpretatius (la forma fonètica i la forma lògica).

Ara bé, la manera com es relacionen els constructes de cada estructura és força diferent de la que trobàvem a la teoria estàndard; allà podíem especificar d'una manera bastant clara el tipus de regla que corresponia a cada relació: així, per exemple, l'estructura profunda i l'estructura superficial estaven relacionades a partir de les regles transformacionals, que no intervenien en cap altra relació. Aquí, en canvi, són els diversos principis els que determinen les relacions que s'estableixen entre les diferents estructures (teoria X', teoria $\theta$, moviment d'$\propto$, principi de projecció, criteri $\theta$, etc.); cosa que es veu complicada, sovint, pel fet que actuen en més d'un nivell.

Convé que vegem quin és el paper del lèxic en aquesta estructura complexa, tenint en compte que presentacions generals de la teoria com la de Sells (1985) no l'hi situen.[9] El lèxic a GB, com a totes les teories generatives antecessores seves, és el lloc on s'emmagatzema la informació pertinent relacionada amb cada entrada lèxica. Pel que fa a la sintaxi (ja que se segueix deixant de banda la semàntica), aquesta informació es concreta en quatre apartats: una matriu fonològica, un tret categorial, un esquema sintàctic, i una graella de papers temàtics (rols $\theta$).

Així, l'entrada lèxica del verb *donar* seria aproximadament com segueix (Zubizarreta, 1986):

| | | | | | |
|---|---|---|---|---|---|
| (9) | donar: | +V, | (AGENT, | destí, | tema) |
| | | | | SP | SN |

En el llistat dels papers temàtics dels predicats s'estableix una distinció entre els papers interns i els externs (Williams, 1981), que queda reflectida a (9), on l'argument extern és representat en majúscules i els interns, en minúscules. Aquesta distinció, doncs, és especificada en el lèxic i una classificació dels verbs segons els seus arguments i el corresponent esquema sintàctic ho ha de tenir en compte. Hi ha verbs que no tenen arguments interns (com *riure*), mentre que n'hi ha d'altres que en tenen un (com *veure*) o més (com *donar*), dels quals alguns tenen argument extern (com *veure* i *donar*) i d'altres no (com *arribar*). La connexió entre els arguments interns i la seva caracterització sintàctica es fa en el lèxic, mentre que la manera com els arguments

(9) Vegeu, però, la presentació de de Geest i Jaspers (1991), que seguim en la nostra descripció.

externs es realitzen a la superfície (si és que s'hi realitzen) depèn de factors de tipus sintàctic que operen a nivell posterior.

Seguint l'establert a la teoria de la X barrada (Jackendoff, 1977), a GB l'enllaç entre les entrades lèxiques i les representacions sintàctiques que se'n deriven, ha de seguir el principi de projecció, segons el qual les representacions de tots els nivells sintàctics es projecten a partir del lèxic d'acord amb les propietats de subcategorització dels elements lèxics (Chomsky, 1981:29). D'aquesta manera, les propietats dels constructes sintàctics són determinades per les propietats dels elements lèxics dels quals són una projecció. A la vegada, es manté la hipòtesi lexicalista, de manera que les transformacions no poden alterar les estructures interiors dels elements lèxics. És important de notar que aquests dos principis atorguen al lèxic un lloc preeminent en la teoria, que abans no havia tingut; el mecanisme de les regles d'inserció lèxica que apareixia a la teoria estàndard, no garantia aquesta immutabilitat de la informació lèxica al llarg dels processos posteriors, sobretot transformatius. Així, amb GB, el lèxic adquireix un paper preponderant en la teoria generativa com mai no havia tingut des del seu naixement. Sens dubte aquests plantejaments suposen un capgirament important en el generativisme, que havia representat una preponderància absoluta de la sintaxi enfront de les altres subdisciplines lingüístiques (com la morfologia o la semàntica, més tradicionalment lligades a la lexicografia).

GB, doncs, atorga una posició determinada al lèxic, amb unes funcions molt concretes en relació a les operacions de tipus sintàctic que es duen a terme en el procés de generació de les oracions. Un plantejament així permet de posar sobre la taula diferents hipòtesis i contrastar les unes amb les altres. No és estrany, doncs, que un dels debats més rics en els darrers anys hagi estat sobre la relació de la sintaxi amb la morfologia; d'una manera especial, l'explicació dels fenòmens de la derivació i de la composició (tradicionalment inclosos en la morfologia) ha centrat bona part d'aquesta polèmica. En general, el centre de la discussió ha estat el veure fins a quin punt hi ha uns processos morfològics independents (i diferents) dels processos sintàctics; la discussió, és clar, pren tot el seu interès quan es comencen a considerar factors universals (o, almenys, multilinguals), ja que hi ha fenòmens que en unes llengües prenen una manifestació sintàctica i en d'altres en prenen una de morfològica. Un dels centres d'interès a GB actualment és la relació entre les estructures predicatives dels elements lèxics de base i les dels seus derivats (per exemple, entre la dels verbs i les dels noms que se'n deriven): es tracta de veure com es projecta l'estructura argumental del predicat de base en els seus derivats i quines modificacions ha de tenir el seu esquema sintàctic, de manera que no es perdin les propietats que convé traspassar de l'un als altres i que es puguin mantenir els principis essencials de la teoria. En gran part, els plantejaments en aquest tema estan lligats al problema general del lligam entre la graella temàtica i l'esquema de subcategorització sintàctica dels predicats (com a Grimshaw, 1990).

Per tant, segons els principis de GB, el lèxic constitueix un element central de la teoria lingüística, de manera que les peces lèxiques usades en la generació d'oracions

resulten fonamentals en la forma final d'aquestes. A la vegada, hi ha un seguit de regles de caire morfològic que contribueixen a enllaçar uns elements lèxics amb altres. Aquestes regles fan possible la reducció del nombre d'entrades lèxiques finals, ja que permeten d'explicar l'aparició d'elements lèxics nous (creativitat en la producció lèxica). A la vegada representen un primer pas en l'explicació de les relacions semàntiques entre entrades lèxiques, ja que en principi s'accepta que la productivitat lèxica va lligada a la composicionalitat en la interpretació; així, les paraules aparentment derivades, que no poden ser interpretades composicionalment (a partir de la interpretació dels seus morfemes i lexemes), són considerades morfològicament irregulars i, conseqüentment, s'ha proposat que siguin llistades en el diccionari (di Sciullo i Williams, 1987).

## 8. LA GRAMÀTICA LÈXICO-FUNCIONAL

En els moments més àlgids de la gramàtica transformacional, en què les transformacions eren utilitzades de manera molt lliure, el seu excessiu poder expressiu va conduir molts lingüistes a buscar-hi alternatives. Es tractava d'intentar construir una gramàtica expressivament adequada, és a dir, que no generés més oracions de les que són considerades naturals en la llengua en qüestió i que no generés la mateixa oració per dos camins diferents. A la vegada, es considerava que el tractament de la informació lingüística en termes d'estructura de constituents, que en principi apareixia com a molt adequat per a l'anglès, no resultava natural per a moltes altres llengües en impedir-hi generalitzacions interessants. Probablement la primera teoria que va aparèixer com una alternativa seriosa fou la Gramàtica relacional de D. Perlmutter. De totes maneres, la primera proposta que ha tingut una gran influència en el desenvolupament posterior de la lingüística teòrica ha estat la Gramàtica lèxico-funcional (LFG, de *Lexical-Functional Grammar*), que ha iniciat el darrer moviment lexicalista en la teoria gramatical. L'obra clàssica de presentació de LFG és Bresnan (1982); de totes maneres, n'existeixen diverses descripcions breus, com Sells (1985) i Wescoat i Zaenen (1991).

Com el seu mateix nom indica, ja des dels seus inicis, LFG ha concedit una gran importància al lèxic.[10] Dit breument, el lèxic és el lloc on s'expressen totes les relacions locals (Wescoat i Zaenen, 1991:119): en el lèxic es descriuen tots els processos que alteren l'estructura argumental dels predicats; així, els fenòmens gramaticals que en altres teories generatives s'explicaven mitjançant transformacions, es descriuen amb operacions de tipus lèxic. Un bon exemple d'això és el tractament de la passiva. Tant a la teoria estàndard, com a GB, la passiva és explicada en termes d'una transformació que implica (entre altres coses) el moviment del sintagma nominal objecte a la posició de subjecte; es tracta, doncs, d'una relació entre estructures sintàctiques: l'estructura generada per les regles de base és relacionada (per mitjà d'una transformació o d'un moviment) amb una altra estructura sintàctica.[11]

(10) L'altra etiqueta, *funcional*, fa referència a la noció de funció sintàctica, que LFG considera un primitiu de la teoria gramatical: així, nocions com la de subjecte o objecte adquireixen un paper descriptiu i un estatus precís: són generalitzacions sobre la manera com els predicats es relacionen amb els seus arguments, independentment de la seva forma sintàctica (cas, règim preposicional...).

(11) Notem, de passada, que aquesta segona estructura sovint és d'unes característiques tals que hauria pogut ser generada directament per les regles de base; així, per exemple, l'estructura configuracional de la passiva *El president fou rebut amb euforia* és idèntica a la de *El president ha arribat amb retard*, és a dir, la mostrada per l'esquema següent: $[_{O}[_{SN}\ ]\ [_{SV}\ [_{V}\ ]\ [_{SN}\ [_{V}\ ]\ [_{SP}\ ]]]]$.

El lèxic ocupa una posició central en la teoria. La informació lèxica constitueix l'enllaç entre els dos nivells d'informació sintàctica que pressuposa LFG, el de l'estructura **c** (és a dir, de constituents) i el de l'estructura **f** (o funcional): les entrades lèxiques dels elements predicatius aporten la doble informació, i determinen la manera com es relaciona l'una amb l'altra. Òbviament, això pot ser així, perquè el lèxic és un component de la gramàtica altament estructurat. Hi trobem, almenys, tres elements diferents: les entrades lèxiques pròpiament dites, les regles morfològiques i les regles lèxiques. Les entrades lèxiques de LFG es caracteritzen per indicar la subcategorització dels predicats per mitjà de les funcions sintàctiques, enlloc de les categories sintàctiques o rols $\theta$. A més, s'hi indica de manera precisa com ha de ser usada la informació que contenen en construir-se les estructures sintàctiques. Vegem un parell d'entrades lèxiques, simplificades, per al nom *noia* i per al verb *veure*:

| | | | | |
|---|---|---|---|---|
| (10) | a. | noia | N | (↑ PRED) = 'noia' |
| | | | | (↑ PERS) = 3 |
| | | | | (↑ NOM) = sing |
| | b. | veure | V | (↑ PRED) = 'veure <(↑ SUBJ) (↑ OBJ)>' |

Entre cometes s'indica el valor de l'atribut PRED (de predicat), que és de tipus semàntic.[12] La fletxa cap amunt (↑) indica que l'atribut que la segueix immediatament ha de ser entès com un atribut del nus mare de l'element en qüestió (en les entrades lèxiques, el nus preterminal). Les equacions presents a les entrades lèxiques serveixen per a construir les estructures **f**, per mitjà de l'operació de la unificació, d'acord amb les anotacions funcionals de les regles sintagmàtiques (vegeu §9).

Les entrades lèxiques dels predicats indiquen entre parèntesis angulats la seva estructura de subcategorització, en la qual cada element és referit per la seva funció sintàctica; així, el verb *veure* a (10b) subcategoritza per un subjecte i un objecte. Quan un predicat té complements no subcategoritzats (els verbs de les estructures anomenades de *raising* o elevació), les funcions no subcategoritzades són indicades fora dels parèntesis angulats, com es pot veure a les entrades lèxiques parcials següents, del verb *semblar*, d'elevació al subjecte, i del verb anglès *expect* ('esperar'), d'elevació a l'objecte:

| | | | | |
|---|---|---|---|---|
| (11) | a. | semblar | V | (↑ PRED) = 'semblar <(↑ XCOMP)> (↑ SUBJ)' |
| | | | | (↑ SUBJ) = (↑ XCOMP SUBJ) |
| | b. | expect | V | (↑ PRED) = 'expect <(↑ SUBJ) (↑ XCOMP)> (↑ OBJ)' |
| | | | | (↑ OBJ) = (↑ XCOMP SUBJ) |

En aquestes entrades, la funció XCOMP és reservada per a clàusules d'infinitiu o gerundi; així, les entrades de (11) servirien per a oracions del tipus de les de (12):

(12) En aquestes entrades, el valor semàntic està extraordinàriament simplificat. Per a una proposta de semàntica par a LFG, podeu veure Fenstad i al. (1987), que incorpora un nivell semàntic d'acord amb la teoria de la Semàntica de situacions de Barwise i Perry (1983).

(12) a. Els nois semblen tenir gana
b. I expected John to arrive on time (lit.: 'Esperava J. arribar a l'hora')

La segona línia de cada entrada serveix per a indicar que l'element no subcategoritzat pel verb d'elevació (el marcat SUBJ per a *semblar* i OBJ per a *expect*) són el subjecte de la clàusula d'infinitiu que té la funció de XCOMP. Notem, de passada, que aquesta representació dels verbs d'elevació comporta, també, un tractament lèxic d'aquest fenomen: és en la mateixa entrada lèxica que s'estableix la relació entre la realització superficial de l'argument i la seva posició en el component de base. A LFG es preveu l'aplicació de les regles morfològiques abans de la construcció de les estructures sintàctiques; per tant, es tracta d'operacions que es duen a terme a l'interior del lèxic, el resultat de les quals serveix d'entrada a les regles sintàctiques. Així, es poden escriure entrades lèxiques per a afixos, com la següent, adaptada de Sells (1985), per a la tercera persona del singular del present dels verbs de la primera conjugació.

(13) -a AFIX (↑ SUBJ PERS) = 3
(↑ SUBJ NOM) = sing
(↑ TEMPS) = pres

Naturalment, les formes irregulars no poden ser tractades d'aquesta manera, sinó que han de ser formulades explícitament en el lèxic. Per tal que no hi hagi interferències entre els processos regulars i els irregulars, cal una bona estructuració de tota aquesta informació. Així, en general, cal pensar en els processos regulars com processos per defecte, és a dir, processos que es realitzen sempre que no hi ha una indicació en el sentit contrari, o sia, sempre que no hi ha una formació irregular (vegeu els mecanismes descrits breument a § 9).

Finalment, LFG preveu l'existència de regles lèxiques, que serveixen per a expressar relacions entre elements lèxics. Aquestes regles s'apliquen també a l'interior del lèxic i, per tant, és només el seu resultat que és utilitzat per les regles sintàctiques. Un exemple característic de regla lèxica a LFG és la de la passiva, la part essencial de la qual trobem a (14):

(14) (SUBJ) → ϕ / (COMPL_AG)
(OBJ) → (SUBJ)

D'acord amb la regla de (14), l'entrada lèxica d'un verb que subcategoritza per a les funcions de subjecte i objecte pot ser convertida en una entrada lèxica, l'esquema de subcategorització de la qual és com s'indica a (15): sense complement agent a (15a) i amb aquest complement a (15b).

(15) a. <(↑ SUBJ)>
b. <(↑ COMPL_AG) (↑ SUBJ)>

Els altres fenòmens que comporten una alteració en l'estructura argumental dels predicats són formulats també per mitjà de regles lèxiques; així, des d'aquest punt de vista, fenòmens com la causativització, la reflexivitat, la intransitivització, l'eliminació de l'objecte, o les construccions de doble objecte, serien tractades com a fenòmens lèxics mitjançant regles lèxiques.

Així doncs, LFG ens porta a una visió del lèxic molt més rica que el simple llistat d'entrades lèxiques. En primer lloc, la informació que aporta el lèxic està integrada plenament en el conjunt de la informació lingüística tractada per la gramàtica, fins al punt que molta d'aquesta darrera deriva, de manera directa o indirecta, de la primera. En segon lloc, els elements en el lèxic estan organitzats, de manera que s'estableixen relacions entre ells: per una banda, els processos flexius permeten de derivar unes formes lèxiques d'unes altres i, per l'altra, les regles lèxiques relacionen unes entrades lèxiques amb les altres, permetent així de formular semblances i paral·lelismes entre elles. D'aquesta visió del lèxic a una concepció lexicogràfica totalment estructurada hi va només un pas. La propera (i darrera) teoria que tractarem ens mostrarà ja una organització totalment estructurada del lèxic.

## 9. LA GRAMÀTICA D'ESTRUCTURA SINTAGMÀTICA ORIENTADA AL NUCLI

HPSG (de *Head-driven Phrase Structure Grammar*) és la més recent de les *gramàtiques d'unificació* (Shieber, 1985). Els seus aspectes fonamentals poden ser estudiats amb un cert detall gràcies a la publicació de dos volums de presentació general (Pollard i Sag, 1987; en premsa). Ja des dels plantejaments inicials, aquesta teoria manlleva deliberadament concepcions i tractaments elaborats originalment en d'altres marcs teòrics. Com Pollard i Sag (1987:1) assenyalen no es tracta tant de copiar d'ací i d'allà, com d'oferir un punt de vista nou a molts plantejaments que han estat desenvolupats en la gran tradició intel·lectual que hi ha darrera l'evolució de les gramàtiques formals els darrers anys.

El primer tret distintiu de HPSG al qual hem de fer referència és que codifica la informació lingüística en estructures de trets. En això, adopta una posició comuna a tots els formalismes i teories basats en la unificació (Shieber, 1985), que ja es troba, encara que parcialment, a LFG. Les estructures de trets són concebudes com *entitats que descriuen o representen un altre objecte tot especificant valors per a diversos atributs de l'objecte descrit* (Pollard i Sag, 1987:28); en general, anomenem *tret* qualsevol parell d'atribut i valor. Normalment, les estructures de trets contenen només informació parcial sobre l'objecte descrit. Així, podem tenir una estructura de trets amb l'única informació que el nom (atribut) de l'objecte és *Barcelona* (valor), que descriurà (parcialment) un nombre indeterminat d'objectes, entre els quals diversos

pobles i ciutats; a aquesta estructura de trets li podem afegir algun tret nou, com que la població (atribut) de l'objecte és de més d'un milió d'habitants (valor) i que és del país (atribut) Catalunya (valor), de manera que llavors l'estructura de trets ja no descriu, en el món actual, cap altre objecte que la capital de Catalunya; encara que aquesta darrera descripció només tingui un sol objecte en el món que se li adeqüi, també es tracta només d'una descripció parcial de Barcelona! En general, aquestes estructures de trets es representen amb matrius d'atributs i valors, com la de (16), que recull la informació (parcial) que acabem de mencionar sobre Barcelona:

(16)
$$\begin{bmatrix} \text{nom:} & \text{Barcelona} \\ \text{població:} & > 1.000.000 \text{ hab} \\ \text{país:} & \text{Catalunya} \end{bmatrix}$$

Quan en una teoria lingüística s'usen aquestes estructures, es fa amb la finalitat de representar les propietats formals que tenen els objectes lingüístics. Per als creadors de HPSG els objectes lingüístics són signes (i citen explícitament Saussure), el significant i significat dels quals reinterpreten, respectivament, com un tipus de situació de parla i com una forma determinada de ser les coses en el món. Convé que notem que en el vessant semàntic del signe adopten una postura absolutament realista, declarant-se així hereus de la tradició semàntica que, provinent de la lògica i la filosofia, va entrar en lingüística a través de l'obra de Montague (1974). De fet, les opcions semàntiques dels creadors de HPSG s'encaminen més cap a la versió anomenada de la semàntica de les situacions (Barwise i Perry, 1983), que cap a versions més clàssiques d'aquesta tradició semàntica.

El segon tret distintiu de HPSG que cal que mencionem és que, a causa de la representació de la informació lingüística amb estructures de trets, aquesta teoria no demana diversos nivells de representació. Tots els aspectes descriptius d'un signe lingüístic determinat poden coexistir en una única estructura. Així, en un únic signe hi pot aparèixer, alhora, informació de tipus fonològic (o gràfic), morfosintàctic, de sintaxi relacional i semàntic. Considerem, per exemple, l'entrada lèxica de la forma verbal catalana *menjà*, adaptada (i simplificada) de les que proposen Pollard i Sag:

(17)
$$\begin{bmatrix} \text{fonologia} & \text{menjà} \\ \text{sintaxi} \mid \text{local} & \begin{bmatrix} \text{nucli} & \begin{bmatrix} \text{major} & \text{V} \\ \text{forma-verb} & \text{finita} \\ \text{auxiliar} & - \end{bmatrix} \\ \text{subcat} & \text{SN[ACC, NORM]}_2\text{, SN[NOM, NORM, SING, 3PERS]}_1 \\ \text{lex} & + \end{bmatrix} \\ \text{semàntica} \mid \text{contingut} & \begin{bmatrix} \text{relació} & \text{menjar} \\ \text{arg1} & 1 \\ \text{arg2} & 2 \end{bmatrix} \end{bmatrix}$$

En ella, hi trobem els quatre tipus d'informació que mencionàvem: la fonològica (representada, de manera simplificada, pel tret FONOLOGIA = menjà), la morfosintàctica (expressada per una part dels trets de l'atribut SINTAXI|LOCAL|NUCLI), la de sintaxi relacional (continguda fonamentalment en el tret SINTAXI|LOCAL|SUBCAT), i la semàntica (expressada com a valor de l'atribut SEMÀNTICA).

Un tractament així té l'avantatge que es poden establir d'una manera molt fàcil i elegant les relacions que hi ha entre els complements exigits sintàcticament i els arguments de la relació expressada; a la vegada, un element rector pot imposar restriccions als seus complements tant d'ordre sintàctic com semàntic. Tots dos factors apareixen a l'entrada lèxica de *menjà*: els índexs 1 i 2 indiquen, respectivament, que el contingut del sintagma nominal subjecte és el mateix que l'argument primer del contingut de *menjar*, i que el contingut de l'objecte és idèntic a l'argument segon del seu contingut.

Finalment, el tercer tret distintiu de HPSG que ens interessa de remarcar és el lloc central que hi ocupa la unificació. Com en els altres formalismes d'unificació, per a HPSG la unificació és una operació fonamental a l'hora de relacionar estructures de trets. La informació continguda en elles es combina gràcies a aquesta operació, de manera que en resulta una altra estructura de trets. Considerem les estructures de trets de (18):

(18) a. $\begin{bmatrix} \text{categoria:} & \text{sn} \\ \text{concordança:} & [\text{nombre: singular}] \end{bmatrix}$

b. $\begin{bmatrix} \text{categoria:} & \text{sn} \\ \text{concordança:} & [\text{persona: tercera}] \end{bmatrix}$

c. $\begin{bmatrix} \text{categoria:} & \text{sn} \\ \text{concordança:} & [\text{nombre: plural}] \end{bmatrix}$

La informació de les estructures de trets de (18a) i (18b) és diferent, però compatible; en canvi, la informació de (18a) i (18c) es incompatible. Per tant, no és possible de pensar en una estructura de trets que contingui tota la informació de (18a) i (18c) alhora, mentre que sí que es pot construir una estructura de trets que contingui la informació tant de (18a) com de (18b): aquesta estructura de trets resultant serà la unificació de les dues originals, que queda consignada a (19):

(19) a. $\begin{bmatrix} \text{categoria:} & \text{sn} \\ \text{concordança:} & \begin{bmatrix} \text{nombre: singular} \\ \text{persona: tercera} \end{bmatrix} \end{bmatrix}$

En general, i informalment, podem definir la unificació de dues estructures de trets A i B com l'estructura de trets C que conté tota la informació de A i B i només la informació de A i B.

L'operació d'unificació pot ser usada per a explicar bona part de les condicions que regulen les relacions sintagmàtiques entre elements lingüístics. Així, en els forma-

lismes d'unificació més simples, com les DCGs (Pereira i Warren, 1980) i PATR (Shieber, 1986), les regles d'estructura sintagmàtica usuals són complementades amb informació addicional respecte a les característiques dels elements lingüístics que hi intervenen, i la informació corresponent a la categoria mare és construïda a partir de la unificació de la informació de les filles.

Equipats amb aquestes dades sobre HPSG, podem passar a considerar-ne la visió sobre el lèxic. HPSG, com LFG (vegeu, abans, §8) i com la gramàtica categorial (Pollard i Sag, 1987:191), adopta la postura de lexicalitzar al màxim la informació lingüística, cosa que li permet aconseguir *una reducció dràstica del nombre i la complexitat de les regles d'estructura sintagmàtica*. Naturalment, aquesta opció comporta un augment notable de la complexitat dels signes lingüístics. Com a mostra d'aquest fet, podem observar a (20) l'entrada lèxica anglesa que proposen per a la forma verbal *likes* (tercera persona del signular de *like*, 'agradar').

(20)
```
[ fonologia      likes
                 [ nucli      [ major        V        ]
                              [ forma_verb   finita   ]
                              [ auxiliar     –        ]
  sintaxillocal
                 subcat       SN[ACUSATIU] 2 , SN[NOMINATIU]
                                                1 [ persona   tercera  ]
                                                  [ nombre    singular ]
                 lex          +
                 [ contingut  [ relació      like ]
                              [ arg1         1    ]
                              [ arg2         2    ]
                              [ localització 3    ]
  semàntica      índexs       [ variable     3
                              [ restricció   [ relació   coincideix-temporalment ]
                                             [ arg1      3                       ]
                                             [ arg2      l_d                     ]
```

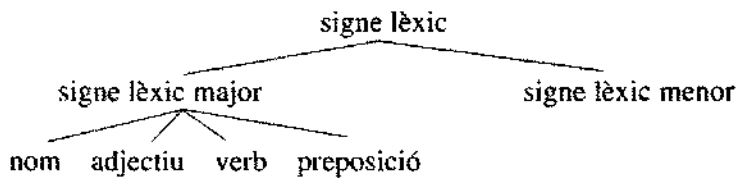
Si ens enfrontem amb entrades lèxiques d'aquesta complexitat (i les entrades lèxiques definitives encara haurien de ser més complexes), resulta evident que calen mecanismes per a estructurar la informació lèxica: ni el model teòric seria explicatiu (hi hauria un nombre molt elevat de relacions que no podrien ser explicades), ni podríem pretendre cap mena d'adequació psicològica per al model (seria impensable que els parlants d'una llengua tinguessin interioritzat un lèxic així, ja que és impossible de memoritzar tota aquesta informació per a cada paraula), ni es podria tractar d'un model computacionalment viable (a causa de la redundància en la informació lèxica, es gastarien quantitats impensables d'espai en la memòria dels programes). La informació lèxica, doncs, ha d'estar convenientment estructurada, de manera que la totalitat de

la informació que la teoria pressuposa per a entrades lèxiques com la de (20) pugui ser deduïda per mecanismes precisos a partir d'unitats més petites d'informació. El lèxic que proposen a HPSG està estructurat a partir de dos grans tipus de relacions: les que són determinades per la jerarquia de tipus lèxics i les que queden explicitades en regles lèxiques. A continuació les descriurem breument tal com queden exposades, per exemple, a Pollard i Sag (1987) i Flickinger (1987); vegeu Shieber (1985) per a una descripció paral·lela més elemental.

Les estructures de trets estan tipificades; és a dir, hi ha diversos tipus d'estructures de trets i les de cada un representen classes diferents d'objectes lingüístics. Així, per exemple, s'estableix una diferència clara entre els signes lèxics i els sintagmàtics: representen dues classes d'objectes lingüístics i, conseqüentment, són de dos tipus diferents; d'aquesta manera, es pot establir la diferència entre la paraula *nenes* i el sintagma nominal *nenes*: la primera és representada per un signe lingüístic de tipus lèxic, mentre que la segona ho és per un de tipus sintagmàtic. Com que cada tipus de signe representa uns objectes lingüístics diferents, és natural que les seves estructures continguin informació diferent; per tant una caracterització dels diferents tipus d'estructures de trets comportarà una definició dels atributs que hi poden aparèixer i dels tipus que aquests poden tenir com a valor. Així, entre les estructures de trets que representen les dues interpretacions de *nenes* hi haurà, almenys, una diferència en el tret LEX: quan *nenes* és interpretat com a paraula el seu valor serà '+', mentre que quan ho és com a sintagma, serà '–'.

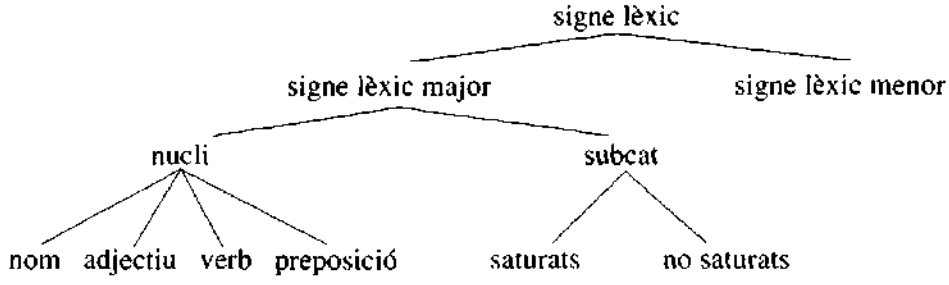
En un primer pas en l'organització del lèxic, podem establir una jerarquia en els nostres tipus, de manera que els més generals (que són els que tenen menys informació) apareguin en els punts superiors de la jerarquia. Així, aquesta estructura jeràrquica podria començar com s'indica a (21):

(21)




Cada tipus en la jerarquia conté una informació determinada. Per exemple, el tipus *signe lèxic* conté la informació SINTAX|LOCAL|LEX = '+', i el signe lèxic major nominal conté, entre altres, la informació SINTAX|LOCAL|NUCLI|MAJOR = N. Ara bé, l'esquema de (21) és una simplificació que convertiria aquest tractament en inviable: és evident que els signes lèxics majors no es classifiquen només d'acord amb els seus trets nuclears, sinó que, almenys, s'han de tenir en compte també les seves propietats de subcategorització. Per tant, cal modificar l'esquema de (21) i convertir-lo en el de (22), que té en compte els dos criteris de classificació dels signes majors:

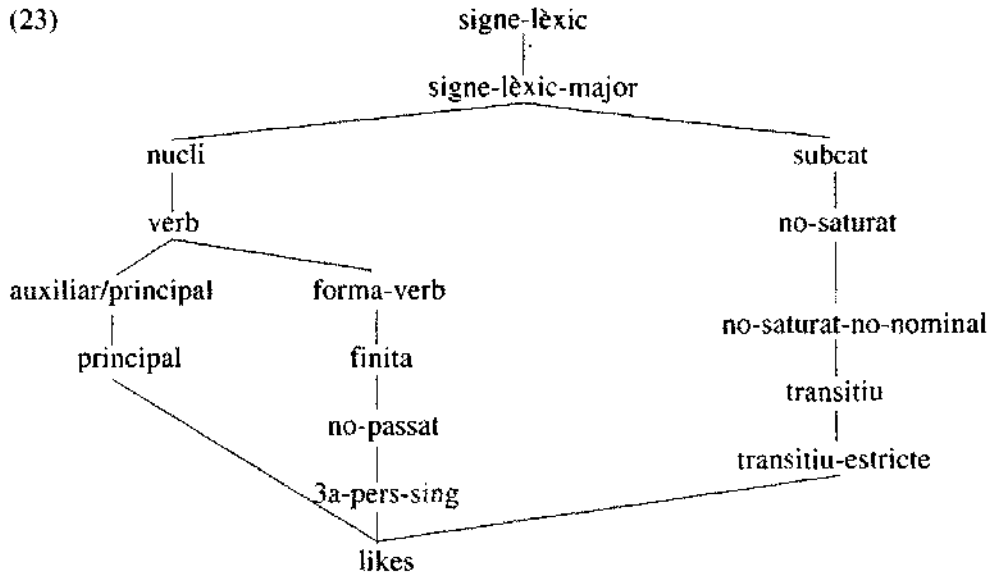
(22)




Entre els diferents tipus de la jerarquia s'hi estableix una relació anomenada de *subsumpció*, de manera que qualsevol element en la jerarquia subsumeix els tipus en posició inferior (per exemple, el tipus *signe lèxic major* subsumeix el tipus *nom*). La relació de subsumpció pot ser definida informalment com la inversa de la relació de *la seva informació és inclosa en*. La relació de subsumpció determina una estructura d'ordre per a tots els elements relacionats, que és la que representem amb la jerarquia. Cada element en la jerarquia rep o hereta la informació dels elements en posició superior (és a dir, dels elements que el subsumeixen) i, a més, com hem vist, pot contenir informació pròpia. Quan ens trobem amb una classificació creuada, com la de (22), els elements inferiors en la jerarquia, reben la informació que els prové dels dos braços de la jerarquia. Així, un verb com *likes* heretarà la informació pròpia del tipus *verb*, però també la del tipus *no saturat* (ja que es tracta d'un verb que té exigències de subcategorització). Com noten Pollard i Sag (1987:193), aquest tractament de la informació lèxica és molt habitual a les llengües amb propietats flexives importants: la forma concreta d'un temps verbal depèn del temps i la persona, per una banda, i del paradigma de conjugació, per l'altra.

D'aquesta manera es pot arribar a simplificar enormement la informació codificada per a cada entrada lèxica particular: de fet en la majoria dels casos la codificació consistirà en la determinació de la forma fonològica de la paraula en qüestió i de la seva caracterització semàntica (exactament, notem, els dos vessants del signe lingüístic segons Saussure); a més, només caldrà determinar en quin punt (o en quins punts) de la jerarquia s'hi ha d'integrar el signe, de manera que hereti tota la informació pertinent dels tipus que el subsumeixen. Vegem-ho amb un exemple. La posició de la forma verbal *likes* en la jerarquia és la indicada per l'esquema de (23), adaptat de Pollard i Sag (1987:205):

(23)




Un plantejament així s'anomena d'herència múltiple, perquè la informació que conté un element lèxic determinat prové de diversos camins en la jerarquia, com queda expressat gràficament a (23) en relació a la forma verbal *likes*.

No obstant, un sistema així, estructurat només a partir de l'herència, no és encara adequat. En primer lloc, no pot tractar les excepcions, és a dir, quan convé que l'element en qüestió hereti quasi tota la informació (però no tota) que li prové d'un camí determinat. A la vegada, no permet tampoc d'establir les relacions que hi ha entre moltes entrades lèxiques: per exemple, les que resulten de la flexió d'una única paraula (la base), o les que són conseqüència dels processos productius de derivació. És indubtable que hi ha una relació entre les formes verbals *cantar* i *canten* i, a més, que aquesta relació és general a molts parells de formes verbals; o que hi ha una relació general entre les formes actives i les passives dels verbs transitius, que és independent de les formes verbals concretes que s'estiguin considerant. Així doncs, de la mateixa manera que consideràvem natural que hi hagués uns mecanismes generals en el lèxic que permetessin evitar la redundància implícita en una descripció completa de moltes entrades lèxiques (les dels noms comuns, per exemple), convé que ens equipem també amb unes altres operacions que ens permetin de relacionar entrades lèxiques que tenen elements comuns. Aquestes relacions són expressades per les regles lèxiques. En concret a Pollard i Sag (1987) s'exemplifiquen les regles lèxiques per a la variació de formes provocades per la flexió i per a la passiva.

La idea general és la següent. La jerarquia de tipus és completa, de manera que caracteritza totes les classes de paraules. Per exemple, la part de la jerarquia de tipus lèxics corresponent al verb català tindrà en compte (adequadament estructurades) la totalitat de les formes verbals catalanes (en totes les seves variants de temps, persona, conjugació, etc). Ara bé, no tindria sentit haver d'afegir porcions de la jerarquia

específiques per a cada verb. El diccionari català contindrà les formes de base dels verbs, d'una manera molt semblant a com són llistats en un diccionari tradicional, amb la informació específica seva (almenys la fonològica i la semàntica). I seran les regles lèxiques les encarregades de generar les diverses formes flexionades, tot dotant-los de la informació mínima perquè el tipus resultant pugui heretar la informació pertinent a través de la jerarquia. Així per exemple, Pollard i Sag (1987:208) proposen l'estructura de trets de (24) com a entrada lèxica (simplificada) del verb anglès *like*:

(24) específica
$$\begin{bmatrix} \text{fonologia:} & \text{like} \\ \text{sintaxillocallsubcat} & \langle [\ ]\,\mathbf{2}\,,\,[\ ]\,\mathbf{1} \rangle \\ \text{semànticalcontingut} & \begin{bmatrix} \text{relació} & \text{like} \\ \text{arg1} & \mathbf{1} \\ \text{arg2} & \mathbf{2} \end{bmatrix} \end{bmatrix}$$
tipus per a heretar: {base, principal, transitiu-estricte}

Aquesta entrada lèxica conté la informació específica de *like* i la indicació de quins són els tipus de la jerarquia dels quals ha d'heretar la resta de la informació (és un verb estrictament transitiu, principal, i de la forma que la gramàtica anglesa anomena la base). A la vegada, proposen una regla lèxica com la de (25) per a obtenir-ne la forma de la 3a persona del singular del present:

(25)
$$\text{base}\begin{bmatrix} \text{fonologia } \mathbf{1} \\ \text{sintaxillocallsubcat } \mathbf{2} \\ \text{semànticalcontingut } \mathbf{3} \end{bmatrix} \Rightarrow \text{3a-pers-sing}\begin{bmatrix} \text{fonologia} & \text{f3A_SING}\,(\,\mathbf{1}\,) \\ \text{sintaxillocallsubcat } \mathbf{2} \\ \text{semànticalcontingut } \mathbf{3} \end{bmatrix}$$

En aquesta regla, f3A_SING és una funció que pren les formes fonològiques de base dels verbs anglesos i les retorna amb el fonema /s/ al final. Evidentment, la definició de la funció ha de ser més complexa per a tractar totes les variants fonològiques que es poden donar en la 3a persona del singular del present del verb anglès.[13] No obstant, la caracterització general és prou clara: la regla pren una estructura de trets d'un tipus determinat (*base*) i en retorna una d'un altre (*3a-pers-sing*) en la qual ha modificat una part de la informació original (en aquest cas, la fonològica). En ser el resultat de la regla una estructura de trets de tipus diferent de l'originària, hereta a través de la jerarquia d'herència múltiple tota la informació que li correspon com a exemplar del nou tipus. Així, en aquest exemple, el fet que el tipus resultant sigui *3a-pers-sing*, que és un subtipus de *no-passat*, fa que hereti, entre altres, la informació semàntica corresponent a la simultaneïtat temporal amb el moment de parla, com es pot veure a l'entrada lèxica expandida de (20). Així doncs, el conjunt de les entrades lèxiques de què disposa la teoria (com a material d'entrada per a l'aplicació de les regles o principis de la sintaxi i la semàntica) és el resultat del poder descriptiu i generatiu de les regles lèxiques i de les relacions d'herència múltiple entre tipus.

(13) A més, cal tenir en compte que la regla ha estat simplificada, ja que tal com està formulada no pot tractar les excepcions (com, per exemple, que la tercera persona del verb *have* no és *haves*).

Finalment, inclourem, simplificada també, la regla per a la passiva que proposen:

(26)

$$\text{base \& transitiu}\begin{bmatrix}\text{fonologia } \mathbf{1} \\ \text{sintaxi llocal llsubcat} \\ \quad \langle [\,]\,\mathbf{3}\,,\,[\,]\,\mathbf{4}\rangle \\ \text{semàntica llcontingut } \mathbf{5}\end{bmatrix} \Rightarrow \text{passiva}\begin{bmatrix}\text{fonologia fpart_pas}\,(\,\mathbf{1}\,) \\ \text{sintaxi llocal llsubcat} \\ \quad \langle(\text{sp[by]}\ \mathbf{4}),\ldots,[\,]\,\mathbf{3}\rangle \\ \text{semàntica llcontingut } \mathbf{5}\end{bmatrix}$$

Aquesta regla accepta com a entrada estructures de trets de tipus *base* i *transitiu* i dóna com a sortida una nova estructura de trets del tipus *passiva* (del qual hereta la informació de tipus semàntic, per exemple), amb una forma fonològica diferent (la pròpia del participi passiu) i amb un esquema de subcategorització diferent també (amb el complement agent –un sintagma preposicional amb la preposició *by*– enlloc del subjecte).

A HPSG, doncs, el lèxic és plenament estructurat. Amb una simplificació evident del diccionari, es poden produir totes les estructures lèxiques que són necessàries per a un bon funcionament de la gramàtica en conjunt. A més, de manera similar a com ocorria a LFG, gràcies al mecanisme de les regles lèxiques, moltes relacions que en altres models gramaticals són tractades amb transformacions aquí poden ser resoltes de manera directa. En aquesta teoria els principis de projecció de les entrades lèxiques cap a les estructures sintàctiques són molt similars als de les teories anteriors; de fet, HPSG recull el tractament de la projecció de la informació lèxica que es va desenrotllar originàriament a partir de la teoria estàndard, formulat habitualment en el principi de la X-barrada. Com és natural, aquest marc teòric ha contribuït enormement a la revifalla actual dels estudis lexicogràfics, de manera que, en part, els plantejaments generals conflueixen amb els derivats de GB. Una bona mostra d'aquest fet la trobem, per exemple, en un volum com Sag i Szabolcsi (1992), en el qual conviuen autors sorgits de les dues tendències que es plantegem problemes gairebé idèntics, entre els quals destaquen l'estructura argumental i la naturalesa dels rols temàtics, els límits del lèxic, la naturalesa del significat lèxic, i la proposta d'alternatives lèxiques a anàlisis sintàctiques.

## 10. CONCLUSIÓ

Hem fet un llarg recorregut; la lingüística moderna ha fet un llarg recorregut. El paper central que tenia el lèxic en els primers models gramaticals que hem analitzat no ha estat retrobat fins a l'últim model que hem considerat. Ara bé, les diferències són tan grans que difícilment es pot parlar de paral·lelisme: el que llavors eren consideracions informals, sense estructuració, s'ha convertit en un model estrictament formalitzat, que ha estat informatitzat; de fet, aquestes estructures no haurien sorgit si no hi hagués hagut una gran interacció entre la lingüística teòrica i la computacional.[14]

(14) Per a alguns no hi ha tal diferència: almenys parcialment, la lingüística computacional té com una de les seves finalitats el treballar cap a la consecució d'un tractament formalment adequat del llenguatge natural (que és, sens dubte, un dels objectius de la lingüística teòrica general).

Entre un i altre extrem, hi ha hagut un llarg període de sequera teòrica sobre el lèxic. Hi ha hagut tota una sèrie d'anys durant els quals semblava que la tasca lexicogràfica era una tasca menor, en oposició a l'activitat «reina» de la lingüística, la sintaxi. A la lexicografia li quedava la tasca de documentar la diversitat de formes i significats de què està format el lèxic d'una llengua, sense pretendre buscar-hi cap sistematicitat (la teoria lingüística mateixa ja s'encarregava de recordar la futilitat d'una tal recerca: el lèxic és el lloc on es consignen les idiosincràsies de la llengua).

Ara bé, vist des de la perspectiva actual, sembla que podem afirmar sense recança que aquesta mena de pas pel desert no ha estat debades. La sintactització de la lingüística ha comportat una rigorositat que no tenia. En la mesura que s'ha anat veient que les propostes estrictament sintàctiques podien ser millorades en incorporar-hi aspectes semàntics i en proposar-ne un tractament lèxic, la rigorositat que la sintaxi ha comportat no ha estat abandonada: no tindria sentit rebutjar per inadequada una proposta formal, sintàctica, per a adoptar-ne una altra, lèxica o semàntica, però informal; això no s'adiria gens als estàndards a què la lingüística generativa moderna ens ha anat acostumant i que són indispensables per a la lingüística computacional. Així doncs, el retorn al lèxic s'ha produït de la millor manera possible: s'ha simplificat molt el component sintàctic de la gramàtica, sense perdre la generalitat que aquesta havia aportat a la lingüística; les explicacions lingüístiques han guanyat en simplicitat (ja que les relacions lèxiques són en general més simples que les que es poden establir entre estructures sintàctiques); i s'ha retornat a una idea natural des del punt de vista psicològic, que el coneixement que els parlants tenim de la llengua és més un coneixement de les seves peces lèxiques que no pas de relacions entre estructures sintàctiques abstractes.

A la vegada, aquest canvi d'orientació està permetent un replantejament del paper de la semàntica en la lingüística. Un tractament amb estructures de trets implica l'eliminació de la separació dràstica entre la informació sintàctica i la semàntica. Ara bé, una perspectiva com aquesta pot servir també per a millorar el tractament de la semàntica lèxica dins la lingüística. La concepció de les jerarquies de tipus permet d'expressar les relacions semàntiques entre les peces lèxiques d'una manera força flexible, exactament igual que s'hi han pogut expressar les relacions sintàctiques que hem comentat. Com és natural aquest és un dels camps d'investigació actuals, especialment en l'àrea de la lingüística o lexicografia computacionals.

Des d'aquesta perspectiva, doncs, la lexicografia està cridada a esdevenir una branca essencial de la lingüística. A més d'haver-se de preocupar per la determinació de les formes d'organització del lèxic i per la seva estructura en general, molts aspectes morfològics, sintàctics i semàntics entren directament en el seu camp d'estudi, en tant que poden ser objecte de representació lèxica.

TONI BADIA
*Universitat Pompeu Fabra*

## REFERÈNCIES BIBLIOGRÀFIQUES


BÁEZ SAN JOSÉ, V. (1988) *Fundamentos de las gramáticas de dependencias*. Madrid, Síntesis.

BARWISE, J.; PERRY, J. (1983) *Situations and attitudes*. Cambridge, Mass. MIT Press.

BIERWISCH, M. (1966) *Der Strukturalismus. Kursbuch*. [traducció en castellà a Tusquets].

BOTHA, R. P. (1968) *The function of the lexicon in transformational generative grammar*. La Haia, Mouton.

BRESNAN, J. (1982) *The mental representation of grammatical relations*. Cambridge. MIT Press.

CHOMSKY, N. (1957) *Syntactic Structures*. La Haia, Mouton.

— (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass. The MIT Press.

— (1970) «Remarks on Nominalization». A: Jacobs, R.; P. Rosenbaum (eds.) (1970) *Readings in English Transformational Grammar*. Waltham, Mass., Ginn.

— (1981) *Lectures on Government and Binding*. Dordrecht, Foris.

CORBIN, D. (1981) «Le monde étrange des dictionnaires (4): La créativité lexicale, le lexicographe et le linguiste». A: *Le dictionnaire*. Actes du Colloque Franco-Néerlandais. *Lexique* 2. PUF.

DROSTE, F.; JOSEPH, J. (1991) *Linguistic Theory & Grammatical Description*. Amsterdam, Benjamin.

FENSTAD, J. i al. (1987) *Situations, Language and Logic*. Dordrecht, Kluwer.

FILLMORE, Ch. (1971) «Types of lexical information». A: Steinberg, D.; Jakobovits, L. (eds.) (1971) *Semantics*. Cambridge, Cambridge University Press. pàgs. 370-392.

FLICKINGER, D. (1987) *Lexical rules in the hierarchical lexicon*. PhD diss. Stanford University.

GEEST, W. de; JASPERS, D. (1991) «Government and Binding Theory». A: DROSTE, F; JOSEPH, J. (1991).

GRIMSHAW, J. (1990) *Argument structure*. Cambridge, Mass. MIT Press.

JACKENDOFF, R. (1977) *X-bar syntax: A study of phrase structure*. Cambridge, Mass. MIT Press.

LYONS, J. (1977) *Semantics*. 2 vols. Cambridge, Cambridge University Press.

MONTAGUE, R. (1974) *Formal philosophy*. New Haven, Yale University Press.

NEWMEYER, F. (1986) (2a ed.) *Linguistic Theory in America*. Nova York, Academic Press.

PEREIRA, F. C. N.; WARREN, D. H. D. (1980) «Definite clause grammars for language analysis». A: *Artificial Intelligence*, 13, pàgs. 231-78.

POLLARD, C.; SAG, I. (1988) *An information-based approach to syntax and semantics: Volume I. Fundamentals*. CSLI Lecture Notes No. 13. Stanford, Califòrnia.

POLLARD, C.; SAG, I. (en premsa) *Topics in constraint-based syntactic theory.*

SAG, I.; SZABOLCSI, A. (1992) *Lexical matters.* CSLI. Stanford, Califòrnia.

SAUSSURE, F. de (1915) *Cours de linguistique générale.* [traducció en català: Barcelona. Ed. 62. 1990.].

SCIULLO, A.-M. di; WILLIAMS, E. (1987) *On the definition of word.* Cambridge, Mass. MIT Press.

SELLS, P (1985) *Lectures on Contemporary Syntactic Theories.* CSLI.

SHIEBER, S. M. (1986) *An introduction to unification-based approaches to grammar.* CSLI.

SOMMERFELDT, K.-E.; SCHREIBER, H. (1974) *Wörterbuch zur Valenz und Distribution deutscher Adjektive.* Leipzig.

TESNIÈRE, L. (1959) *Éléments de syntaxe structurale.* París, Klincksieck.

TUSÓN, J. (1980) *Teorías sintácticas y análisis sintáctico.* Barcelona, Teide.

WESCOAT , M. T.; ZAENEN, A. (1991) «Lexical Functional Grammar». A: DROSTE, F.; JOSEPH, J. (1991).

WILLIAMS, E. (1981) «Argument structure and morphology». A: *Linguistic Review,* 1, pàgs. 81-114.

ZUBIZARRETA, M.-L. (1986) *Levels of representation in the lexicon and syntax.* Dordrecht, Foris.

LLUÍS PAYRATÓ

# LEXICOGRAFIA I VARIACIÓ FUNCIONAL[1]

## 1. LES MODALITATS DE LA VARIACIÓ I LA VARIACIÓ FUNCIONAL

Una de les formes més conegudes i acceptades per compartimentar la noció de variació lingüística separa les varietats dialectals (o *dialectes*) de les varietats funcionals o estilístiques (o *registres*).[2] Les primeres s'associen amb característiques dels usuaris, les segones amb característiques dels usos (dels àmbits o contextos d'ús).

Entre les varietats dialectals es poden distingir les geogràfiques (o *dialectes* en sentit estricte),[3] les històriques i les socials. Aquestes últimes amaguen en realitat un calaix de sastre amb distincions de diferents menes: varietats culturals, varietats associables amb distincions de classe social, varietats ètniques,[4] varietats «anormals» (en el cas de les pertorbacions lingüístiques), varietats de gènere (parla *masculina* o *femenina*), i varietats corresponents a grups més o menys tancats i marginals (els argots) o, en el pol oposat, a grups més oberts i inclusius (els estàndards).[5]

Pel que fa a les varietats associables amb els usos, trobem primer de tot una terminologia molt (segurament massa) rica: varietats *funcionals*, *contextuals* i *estilístiques*, o bé, més sintèticament, *registres* i *estils*. Donat que *estil* és un terme molt marcat en diferents sentits, sembla que *registre* és més adequat, en la mesura que té un to més neutre i més tècnic. De tota manera, no tothom estaria d'acord a fer sinònims *registre* i *estil*: en algunes formulacions aquestes dues etiquetes recobreixen àrees diferents (sovint reservant *estil* per a la caracterització –descripció de les peculiaritats– d'uns usos personals o propis d'un moviment o període històric). Pel que fa a *funcional*,

(1) La confecció d'aquest treball s'ha beneficiat d'un ajut a la investigació de la DGICYT (projecte PB90-0505). La primera part d'aquest article reprodueix els apartats 3-6 de Payrató (1995), on es fa una exposició una mica més detallada de la variació funcional.

(2) Un dels primers treballs en què es proposa és el de Halliday, McIntosh i Strevens (1964). La proposta s'ha reformulat en diverses ocasions (cf. Halliday (1976) i Gregory i Carroll (1978)), però continua sent, en essència, la mateixa. Cf. també Marí (1983), Lamuela (1986) i Payrató (1988). Segurament el primer treball català en què es parla d'aquestes qüestions és el de Vallverdú (1973), informació que agraeixo a Amadeu Viana. I uns dels últims deuen ser les dues contribucions de Marí (1993, 1994), la primera l'entrada dedicada al terme *registre* en el *Tercer Suplement* (volum 18) de la *Gran Enciclopèdia Catalana*, Barcelona 1993, i la segona el rasseguiment de la història del terme en

pot semblar un terme massa restrictiu (en sentit estricte fa pensar només en les finalitats, i no en d'altres components de les varietats) i pot suggerir un aparellament confusionari amb el terme *formal* (quan en realitat tota mena de variació és formal, de forma). Tanmateix, el terme *contextual* no soluciona cap problema perquè dóna la imatge –encara més esbiaixada– que només unes varietats «tenen» context o en depenen (quan qualsevol varietat, de fet, incloses les dialectals, és necessàriament contextual, està lligada a un context determinat: històric, social, geogràfic i situacional). Utilitzaré, en conclusió, preferentment, els termes de *registre* i de *varietat funcional*, en el sentit ampli de modalitat lingüística *situacional*, associada als contextos d'ús.[6]

## 2. ELS ESTUDIS SOBRE LA VARIACIÓ FUNCIONAL

Per bé que, com és prou conegut, la tradició d'estudis centrada en l'estilística és força rica i antiga, sobretot en el món anglosaxó i en el germànic (cf. Turner 1973 i Spillner 1987), normalment s'ha decantat més cap al vessant estètic i literari (l'estil com a traça personal en l'escriptura) que no cap al funcional i sòcio-cultural (l'estil com a varietat lingüística). En aquest segon vessant la majoria d'estudis, i en especial els més difosos i influents, provenen de la tradició anglosaxona. Les bases s'estableixen al llarg de la dècada dels seixanta, en especial amb el treball de Joos (1962) i amb dos treballs publicats el mateix any: el de Halliday, McIntosh i Strevens (1964), i el d'Enkvist, Spencer i Gregory (1964). Abans, però, Kenyon (1948) ja havia separat clarament el concepte de *varietat funcional* del de *nivell cultural*, barrejats sovint en tractaments que han provocat múltiples confusions i malentesos, i Reid (1956) (veg. Halliday (1976)) havia utilitzat per primera vegada el terme *registre* (i segurament no pas per atzar en relació al bilingüisme, donat que la qüestió de la variació lingüística i de les varietats lingüístiques en particular es fa encara més palesa en aquest cas).

La dècada del setanta s'obre amb l'article de Ferguson (1971), que desvetllarà l'interès pels registres simplificats, i amb els primers estudis contrastius (Ure (1971), Ure i Ellis (1974)); i es tanca amb l'estat de la qüestió sobre els marcadors situacionals de Brown i Fraser (1979). A més, amb les obres de Turner (1973) i Gregory i Carroll (1978), inclou els que es poden considerar els dos primers manuals sobre la variació funcional. La dècada del vuitanta, per últim, presenta ja força treballs de contrastació de registres (al marge de molts d'altres de descripció de varietats particulars). Aquests treballs s'apliquen normalment, com era d'esperar, al cas de l'anglès, on ha continuat desenvolupant-se amb més força aquesta línia d'investigació (cf. Biber (1988), Ghadessy (ed.) (1988) i Young (1990)). Molt pocs estudis similars tenen com a objecte d'altres llengües, si bé es poden trobar nombrosos treballs monogràfics sobre registres aïllats (cf. en el cas del català López del Castillo (1976), Marí (1983, 1987), Lamuela (1986), Payrató (1988, 1992)).

la lingüística catalana, publicada dins el volum 308 d'*Escola Catalana*, el primer monogràfic en català dedicat als registres, on també es poden consultar d'altres treballs sobre diferents aspectes de la variació funcional (cf. Salvador 1994 i Viana 1994) i sobre aplicacions didàctiques.

(3) L'abast del terme *dialecte* en relació al de *varietat* fluctua al llarg dels estudis: o bé s'usa com a sinònim de varietat (amb l'extensió semàntica màxima), o bé com a identificador de varietat dialectal, o bé en el sentit de varietat estrictament geogràfica (amb l'abast més reduït de totes tres accepcions). Una polivalència similar es troba en l'ús del terme *varietat social*.

(4) Per bé que el català no sigui la millor llengua per exemplificar aquesta mena de varietats, conté casos prou clars com el de les varietats catalanes de les comunitats gitanes (per exemple les de Barcelona o de Perpinyà).

(5) Si bé la noció de *varietat estàndard* té components que l'acosten a la de registre, s'acostuma a incloure entre les varietats dialectals. En realitat, no acaba d'encaixar en la distinció entre les varietats dialectals i les varietats funcionals, s'hi situa a mig camí. De fet, encara que la funció de la varietat estàndard és justament neutralitzar la variació per assegurar la intel·ligibilitat d'uns certs usos lingüístics, l'estàndard acaba reproduint al seu si la variació present en la comunitat de parla en què s'utilitza: l'estàndard sol contenir marques tant de la identitat dels usuaris com de les situacions d'ús en què es fa servir.

(6) En realitat, la col·lecció de termes no s'acaba amb els apuntats, i es pot completar amb les denominacions, força usuals, que va popularitzar Coseriu (veg. Spillner 1987) a l'hora de parlar de les varietats tant dialectals com funcionals: varietats *diatòpiques* (geogràfiques), *diafàsiques* (funcionals) i *diastràtiques* (socials),

## 3. LES MODALITATS DE LA VARIACIÓ FUNCIONAL

La subcategorització de la variació funcional es veu abocada ineludiblement a la idea de *continuum*, concepte utilitzat també molt sovint per donar la imatge més aproximada del que representa la variació històrica i generacional (varietats succeint-se ininterrompudament), la social (varietats destriables indefinidament, tot associant-les amb determinats estrats i grups socials) i la geogràfica (dialectes que se separen d'altres dialectes o que s'hi aproximen sense solució de continuïtat).

La situació de *continuum* de parla és tan evident en el cas de la variació funcional (veg. Payrató 1988), que tant els factors que s'han fet servir habitualment en les classificacions com els registres resultants tenen un aire d'arbitrarietat evident, però més per falta de tradició que per altres raons de més pes (l'arbitrarietat no sembla en absolut superior a la que se sol admetre en la delimitació de classes socials, de generacions, d'èpoques històriques o de zones geogràfiques). Els factors distingits tradicionalment (cf. Halliday (1976), Gregory i Carroll (1978), Spillner (1987)) són tres, desglossables en quatre: el camp, el mode i el tenor (funcional i interpersonal). El *camp* representa de què es parla i què es fa: és l'essència i l'objecte de l'acte (i conseqüentment el *tema*). El *mode* es refereix al *canal* de producció, tot incloent la planificació del text. El *tenor funcional* té a veure amb la finalitat o *propòsit* del text, i el *tenor interpersonal*, per últim, fa referència a les relacions establertes entre els interlocutors, i es concreta bàsicament en la *formalitat*. Donat que aquest últim factor es pot separar del precedent, sembla preferible anomenar-lo amb un rètol també diferent (com per exemple el de *to*).[7]

Cap dels quatre factors esmentats resulta fàcil de quantificar o categoritzar. El camp admet un tall fonamental entre caràcter *genèric* i *tècnic*, que és a la base de la distinció entre el registre col·loquial (el llenguatge corrent) i els registres tècnics (o tecnolectes). El mode amaga una gran riquesa de registres darrere el que (només en aparença) sembla una clara dicotomia: *oral / escrit*. Cal completar aquesta distinció amb la consideració del grau de preparació o d'espontaneïtat del procés de producció del text (des de productes recitats de memòria o llegits a d'altres absolutament espontanis o improvisats a partir d'un escrit).[8] Al seu torn, el tenor funcional (o simplement *tenor*) presenta tants tipus com funcions es vulgui reconèixer al llenguatge. En una de les formulacions més simples possibles, es pot distingir un tenor *informatiu* i un d'*interactiu* (que es correspondrien, a grans trets, amb les funcions *ideativa* i *interaccional* apuntades per Halliday). Si recollíssim la formulació clàssica de Jakobson, podríem parlar de tenors lligats a les funcions *referencial*, *conativa*, *expressiva*, *fàtica*, *poètica* i *metalingüística*. Des d'una perspectiva més textual i discursiva, encara, s'ha parlat sovint dels tenors *expositiu / explicatiu*, *argumentatiu*, *descriptiu*, *narratiu*, *instructiu*, *estètic* o *retòric*, *interactiu* i fins i tot, a vegades, de *predictiu* i de *persuasiu* (veg. Castellà (1992)). Per últim, pel que fa al *to* o *tenor interpersonal*, la gradació més senzilla establiria

a les quals cal afegir les (ja anomenades amb més tradició) varietats *diacròniques* (històriques).

(7) De fet, el nombre exacte i els tipus concrets de factors categoritzadors dels registres (que realment són pertinents) formen part, tal com s'exposa més endavant, del que han d'esbrinar els estudis sobre variació funcional. Sense aquests estudis, que aportin una evidència empírica suficient, resulta força arbitrari discutir la preeminència d'uns trets sobre d'altres. Cf. per exemple Hasan (1973) i Bell (1984). La primera autora prefereix distingir cinc factors: tema (*subject matter of discourse*), situació (*situation-type for discourse*), rols dels participants (*participant roles within discourse*), mode (*mode of discourse*) i medi (*medium of discourse*). Bell (1984), en canvi, per la seva part, fonamenta la variació estilística en l'auditori, exclusivament.

(8) Gregory i Carroll (1978) ho exposen ben clarament a través d'un esquema prou conegut i que es pot presentar, en termes bíblics, com l'arbre de l'oralitat i l'escriptura. Ochs (1983) ha analitzat amb deteniment la distinció entre registres planificats i improvisats. Veg. Castellà (1992) pel que fa a distincions suplementàries sobre el conjunt d'aquests factors.

un pol *formal* i un d'*informal*. Una categorització més acurada del *continuum* de la formalitat ja esdevé, però, complexa, i ha estat duta a terme de maneres diverses segons els autors.[9] Molt sovint s'ha recorregut a una gradació triple, corresponent a un ús *col·loquial*, *estàndard* i *solemne* (en el sentit d'*informal*, *neutre* i *formal*, respectivament).

## 4. VARIETATS FUNCIONALS I COMUNITAT DE PARLA

Si la caracterització dels registres es vol dur a terme combinant els quatre factors anteriors, amb els *continua* respectius, és comprensible que el resultat final no sigui altre que una graella (o *continuum* de parla) sobre la qual són admissibles desplaçaments múltiples (veg. Payrató 1988). Creuant, simplement, el camp i el to, amb les mínimes distincions possibles (genèric / tècnic i formal / informal), s'obtenen ja immediatament quatre registres (veg. Hudson (1980)):

(a) Genèric - informal: «Passa'm la sal!»
(b) Genèric - formal: «Passi'm la sal!»
(c) Tècnic - informal: «Passa'm el clorur sòdic!»
(d) Tècnic - formal: «Passi'm el clorur sòdic!»

Si es creuen els quatre factors amb distincions binàries tot subdividint el mode (genèric / tècnic; oral planificat / oral espontani, escrit planificat / escrit espontani; informatiu / interactiu; informal / formal), el resultat és aleshores ja de 32 modalitats, i el *continuum* de parla funcional es fa cada vegada més evident.

Ara bé, com que els registres constitueixen conjunts de tries lingüístiques (i paralingüístiques, de fet) determinades sòcio-culturalment, no totes les possibilitats de la graella sencera es fan efectives en el si d'una comunitat comunicativa (o comunitat de parla). Vist així, és lògic que la recerca sobre les varietats funcionals no solament s'hagi de plantejar descriure modalitats lingüístiques (en termes de freqüència o de presència / absència d'elements i d'estructures) sinó que hagi de descriure pròpiament el repertori funcional d'una comunitat de parla, tot descobrint quins són els registres que realment (i no només a priori) hi funcionen. Això equival, a més, a establir el pes efectiu de cada un dels criteris de subcategorització anteriors en cada comunitat comunicativa i, conseqüentment, en últim terme, a plantejar el valor general –no associat a una comunitat en particular– de cada factor.

(9) Joos (1962), en una de les més conegudes, va utilitzar cinc graus, amb les etiquetes originals de *frozen*, *formal*, *consultative*, *casual* i *intimate* (veg. sobre la seva formulació Viana (1987)). Labov (1972) també va distingir cinc graus, però en el seu cas el resultat té més a veure amb procediments d'elicitació de la informació que no pas amb autèntics registres (els cinc graus corresponen a la lectura de parelles de paraules contraposades, a la lectura de llistes de paraules, a la lectura de text, a la conversa dirigida i a la conversa lliure).

## 5. EL REFLEX DE LES VARIETATS FUNCIONALS EN ELS DICCIONARIS

### 5.1 Els components dels registres i les marques lexicogràfiques

Els diccionaris s'enfronten amb el mateix problema que qualsevol treball lingüístic que descansi sobre la recollida de dades i que no faci abstracció de la variació funcional: ¿Quin ha de ser el registre que forneix l'evidència empírica? ¿Quina és la varietat funcional referencial? En moltes ocasions aquestes preguntes no es responen, i se sobreentén que la teorització es construeix sobre una discutible varietat neutra o sobre una dubtosa llengua «comuna» o no marcada.[10]

El fet és que la diversitat funcional ha estat, en general, molt desatesa pel conjunt dels estudis lingüístics, no només a l'hora de fer-ne un objecte d'anàlisi sinó també a l'hora de considerar la recollida de les dades i les varietats que les proporcionen. Els diccionaris han reflectit, també en general, aquesta situació, almenys fins fa molt poc anys. Si prenem, per exemple, com una primera mostra en la tradició lexicogràfica catalana, l'obra de Fabra (1932), ens adonarem que les marques de context situacional són més aviat pobres i confuses, per bé que no pas gaire més que certes indicacions d'altres menes. En efecte, només que repassem la llista d'abreviatures ens adonarem que, pel que fa a les varietats històriques, coexisteixen les expressions *antigament*, *arcaic* i *obsolet*, i pel que fa a les funcionals, *propi del llenguatge familiar* i *popularment* (i encara, en part, *habitualment / ordinàriament*, d'una banda, i *emprat irònicament / poèticament*, d'una altra). Si hi afegim *vulgarment*, que té a veure en principi amb la variació sòcio-cultural, i la llarguíssima sèrie d'expressions referides al camp tècnic, el panorama es completa pel que fa a aquesta obra en particular.[11]

Tant en el cas de Fabra (1932) com en molts d'altres, per exemple en el *Diccionari de la Llengua Catalana*,[12] les nocions es recobreixen i no són fàcils de delimitar. Per conseqüent, el (suposablement posterior) marcatge de les unitats lèxiques no és gens clar, cosa que es descobreix de seguida encara que sigui fent-ne una repassada només d'un conjunt.[13] Segurament el cas més evident d'aquesta situació es dóna en la parella *familiar / popular*, d'ús molt estès –curiosament i estranyament– en la tradició lexicogràfica. Fins i tot oblidant que *popular* ja té un altre sentit tècnic en lingüística (mot *popular* en oposició a mot *culte*), es fa difícil d'entendre com ha pogut fructificar i arrelar tan fort una distinció com aquesta, amb tan poc fonament, i que molt sovint va flanquejada per les marques addicionals de *vulgarisme* i *(propi del llenguatge) infantil* (aquesta última com a marca prèvia a la definició o com a especificació en el cos mateix de la definició). En principi, *familiar* i *popular* no s'oposen de cap manera com a termes del llenguatge corrent, i no es veu com pot establir-se entre ells una frontera precisa de caire tècnic, de manera que cal buscar d'altres rètols si es vol etiquetar alguna distinció concreta en la línia del que s'ha exposat en els apartats precedents (sobretot pel que fa a la (in)formalitat). Justament el problema de fons que presenten

(10) Aquest fenomen es dóna sovint en el cas del català. Quan es parla, per exemple, de la (in)correcció o (in)adequació d'un terme o d'una construcció, gairebé mai –sobretot tradicionalment– no es fa esment de la varietat funcional implícita, d'aquesta varietat que es fa passar com a neutra, comuna o no marcada, i que sol identificar-se, encara que no es digui, amb un registre genèric o tècnic, escrit, preparat i més aviat formal.

(11) Per ser més concret, aprofito les dades que Bernal i Coma (1993) ofereixen després d'un buidatge del diccionari: *en llenguatge col·loquial* apareix en quatre entrades (*el*, *en*, *hi*, *veure*), *familiarment* és la marca més freqüent (156 casos), *habitualment* apareix en quatre casos (*atmosfera*, *escarabat*, *imbecil*, *tinguda*), *en llenguatge infantívol* en vint, *ordinàriament* només en un (*família*) i *popularment* només en dos (*corea*, *doctor*); *vulgarment*, per últim, és utilitzat en tretze ocasions (*acollonir*, *adinyar*, *burilla*, *clapar*, *collonada*, *collonut -da*, *dinyar-la*, *endinyar*, *llufa*, *nesprer*, *nombre*, *papagai*, *pebrot*).

(12) En aquesta obra, cinquanta anys més tard, la situació no ha canviat gaire, ja que s'hi troben les marques de *col·loquial*, *familiar*, *popular* i *vulgar*, d'una banda, i les de *despectivament / pejoratiu*, d'una altra. Segons la informació que m'ha facilitat l'editorial, la nova edició del DLC, prevista per inicis de 1994, unificarà les marques d'informalitat sota el rètol *col·loquial*. Quant als percentatges, respecte d'un fons de 150.000 fitxes, 157 tenen la marca de *vulgarisme*, 415 la de *col·loquial*, 1.784 la de *despectivament*, 86 la d'*irònicament*, 28 la de *pejorativament*, 33 la de *poèticament* i 115 la de *popular*. En el DLC actual hi ha marques de *col·loquial* només en quatre entrades: *inrevés*, *jungla*, *vaca* i *veure* (agraeixo aquestes informacions a *Enciclopèdia Catalana* i, molt especialment, a Xavier Carrasco).

molts diccionaris és que no ofereixen definicions del que etiqueten, i esdevé gairebé impossible esbrinar els criteris emprats. Així, per exemple, en una de les poques definicions consultables, Castellanos i Castellanos (1979:9) apunten que «L'abréviation *fam* indique un terme ou une prononciation propres au langage familier (habituels dans les conversations). L'abréviation *pop* indique un terme ou une prononciation propres au langage populaire». Paral·lelament, en la mateixa col·lecció de diccionaris bilingües, Batlle i Haensch (1981:48) associen *popular* i *vulgar*, com si l'excés de popularitat portés al vulgarisme: *pop* es refereix a popular i *pop!* a «(massa) popular, vulgar». Prou curiosament (i caòticament) també, per últim, i encara dins la mateixa col·lecció, mentre que el DCASCAT marca *fam* (*familiar*) i *pop* (*popular, vulgar*), el DCATCAS marca *fam* (*familiar*) i *vulg* (*vulgar*).[14]

En definitiva, en relació als factors associables amb el registre vistos en els apartats anteriors, cada un té una incidència diferent des d'un punt de vista lexicogràfic. El *camp* és, sense cap mena de dubte, el factor que més s'ha tingut en compte des d'aquesta perspectiva, amb una –proporcionalment incomparable– quantitat de submarques o subdistincions corresponents sobretot als registres tècnics i científics (i, excepcionalment, a alguna altra varietat associada als registres corrents o col·loquials, com la del llenguatge infantil, tant en el sentit de varietat usada pels infants com de varietat usada pels adults per parlar amb infants).[15] El *mode*, en canvi, no s'hi sol veure representat (ni pel que fa al canal –oral / escrit– ni pel que fa al grau de planificació del text), com tampoc el *tenor funcional* (o simplement *tenor*). El *tenor interpersonal* o *to*, per últim, queda marcat per les etiquetes vistes més amunt (*familiar / popular*). En realitat, donat que hi ha correlacions òbvies entre els quatre factors, les etiquetes lexicogràfiques es poden interpretar sovint anant més enllà del seu significat literal. Així, per exemple, *familiar* i *popular* s'associen amb (possibles etiquetes) com *oral*, *espontani* i *interactiu*, mentre que una etiqueta com *formal*, en canvi, si s'utilitza,[16] s'associarà més aviat amb *escrit* i *planificat* (o com a mínim amb *oral preparat* o *oral llegit*).

### 5.2 Aspectes culturals, enunciatius i textuals

El vessant pragmàtic de la llengua, la dimensió del seu ús efectiu, queda associat necessàriament amb uns àmbits o contextos d'ús que es componen de situacions comunicatives concretes i agrupables en funció de la seva coherència i de les seves similituds. L'anàlisi d'aquests contextos (i el seu reflex en les obres lexicogràfiques) no s'exhaureix amb el que s'ha exposat més amunt. Els quatre components anteriors, sobretot el camp i el to, es poden aprofundir o bé arrodonir amb aspectes complementaris. Així, el camp es pot beneficiar –encara que sigui el genèric– d'informacions de caire cultural sobre l'ús de certs termes (cf. casos com els de *bandera* o *senyera*).[17] El to, per la seva part, es pot arrenglerar amb altres components de l'enunciació, en el sentit de posada en funcionament de la llengua i alhora de reflex, en el producte textual, de

(13) Unes quantes mostres del caos actual, corresponents al DLC: *saltataulells* té la marca de *familiar*, però *saltamarges* o *saltaterrats* no; *pispa* és *familiar* però *pispar* no; *tifa* no té la marca de *familiar* però *fer el tifa* sí; *burro* és *vulgar* en l'accepció de 'ase' però no en les altres (*burra* tampoc, però *matxo* curiosament és *popular*); *xona* és *vulgar* en l'accepció de 'vulva' i *familiar* en la de 'dona bruta'; *tenir bola (a algú)* és *vulgar* (i no *familiar*), mentre que *bola* en el sentit de 'mentida' no és ni *familiar* ni *vulgar*; *estar tocat del bolet* és *familiar*, però *bolet* en el sentit de 'cop' és *popular* (com *cleca*, però no com *bufa*, que només és *popular* com a interjecció; *hòstia* ja és entrat com a *vulgar*); *cagar*, *cagar-la*, *cagarada*, *cagarel·la*, *cagarina*, *cagarro*, *cagat -ada*, *caguera*, *cagarró*, etc. no tenen marca de cap mena, però *cagada* ('errada') i *fer cagarel·la* presenten la de *vulgar*, i *sembla que tingui cagarines* és *familiar*; *mamellam* és *popular*, però *pitam* o *pitram* no; *guipar* és *popular* però *clissar* no; *putada* és *vulgar* però *cabronada* no; *carall* és *vulgar* com a interjecció però no com a substantiu, i el mateix passa amb *collò* i *collons* (si bé *cony* ho és sempre); per últim, en una mena d'espiral absurda, *conill*, com a adjectiu, agrupa la doble marca de *familiar* i *popular*, i *lladre* ('tipus d'endoll') és *popular* dins l'electrotècnia, mentre que termes com *bleda* (adjectiu), *calçasses*, *caparró*, *capsigrany*, *llet* ('semen'), *pixar*, *pixera*, *nyonyo*, *nyap*, *nyigui-nyogui*, *xano-xano* o *xup-xup*, entre molts d'altres, no estan marcats; *ésser cul i merda* és *despectiu* sota l'entrada *cul* però *vulgar* sota l'entrada *merda*, paraula, per cert, que reuneix tots els trumfos: és *vulgar* en el sentit de 'excrement', *familiar* en el de 'brutícia' i 'cosa dolenta' i *popular* en el de 'fatxenda'. No es pot demanar gaire més.

(14) Coromines (1981-1991:

(sobretot) l'emissor, però en realitat de tots els interlocutors. Tanmateix, pocs diccionaris van més enllà de distincions com *irònicament* o *poèticament* (veg. Fabra (1932)), o d'altres com *despectiu / pejoratiu* i *eufemístic* (cf. Batlle i Haensch (1981) i Castellanos i Castellanos (1979)). El LDELC, en canvi, és força més complet en aquesta dimensió, i distingeix, com a etiquetes que mostren actituds (de l'emissor), les cinc següents:

| | |
|---|---|
| «*apprec* | appreciative - shows that the speaker likes or approves of something |
| *derog* | derogatory - shows that the speaker dislikes or disapproves of something |
| *euph* | euphemistic - a polite or indirect word for something unpleasant or embarrassing |
| *humor* | humorous - shows a joking or ironic attitude |
| *pomp* | pompous - shows a foolishly self-important attitude» |

Aquest mateix diccionari és també excepcional, per últim, en la mesura que inclou força més informació que l'habitual en obres lexicogràfiques tradicionals sobre els *cotextos* (o contextos lingüístics) de molts termes. Com és prou sabut, les unitats lingüístiques estableixen relacions mútues en el marc del text. Entre aquestes, la *col·locació*, en el sentit de l'aparició conjunta, regular i freqüent, de paraules, és probablement la més important, en la mesura que permet –sobretot als parlants no natius– combinar amb encert les unitats per formar sintagmes habituals, no (involuntàriament) sorprenents. Si prenem com a exemple les denominacions de color, en català un diccionari podria aportar informacions del tipus que *blau marí* o *blau cel(est)* són combinacions habituals (mentre que no ho són *groc solar* o *gris estrellat*), o que *pujat / -ada* s'aplica predominantment als colors (més que no, posem per cas, als sons o als sentiments).[18]

## 7. CLOENDA

Encara que els estudis sobre els registres no tinguin la tradició que podem trobar en els d'altres varietats, sobretot de les geogràfiques i de les històriques, si els lexicògrafs s'haguessin preocupat més del problema de la variació associada als contextos d'ús i n'haguessin analitzat amb més detall els factors, segurament la gran majoria de diccionaris que es poden consultar avui presentarien un conjunt de marques funcionals (i un consegüent marcatge de les unitats lèxiques) més coherent i entenimentat. La remarca sembla justa tant en general (en el cas de les llengües romàniques i germàniques, si més no) com en el cas particular de la tradició catalana, en què els estudis sobre els registres són proporcionalment ben escassos.

Certament la situació ha canviat en els últims anys, sobretot fora de l'àmbit català, amb l'aparició d'algunes obres amb nombroses i clares indicacions sobre l'estil i l'ús dels termes (cf. LDELC). Tanmateix, el problema de fons que cal continuar plantejant-se és quines han de ser les marques funcionals o estilístiques que ha de contenir el

s.v.) ha criticat també el concepte de *familiar* en favor del de *col·loquial*, malgrat l'anglicisme (veg. Payrató (1988:212)). En d'altres tradicions es troben distincions similars (i també confusions o imprecisions similars). Slaby i Grossmann (1983:XX), per exemple, apunten que «Las traducciones españolas que no vayan acompañadas de indicaciones especiales, corresponden al estilo corriente y castizo [sic]; las que se señalan com <fam> sólo se emplean en el lenguaje familiar. El signo <pop> caracteriza las expresiones populares, <vulg> las que deben considerarse como vulgares o indecentes. Por el contrario, <Lit> significa que la expresión pertenece al estilo culto; <poet> que sólo se usa en poesía o prosa poética».

(15) Aquestes submarques, però, varien considerablement d'un diccionari a un altre, cosa que es pot apreciar fins i tot dins les col·leccions de diccionaris d'una mateixa editorial (cf. els diccionaris bilingües d'Enciclopèdia Catalana, en el cas del català). El problema de fons pel que fa al camp és el de decidir el límit d'aquestes subespecificacions i el pes de cada varietat en el diccionari. En el DLC, per exemple, resulta molt xocant l'aparent desproporció entre termes tècnics i termes propis del llenguatge infantil: mentre que apareixen paraules com *eriòfor*, *exòmfal*, *pumpellyta*, *rinobàtids* i *desoxiribonucleic-polimerasa*, no es recullen *bup-bup* ('gos'), *tete* ('germà gran'), *ring-ring* ('mel·lic') i *(fer un) riu* o *(fer un) pis* ('orinar').

(16) El recent LDELC, per exemple, estableix una gradació, bastant més afinada que la doble tradicional (*familiar/popular*), en tres graus: *informal* («used especially in conversation, in letters between friends, etc.»), *formal* («used especially in official papers, business letters, public speeches, etc.») i *slang* («very informal, rarely used in writing,

diccionari. La solució l'ha de donar, com és obvi, cada diccionari en particular, en funció de les seves característiques i de l'ús que se'n farà, de la utilitat que prestarà als usuaris. És arriscat generalitzar, fent abstracció de si es tracta d'un diccionari monolingüe o bilingüe, si va destinat a algun àmbit en particular (escolar, tècnic), si és de caràcter normatiu o descriptiu, etc. En principi, però, qualsevol obra lexicogràfica hauria de senyalar, com a mínim, tots aquells elements lèxics que un parlant competent comunicativament en la llengua objecte del diccionari reconegués com a clarament *marcats* funcionalment, és a dir, propis exclusivament d'un registre específic o d'un grup coherent de registres definibles a partir dels factors apuntats més amunt (camp, mode, tenor i to). Si es descarta el tenor, tal com també s'ha assenyalat abans, el conjunt de marques podria quedar reduït a *genèric / tècnic*, *oral / escrit* i *formal / informal*, de les quals la primera és eliminable justament perquè es pot considerar l'alternativa no marcada o implícita, mentre que les dues parelles d'opcions restants es podrien agrupar i sintetitzar en l'última: *formal / informal*. La tríade mínima resultant (*tècnic / formal / informal*), que ja no sembla recomanable de simplificar més, tal com és costum en la tradició lexicogràfica catalana,[19] hauria d'anar acompanyada de marques de varietats d'altres menes i de la significació sòcio-cultural del termes: *infantil*, *argot*, *tabú* i *eufemisme*. Finalment, les informacions complementàries es podrien dedicar a continuació a valors culturals específics d'algunes paraules, a l'actitud enunciativa de l'enunciador (*apreciatiu*, *despectiu*, *humorístic* o *irònic*) i al cotext habitual de les unitats.

LLUÍS PAYRATÓ
*Universitat de Barcelona*

## REFERÈNCIES BIBLIOGRÀFIQUES

BATLLE, Ll. C.; HAENSCH, G. (1981) *Diccionari alemany - català*. Barcelona, Fundació Enciclopèdia Catalana.

BELL, A. (1984) «Language style as audience design». A: *Language in Society*, 13, 1984, pàgs. 145-204.

BERNAL, E.; COMA, M. (1993) *Estudi sobre els valoratius sociolingüístics del Diccionari General de la Llengua Catalana de Pompeu Fabra*. Treball de curs de l'assignatura «Lingüística aplicada catalana», sota la supervisió de Mercè Lorente. Dept. de Filologia Catalana, Universitat de Barcelona, 1993.

BIBER, D. (1988) *Variation across speech and writing*. Cambridge, Cambridge University Press.

used especially in the private language of particular social groups»). Per una altra banda, també utilitza l'etiqueta de *tabú* («a very offensive word which should always be avoided»), amb menys connotacions sòcio-culturals negatives que la de *vulgarisme* (cf. també apartat següent).

(17) Per exemple, el ja esmentat LDELC inclou notes culturals sobre l'ús de certs mots en apartats específics dins les entrades dels termes. Tot relacionant camp i tenor, l'última edició d'Smith (1979) ofereix un apèndix sobre llengua i ús que recorda els llibres tradicionals d'ensenyament d'idiomes basats en els diàlegs propis de cada situació comunicativa.

(18) És molt difícil deduir aquestes informacions de la majoria de diccionaris. Er el DLC, per exemple, *blau mari* no apareix, *blau cel* tampoc i *blau celest* és entrat sota *celest* (cf. en canvi els exemples s.v. *verd*). En el cas de *pujat / -ada*, aquest diccionari dóna «Que ateny un alt grau d'intensitat. *Un vermell molt pujat*». Això permet inferir que el terme es pot aplicar als colors (com a mínim al vermell), però no que se sol aplicar exclusivament als colors (cf. en canvi s.v. *viu*). Aquestes qüestions poden semblar nímies o capcioses si pensem en un parlant que consulta un diccionari d'una llengua en la qual és competent comunicativament, però esdevenen importants en molts d'altres casos (fins i tot en el primer supòsit, quan el parlant no té una experiència directa d'un mot o una construcció).

(19) La tradició catalana sempre ha tendit a marcar el tret *informal* (amb les denominacions ja apuntades) i no el *formal*, i alhora més el *vulgarisme* que el *cultisme*. Aquest sistema crea la sensació –força estesa– que certs elements (*formals*, *cultes*) són apropiats en el fons en *qualsevol* àmbit mentre que d'altres (*informals*, *vulgars*) només poden utilitzar-se en els propis de la indicació.

BROWN, P.; FRASER, C. (1979) «Speech as a marker of situation». A: Scherer, K. R.; Giles, H. (eds.) (1979) *Social markers in speech*. Cambridge, Cambridge University Press, pàgs. 33-62.

CASTELLÀ, J. M. (1992) *De la frase al text. Teories de l'ús lingüístic*. Barcelona, Empúries.

CASTELLANOS, C.; CASTELLANOS, R. (1979) *Diccionari català - francès francès - català*. Barcelona, Enciclopèdia Catalana.

COROMINES, J. (1981-1991) *Diccionari etimològic i complementari de la llengua catalana*. Barcelona, Curial.

*Diccionari de la llengua catalana* (1982). Barcelona, Enciclopèdia Catalana (2a. ed., 1983). Esmentat com DLC.

*Diccionari castellà - català* (1985). Barcelona, Enciclopèdia Catalana. Esmentat com DCASCAT.

*Diccionari català - castellà* (1987). Barcelona, Enciclopèdia Catalana. Esmentat com DCATCAS.

ENKVIST, N. E.; SPENCER, J.; GREGORY, M. J. (1964) *Linguistics and style*. Londres, Oxford University Press.

FABRA, P. (1932) *Diccionari general de la llengua catalana*. Barcelona, Edhasa (7a. ed., 1977).

FERGUSON, Ch. A. (1971) «Absence of Copula and the Notion of Simplicity». A Study of Normal Speech, Baby Talk, Foreigner Talk, and Pidgins». A: Hymes, D. (ed.) (1971) *Pidginization and Creolization of Languages*. Cambridge, Cambridge University Press, pàgs. 141-150.

GHADESSY, M. (ed.) (1988) *Registers of written English. Situational factors and linguistic features*. Londres, Pinter.

GREGORY, M.; CARROLL, S. (1978) *Language and Situation. Language Varieties and their Social Contexts*. Londres, Routledge & Kegan Paul. [Traducció en castellà (1986) *Lenguaje y situación. Variedades del lenguaje y sus contextos sociales*. México, Fondo de Cultura Económica].

HALLIDAY, M.A.K. (1976) *Language as Social Semiotics. The Social Interpretation of Language and Meaning*. Londres, Arnold. [Traducció en castellà (1982) *El lenguaje como semiótica social. La interpretación social del lenguaje y del significado*. México, Fondo de Cultura Económica].

HALLIDAY, M.A.K.; MCINTOSH, A.; STREVENS, P. (1964) *The Linguistic Sciences and Language Teaching*. Londres, Longman.

HASAN, R. (1973) «Code, register and social dialect». A: Bernstein, B. (ed.) (1973, II) *Class, codes and control*. Londres, Routledge & Kegan Paul, pàgs. 253-292.

HUDSON, R. A. (1980) *Sociolinguistics*. Cambridge, Cambridge University Press. [Traducció en castellà (1981) *Sociolingüística*. Barcelona, Anagrama].

LABOV, W. (1972) *Sociolinguistic patterns*. Filadèlfia, University of Pennsylvania Press. [Traducció en castellà (1983) *Modelos sociolingüísticos*. Madrid, Cátedra].

JOOS, M. (1962) *The five clocks*. Bloomington, I.J.A.L..

KENYON, J. S. (1948) «Cultural levels and functional varieties of English». A: *College English*, 10, 1948, pàgs. 31-36. Reimprès a: Allen, H. B. (ed.) (1958) *Readings in applied English linguistics*. Nova York, Appleton, pàgs. 215-221.

LAMUELA, X. (1986) «Enquadrament sociolingüístic del llenguatge jurídico- administratiu». A: Duarte, C.; Alamany, R. (eds.) (1986) *Tradició i modernitat en el llenguatge administratiu*. Barcelona, Escola d'Administració Pública de Catalunya, pàgs. 47-63. Reimprès a: Lamuela, X. (1987) *Català, occità, friülà: Llengües subordinades i planificació lingüística*. Barcelona, Quaderns Crema, pàgs. 93-114.

*Longman Dictionary of English Language and Culture* (1992). Harlow, Longman, Burnt Mill. Esmentat com LDELC.

LÓPEZ DEL CASTILLO, Ll. (1976) *Llengua standard i nivells de llenguatge*. Barcelona, Laia.

MARÍ, I. (1983) «Registres i varietats de la llengua». A: *Com*, 3, 1983, pàgs. 27-39.

— (1987) «Varietats i registres en la llengua dels mitjans de comunicació de masses». A: *Actes de les Segones Jornades d'Estudi de la Llengua Normativa*. Barcelona, Publicacions de l'Abadia de Montserrat, 1987, pàgs. 9-30.

— (1993) «Registre». *Gran Enciclopèdia Catalana*. Tercer Suplement, vol. 18, s.v., Barcelona.

— (1994) «La difusió del concepte de registre en la lingüística catalana». *Escola catalana* 308, 24-27.

OCHS, E. (1983) «Planned and unplanned discourse». A: Ochs, E.; Schieffelin, B. (1983) *Acquiring conversational competence*. Londres, Routledge & Kegan Paul, pàgs. 129-157.

PAYRATÓ, Ll. (1988) *Català col·loquial. Aspectes de l'ús corrent de la llengua catalana*. València, Universitat de València, (2a. ed., 1990).

— (1992) «Pragmática y lenguaje cotidiano. Apuntes sobre el catalán coloquial». A: *Filología Románica*, 9, 1992, 143-153.

— (1993) «La variació funcional: els registres». A: Payrató, L. (ed.) (1995) *Oralment. Estudis de variació funcional* (en preparació).

REID, T.B.W. (1956) «Linguistics, structuralism, philology». A: *Archivum Linguisticum*, 8, 1956.

SALVADOR, V. (1994) «Registres i discurs literari». *Escola Catalana* 308, 38-41.

SMITH, C. (1979) *Collins Spanish - English English - Spanish Dictionary*. Collins - Grijalbo, (3a. ed. 1992).

SLABY, R. J.; GROSSMANN, R. (1983) *Diccionario de las lenguas española y alemana. II. Alemán - español*. Barcelona, Herder.

SPILLNER, B. (1987) «Style and register». A: Ammon, U.; Dittmar, N.; Mattheier, J. (eds.) (1987) *Sociolinguistics*. Berlín, Mouton de Gruyter, Berlín, pàgs. 273-285.

TURNER, G. W. (1973) *Stylistics*. Harmondsworth, Penguin.

URE, J. N. (1971) «Lexical density and register differentiation». A: Perren, G. E.; Trim, J. L. M. (eds.) (1971) *Applications of linguistics: Selected papers of the 2nd. International Congress of Applied Linguistics*. Cambridge, Cambridge University Press, pàgs. 443-452.

URE, J.; ELLIS, J. (1974) «El registro en la lingüística descriptiva y en la sociolingüística lingüística». A: Uribe-Villegas, O. (ed.) (1974) *La sociolingüística actual: Algunos de sus problemas, planteaminentos y soluciones*. México, Universidad Autónoma de México, pàgs. 115-164. [Versió anglesa (1977) «Register in descriptive linguistics and linguistic sociology». A: Uribe-Villegas, O. (ed.) (1977) *Issues in sociolinguistics*. La Haia, Mouton, pàgs. 199-243.

VALLVERDÚ, F. (1973) *El fet lingüístic com a fet social. Assaig de lingüística institucional*. Barcelona, Edicions 62, (2a ed., 1977).

VIANA, A. (1987) «Els cinc rellotges de Martin Joos». A: *L'Espill*, 23/24, 1987, pàgs. 43-59.

— (1994) «Sobre els gèneres». *Escola Catalana* 308, 42-44.

YOUNG, L. (1990) *Language as behaviour, language as code. A study of academic English*. Amsterdam, Benjamins.

JOAN J. PUJADAS/DOLORS COMAS

# LEXICOGRAFIA I CULTURA: ANTROPOLOGIA I SISTEMA SOCIAL (NOMS, RENOMS I NOMS DE CASA AL PIRINEU ARAGONÈS)

Els antropònims constitueixen un sistema de denominacions mitjançant els quals s'aconsegueix reduir la diversitat individual a una racionalitat classificatòria socialment significativa. Anomenar, denominar, és quelcom més que una modalitat pràctica per a identificar les persones. Denominar constitueix una operació lògica que funciona en diferents sentits: individualitza els elements d'un conjunt, estableix categories entre les diferents classes d'un sistema i és una forma de fixar significacions. Aquesta idea, que Lévi-Strauss va desenvolupar brillantment ja fa anys en *El pensament salvatge*, obliga a considerar els sistemes de denominació en la seva dimensió cultural, atès que aquesta proporciona els codis que dotaran de significació als termes onomàstics. Aquesta perspectiva és la que adoptarem en aquest article, en què analitzarem el sistema antroponímic a l'Alt Aragó considerant els seus diversos components: noms de pila, cognoms, renoms i noms de casa.

## 1. EL CAMP DE L'ANTROPONÍMIA

L'estudi sistemàtic dels gustos i criteris amb què una societat organitza el sistema apel·latiu i identificador dels seus membres no és interessant només per la seva curiositat filològica i folklòrica, sinó perquè remet a un terreny molt més ampli en

*Caplletra 17* (Tardor 1994), pp. 59-80

què s'hi veuen implicades les relacions i jerarquies familiars, els usos i costums locals, el paper de l'Església i de l'Estat com a modeladors o reguladors de determinades pràctiques i, en definitiva, les transformacions que tot sistema antroponímic va experimentant en el context dels canvis estructurals més amplis de cada societat.

La formació del nom de persona ha sofert al llarg dels segles nombroses modificacions. Entre els grecs consistia en un nom propi, seguit eventualment d'un topònim referit al lloc d'origen de la persona. Entre el romans els patricis tenen un sistema ben sofisticat que es compon de *praenomen*, *nomen gentilicium*, *cognomen* i *agnomen*. En l'època altmedieval s'estén per tot Europa el sistema antroponímic format per nom propi seguit d'un sobrenom referit a l'ofici, tret físic o característica diferenciadora de cada subjecte. Als segles XIII i XIV segons Moreu (1981:18) comença a estendre's per Catalunya i França l'ús de sobrenoms precedits del terme *àlies*, darrera del nom i cognom de cada persona. Les dones en casar-se prenien el cognom del marit, mantenint com a sobrenom el seu cognom de solteres. El cas era a la inversa si era un home el que es casava amb una pubilla. A partir del s. XVIII comença a generalitzar-se a tot Espanya la duplicació cognatícia del cognom, és a dir, l'adopció del primer cognom del pare i de la mare, la qual cosa evita les confusions d'identificació derivades del repertori finit de cognoms (sobretot en poblacions d'elevada endogàmia) i de l'ús repetitiu d'un nombre limitat de noms de pila.

Fins fa poc temps, però, la major part de les zones rurals, pesqueres o de muntanya, amb unitats de població reduïdes, han mantingut l'ús de *sobrenoms, renoms, malnoms i àlies*, a més a més o en substitució parcial dels binomis patronímics cognaticis. Aquest dualisme onomàstic entre un sistema oficial i legal (compost per nom i cognoms) i un altre sistema «popular» (format per nom i renom), ha tendit a ésser interpretat com una manifestació del conservadorisme i tradicionalisme de les poblacions rurals.

Actualment es fa més èmfasi en d'altres interpretacions menys abstractes o estereotipades. En primer lloc, l'ús de sobrenoms és un recurs útil per a diferenciar individus homònims. A Ansó, per exemple, vam trobar setze persones anomenades Antonio López en el període comprés entre 1924 i 1946. Com que pertanyien a dotze cases diferents, cadascun d'ells tenia un nom de casa que permetia distingir-los dels altres. En aquest mateix període, hi ha també a Ansó dotze *Chusé* Gastón, nou *Chusé* Puyó, vuit Miguel Aznárez, vuit *Chusé* Pérez i vuit *Chusé* Añaños. En suma, de les duescentes vuitanta-cinc cases ansotanes del present segle, cent vint-i-tres (és a dir, el 43%) detenien algun dels sis cognoms esmentats.[1] El cas d'Echo és molt similar, ja que un 37% de les cent noranta-vuit cases d'aquest període tenien un dels següents cognoms: Coarasa, Orensanz, Mange, Larripa, Lagraba, López.[2] Moreu (1981: 25) cita el cas del poble del Lloar (Priorat, Catalunya), que al segle passat tenia gairebé un 50% de famílies amb el cognom Sabater.

En segon lloc, la pervivència i l'ús dels sobrenoms manifesta la insistència de cada comunitat en preservar la pròpia identitat social davant d'altres sistemes de classificació

(1) A Ansó hi ha un repertori total de setanta-quatre cognoms. La freqüència d'aparició dels cognoms més comuns, considerant únicament els caps de família, és la següent: López (27), Gastón (25), Pérez (20), Añaños (18), Aznárez (17) i Puyó (16).

(2) A Echo hi ha un repertori total de seixanta-set cognoms. La freqüència d'aparició més comuns, considerant únicament els caps de família, és la següent: Coarasa (22), Orensanz (13), Mange (10), Larripa (10), Lagraba (9) i López (9).

o d'inclusió més amplis. Posar renoms o malnoms forma part d'un discurs local, és una forma específica d'apropiació de la realitat social. Antany els nens d'Echo i Ansó, com els de moltes altres poblacionss altaragoneses, jugaven a identificar els adults del seu poble. Es tractava d'un joc d'endevinalles que es projectava sobre l'espai social de la comunitat i que tenia una funció lúdica de socialització. Els nens aprenien així a classificar els individus com a membres de famílies i de cases, que eren les unitats socials bàsiques.[3]

En tercer lloc, el fet de posar renoms a individus i a grups (o cases) permet introduir un plus d'informació sobre els trets idiosincràtics o els orígens socials, professionals i geogràfics d'ancestres o de persones contemporànies, que afegeixen significació i intensitat semàntica a un camp que acostuma a ser simplement designatiu. Davant l'opacitat significativa de dir-se *Chusefa* López, els ansotans d'inicis de segle podien distingir entre una *Chusefa* de *casa Tuerto Nabarré*, una *Chusefa* de *casa Bastero*, una *Chusefa Jaque*, una *Chusefa Munrial d'os Rios*, etc.

Mentre que el cognom López és un terme denotatiu desproveït de significació específica, els renoms tenen sempre una gran quantitat de connotacions. Per exemple, tenen la virtut d'informar-nos sobre el caràcter més o menys antic de cada unitat (... «se trata de una casa de las de siempre»), sobre la condició sòcio-econòmica de cada grup i, sobretot els malnoms donen peu a tot tipus de disquisicions sobre l'origen i significació de cada denominació. Malnoms com *Nabarré* (Ansó) o *Murillo* (Echo) poden referir-se tant a topònims o gentilicis, que indiquen el lloc d'origen, com a patronímics convertits en noms de casa. *Royo* (Ansó) i *Royé* (Echo) poden ser tant patronímics, com poden tenir el seu origen en el color *royo* «vermell», o en el *rollo* «rodillo para apisonar la era». *Petré* (Ansó) o *Pericón* (Echo) poden ser tant hipocorístics del nom Petra/Pedro com derivats de *petral* «cingla» i de *pericotear* «tafanejar». Però amb tot sempre hi ha termes de significació unívoca, sigui certa o fictícia, com *Broxeta* (Ansó) «bruixa», *Zarzalla* (Ansó) «tartamut», *Baldragas* (Ansó) «desacurat», *Julepe* (Ansó) «Julepe» (joc de cartes), *Miñon* (Echo) «persona baixa», *Pixé* (Echo) (derivat de *pixar*, orinar), *Cabalero* (Echo) «fill no hereu».

Com es pot comprovar, l'ús de sobrenoms és un recurs sociolingüístic que permet donar especificitat, contingut i vida als termes utilitzats per a organitzar el sistema d'identificació d'individus i cases. Moreu (1981:26) parla del malnom com d'una «insurrecció davant l'arbitrarietat del signe», ja que un «malnom 'motivat' respon millor a la nominació». El malnom serveix, doncs, com a crònica o testimoni d'una societat, dels seus personatges, de les seves característiques personals i de les seves circumstàncies socials, així com de la ironia, del sentit crític i, fins i tot, de la crueltat amb què alguns veïns veuen els altres i els designen.

En la bibliografia antropològica consultada no es marca la diferència entre els noms de casa i els renoms individuals. Brandes (1975) i Hoyer (1976) fan servir el terme *nickname* per a referir-se a l'ús de malnoms individuals en dues comunitats de Castella

(3) De Javierregay, població propera a Echo, tenim la descripció del joc «*Panecito dame pan*», que consisteix a repetir la fórmula «panecito dame pan» i la resta de jugadors successivament han de respondre: «No tengo, ves a casa...» D'aquesta forma es van repassant els noms de tot el poble. A Bielsa es jugava al joc de *Las adivinetas* en què cada jugador ha d'endevinar el nom d'una casa del poble, que la resta dels jugadors han pensat. A mesura que va recitant els noms, el guien, dient-li «frío» o «caliente» (Pujadas, 1977).

(Navanogal i Villavieja del Río), mentre que Barrett (1978) utilitza el mateix terme per a referir-se, en canvi, al sistema de noms de casa usat a Benabarre (Osca) i que és molt similar al que nosaltres hem estudiat a d'altres pobles de l'Alt Aragó. Per la seva banda Collomp (1980) usa el terme *sobriquet* per a designar el renom personal i distingir-lo del nom de casa, hereditari i identificador de grups familiars. Altres autors, com Juliano (1984) i Iszaevich (1981), usen el terme *renom* per a referir-se als noms de casa de Barberà de la Conca (Tarragona), fent servir el mateix terme per als malnoms individuals.

Malgrat les seves evidents interconnexions i indubtable analogia en una multiplicitat d'aspectes, ens sembla oportú afirmar que un *sistema de noms de casa*, com el que es pot observar a la zona pirinenca i pre-pirinenca aragonesa (així com a la major part de zones rurals on l'herència és de tipus indivís), constitueix un paradigma antroponímic diferent i, fins i tot, antagònic respecte a un *sistema de renoms individuals*.

També el nom de pila és un element onomàstic carregat de significat. Identifica les persones, individualitzant-les respecte a qui té els mateixos cognoms i renom. Però al mateix temps també categoritza, tal com intentarem demostrar, ja que indica la posició de l'individu dins la família. Per això el nom té una funció comunicativa, informant d'aspectes significatius de caràcter social. A més, el nom de pila té el poder d'evocar determinats personatges, situacions o tradicions culturals, cosa que condiciona la preferència o rebuig de certs noms (Zonabend, 1980). Però l'interès d'estudiar aquesta forma de denominació no es deu als seus aspectes merament psicològics, ni a la simple curiositat. Darrera l'aparent llibertat per a posar nom a un nounat, hi ha unes regles tàcites que constitueixen l'expressió d'una estructura social concreta i que restringeixen a la pràctica el repertori onomàstic del qual s'escull. Fins i tot avui en dia, en què es valora l'originalitat en la tria de noms singulars i exòtics, es produeix un mimetisme pel qual es seleccionen certs noms, imposats per la moda del moment i se'n rebutgen d'altres.

Als països de tradició catòlica, l'acció de l'Església ha tingut històricament una importància decissiva en el repertori dels noms de pila, ja que en el segle XVI s'havien eliminat ja els noms pagans i també els procedents de l'Antic Testament, que als protestants, en canvi, els agradava d'utilitzar. La idea, concretada en el Concili de Trent, era que, mitjançant el nom, l'individu rebia les virtuts morals i psicològiques del seu sant patró, a qui hauria de tractar d'emular. Així es van difondre certs noms místics i devocionals, associats al culte marià o al d'alguns sants.

Resulta impossible explicar per què a cada zona s'imposà una determinada gamma de noms i no una altra. És significatiu, en tot cas, que enfront a l'amplitud del santoral eclesiàstic, el repertori de noms existent a cada poble acostuma a ser relativament reduït i estable, tal com succeeix al Pirineu aragonès fins a dates força recents. A més, uns pocs noms es consagren com a capdavanters per la seva freqüència d'aparició: José (*Chusé*), Antonio, María o Josefa (*Chusefa*) en el cas d'Echo i Ansó. Aquesta

estabilitat s'explica per les regles de transmissió que intervenen en l'elecció del nom, institucionalitzades mitjançant el *sistema de padrinatge*. L'ús consagrava el privilegi del padrí i de la padrina d'imposar el seu propi nom al nounat. D'altra banda, l'elecció dels padrins no és aleatòria, sinó que es busquen entre els parents més pròxims del pare i de la mare del qui és batejat. Així, uns mateixos noms es van transmetent dintre d'una parentela. És una altra forma simbòlica de representar la continuïtat del patrimoni i del llinatge.[4]

## 2. ELS NOMS DE CASA

Un cop argumentada l'existència d'un sistema antroponímic diferent al dels renoms individuals, cal preguntar-se per la lògica social que sustenta la cristal·lització de noms de casa com a base de l'estructura onomàstica de l'Alt Aragó. Mitjançant quins mecanismes la pràctica de posar renoms es converteix en un paradigma tancat, denotatiu, hereditari, col·lectiu i que designa no solament els individus, sinó també les seves propietats? Per què els renoms perden amb el temps el caràcter mordaç i irònic que tenien quan es tractava de veritables malnoms?

Tinguem en compte, en primer lloc, que qualsevol recurs onomàstic creat i utilitzat per un col·lectiu humà constitueix un sistema comunicatiu que comparteix sempre dues dimensions: designativa i significativa. Quan més institucionalitzat i cristal·litzat està un sistema onomàstic, menor poder informatiu i connotatiu té a causa del principi de la probabilitat d'aparició.[5] El sistema dels noms de casa altaragonès comparteix alguns trets d'un sistema de renoms individuals, però els seus objectius i finalitats són totalment diferents, fins al punt de convertir-se en una nomenclatura tan formal i tancada com la del sistema de cognoms, però referida no sols a llinatges, sinó també a unitats de propietat-residència. Passem a veure els principals trets estructurals dels mecanismes que regulen la reproducció i perpetuació del sistema antroponímic.

### *a) Formes de reproducció onomàstica*

La gran estabilitat observada en el manteniment del sistema de noms de casa a Echo i Ansó reposa sustancialment en el principi normatiu de la reproducció íntegra dels patrimonis, associada a la troncalitat familiar. La insistència *èmic* en la preservació del nom s'ha d'entendre com un mecanisme metonímic, en què el nom representa la totalitat de l'univers casal.

El model normatiu de transmissió hereditària i, per tant, de perpetuació del nom de casa, mitjançant el fill primogènit, ha estat i continua essent dominant, encara que en una proporció menor de la que es sol pensar, ja que en el període comprès entre 1900 i 1960 solament s'acompleix en un 60,4 % dels casos a Echo i en un 65 % a Ansó, tal com es pot comprovar a la taula I (cf. Comas d'Argemir, 1980: 40).

(4) Cf. Zonabend (1979), el número especial de la revista *L'Homme* (1980, XX, nº 4), així com el llibre que recull un col·loqui sobre el tema: Dupâquier, J.; Bideau, A. et Ducreux, M.E. (1984).

(5) Cf. Lyons (1971: 85 i ss.).

TAULA I
TRANSMISSIÓ DEL PATRIMONI. ECHO I ANSÓ (1990-1960)

| Tipus d'herència | Echo | % | Ansó | % |
|---|---|---|---|---|
| Fill primogènit | 90 | 49,0 | 101 | 46,6 |
| Fill no primogènit | 21 | 11,4 | 40 | 18,4 |
| Filla | 51 | 27,7 | 68 | 31,3 |
| Altres parents | 12 | 6,5 | 6 | 2,8 |
| No parents | 1 | 0,5 | – | 0,0 |
| Divisa (diferents fills) | 9 | 4,9 | 2 | 0,9 |
| TOTALS | 184 | 100,0 | 217 | 100,0 |

La continuació de l'estirp mitjançant l'herència femenina es situa al voltant del 30 % dels casos estudiats que reponen a aquest model. A diverses cases d'Echo i Ansó l'herència per línia femenina s'ha produït en dos o en tres esglaons generacionals des de començaments de segle.[6] Es perd llavors la continuïtat del cognom, essent el nom de casa el marcador del llinatge. En els casos de crisi reproductiva de les cases s'ha vingut practicant secularment la institució de l'acolliment, que garantia tant la integritat patrimonial, com la perpetuació del nom de casa. En el cas de propietaris solters, o bé casats però sense fills, es feia hereu a algun parent pròxim: germà/na menor casat i amb descendència, o, amb molta més freqüència, algun nebot/da casat, que passava immediatament a residir a la seva nova casa, amb les mateixes funcions i atributs que, en el seu cas, hauria tingut un fill de l'instituent. En el que va de segle hi ha hagut una vintena escassa de casos d'acolliment en els dos pobles estudiats.[7]

*b) Aparició de noves unitats onomàstiques.*

Un dels fenòmens més interessants és el de la generació de nous noms de casa, perquè ens permet posar de manifest els criteris i valors usats pels veïns locals en el procés de creació de nous termes antroponímics. En el transcurs d'aquest segle, l'aparició de noves cases ha estat en general poc freqüent en aquesta zona del Pirineu, i, en conjunt, és més elevada la nòmina de cases tancades i renoms caiguts en desús que la creació de noves cases portadores d'un nom nou. Ha estat comú, en canvi, el desplaçament d'uns renoms per altres com a resultat de la compra d'una casa, per l'emigració de tota una família, o bé per la fusió de dues cases.

Existeix, per tant, un procés que podríem denominar *nomofàgic*, en què uns noms prevalen sobre uns altres i alguns noms nous substitueixen a d'altres vells. La major fluïdesa en l'aparició i desaparició dels sobrenoms s'opera, lògicament, entre les cases de petit patrimoni, per a les quals la reproducció econòmica és o ha estat més problemàtica i entre les quals l'emigració o la fusió de patrimoni ha estat obligada en un

(6) Aquest és el cas de les cases d'Echo *Bisaltico, Bizente, Caberé, Grábiz, Graitón, Poli, Reluchero* i *Ubero de Santa Luzía*, o de les ansotanes *Albeita, Blanquis, Garay, Garzía, Latre, Marconé, Marcos, Notario, Puyó* i *Ximena*.

(7) Això va succeir a Echo amb *casa Blascosanz, Clareta, Isidro o Madrilejo*, en què es van nomenar hereus a nebots dels caps de llinatge. El mateix es donà a Ansó amb *casa Benabides, Jazunto, Marconé, Pablo, Talecón* i *Tía Ramona Abau*. A *casa Don Pascual* d'Echo, es nomenà hereva una cosina, i a casa *Marosa* d'Ansó es nomenà hereu a una persona sense cap relació de parentiu, cosa que constitueix un cas extrem, però previst, d'acolliment.

moment determinat. Veiem alguns exemples de formació de nous noms de casa per la constitució de noves unitats domèstiques per part de consorts no hereus.

A l'últim terç del segle XIX apareix a Echo *casa Ansotana*, com a resultat de la instal·lació de Ramona Ara (nascuda a Ansó el 1859) i Alejandro Coarasa (de *casa Coarasa*) en una finca propietat de *casa Escultor*. Comença a arrelar la denominació «de la ansotana», per no incórrer en una homonímia amb la casa originària del marit, ni amb la denominació antiga, corresponent a un tronc familiar desaparegut. A finals del segle XIX, un fill hereu de *casa Marica*, Pedro Boli Gastón (nascut el 1871), que fou guarda forestal, funda una nova casa, que s'anomenà *Pedro Marica* per fusió del nom de pila i del nom de casa del fundador. Cap al segon terç d'aquest segle apareix *casa Concheta*, que és el terme hipocorístic derivat del nom de Concepción Orensanz (originària de *casa Cristineta* i nascuda el 1904), casada amb Celedonio Coarasa (de *casa Ansotana*, nascut el 1902), que compraren el seu actual habitatge a la gent de *casa La Pastora* poc després de la guerra civil.

A començaments d'aquest segle, Joaquin Mendiara (nascut el 1888), de *casa Mendiara*, compra a Ansó un edifici urbà de *casa Champuyó* i s'hi instal·la amb la seva segona esposa, Manuela Añaños (de *casa Tixinero*, nascuda el 1895). El renom amb què es coneix la casa des de llavors és un derivat del nom de la casa originaria d'ell: *Mendiareta*. El mateix ha succeït amb la fundació de *casa Saletas* per part de Sebastián López Romeo (nascut el 1886), procedent de *casa Salas*. Un altre cas interessant és el de la distinció onomàstica de dues cases ansotanes conegudes per *Munrial*, que derivaven del cognom Monreal que tenien els seus propietaris a mitjan segle passat. Una d'elles tenia alguns membres de la seva família a Amèrica, que van pagar els estudis d'un dels fills de la casa. Això va portar a la gent del poble a identificar-los com *Os ricos*, cosa que va permetre distingir dos renoms de casa, *Munrial dos Ricos* i *Munrial dos Pobres*.

Malgrat aquests exemples, la formació de noves unitats domèstiques no sempre es correpon amb la creació de nous renoms de casa, com comentarem més endavant. De vegades s'utilitza senzillament el nom de la casa d'origen d'un dels membres del grup nou i, per tant, la redenominació de cases genera homonímies amb cases ja existents (com succeeix amb les de *Gallé, Popolín, Ferrero, Chuané o Catarecha* a Echo).

Un altre factor important per l'aparició de nous noms de casa es correspon amb la instal·lació en el poble de persones foranes, que de forma global són identificades antroponímicament bé a partir del seu cognom, bé del seu ofici o activitat. Dins del primer grup podem consignar l'exemple de *casa Cheto*, terme eufònic derivat del cognom del fundador, pare d'Acisclo Echeto, que, provinent de Navarra, es traslladà a Echo a mitjan segle XIX i instal·là una botiga i un cafè al poble. Hi ha altres exemples de cognoms convertits recentment en noms de casa, com succeeix amb *casa Sanclemente*, que prové de Santiago Sanclemente (nascut el 1902), antic secretari de l'Ajuntament; *casa Ibor*, de José Maria Ibort (nascut el 1914), tinent coronel retirat

de l'exèrcit; *casa Borderas*, dels germans Antonio i José Luis Borderas (nascuts el 1929 i el 1930 respectivament), que són pròspers constructors, instal·lats fa tres decenes d'anys a la comunitat. Un últim cas, molt curiós, és el de l'aparició del renom *Galindo*, referit a la casa de Félix Galindo (nascut a Echo el 1932), el cognom del qual desplaça el nom tradicional de la casa dels seus avantpassats, *casa Jenaro*. Cal suposar que la notorietat personal d'aquest home, secretari de l'Ajuntament, influeix en el fet que es produeixi aquesta sobreposició del seu cognom sobre la referència menys personalitzada del nom tradicional que tenia la casa.

Durant l'últim terç del segle XIX, Pascuala Regla (de *casa Lo Pipau*, nascuda el 1852) es casà amb un guàrdia civil foraster i el seu renom particular *La Zibila* s'acabà consagrant com a nom de casa, que arriba fins als nostres dies. El nom de *casa El Herrero* té el seu origen en l'ofici de l'ansotà José Pérez Betés (nascut el 1871), que a finals del segle XIX es casà amb Manuela Yuste de Embún i s'instal·là a Echo. Posteriorment, el seu fill José Pérez Yuste (nascut el 1900) i el seu nét José Pérez Boli (nascut el 1938) mantenen tant l'ofici com el nom de casa. Un altre ansotà, Pascual Galé (nascut el 1912), es va casar a Echo amb María Marraco (de *casa Lorenz*) i hi instal·là una fleca, d'aquí el sobrenom que rep: *La Panadería*. L'ús d'aquest terme, en lloc del tecnònim més usual, *El Panadero*, es deu a l'existència d'una *casa La Panadera* i a l'intent d'evitar l'homonímia. Un altre nom de casa relacionat amb ofici o ocupació es *Maestro Música*, renom amb què popularment és conegut Patricio Castellano (nascut el 1934, empleat municipal i professor de solfeig), d'implantació encara molt recent, però que ja ha cristal·litzat com a nom de casa.

Per a acabar, volem presentar dos casos ansotans de formació de renoms, que acaben convertint-se en noms de casa. Un d'ells, *casa Tanasio*, prové de la derivació eufònica del nom propi del fundador, Atanasio Añaños (nascut el 1840); el nom de pila es transmet als successius hereus de la casa en generacions posteriors i això reforça la motivació que va donar origen al nom de la casa. Un altre cas molt interessant és l'explicació, transmesa per tradició oral, de l'origen de *casa Lobo*, que sembla provenir del malnom personal atribuït a un carrabiner de molt mal caràcter que va viure bastants anys a la casa com a llogater.

Tots els exemples citats ens il·lustren sobre com l'aparició de noves unitats antroponímiques requereix dos factors: la possessió d'una rellevància social per part del cap de família o del seu cònjuge, que ajudi a popularitzar un renom personal que s'anirà imposant a poc a poc al nom vell de la casa on s'habita, i, en segon lloc, les propietats immobles. Ja que tret d'alguna rara excepció, els no propietaris queden exclosos del sistema de noms de casa i són identificats mitjançant el nom de la seva casa d'origen o de la casa en què habiten.[8]

*c) La crisi de reproducció: persistència i desaparició de noms de casa.*

Àdhuc en condicions de reproducció social tan difícils com les que està travessant el Pirineu aragonès en les últimes tres dècades, la tendència a la preservació dels noms

(8) Resulta molt expressiu el cas de deu veïns que habiten a un bloc de pisos existent a Echo, denominat Joan XXIII. Tant els sis grups domèstics formats per immigrants, com els quatre constituïts per gent nascuda al poble són genèricament coneguts com «los de Juan XXIII». No s'ha creat cap nom específic per a cada família.

de casa és extrema. Això no ens ha d'estranyar, ja que darrera d'aquests renoms es troben condensades simbòlicament les històries col·lectives de cada poble. Encara que no és freqüent la venda total del patrimoni per part del emigrants, trobem alguns exemples en què grups domèstics constituïts per no hereus es fan amb casa pròpia per compra a famílies emigrades. Aquest fet no implica, amb tot, la desaparició de l'antic nom de casa, que es manté i s'identifica amb els nous propietaris.[9] Això succeeix també quan les cases són llogades: gairebé sense excepcions es manté la denominació tradicional de l'antiga casa. Sembla com si es generés un mecanisme de reforçament de la identitat local i d'un ordre patrimonial i familiar que els habitants de les comunitats estudiades es resisteixen a abandonar.[10]

Però malgrat l'esforç per a conservar els patrimonis onomàstics locals, són innombrables els casos de noms de casa desapareguts al llarg del present segle. Un bon nombre de pallers, magatzems i cases desocupades que pertanyen a les actuals famílies del poble foren abans de famílies d'escàs patrimoni, que van començar a emigrar d'Echo i Ansó ja a les dècades dels anys vint i trenta, a redós dels canvis sòcio-econòmics de l'època i de les importants migracions transatlàntiques.[11] Altres cases es van quedar buides amb motiu de la guerra civil, especialment a Echo. Moltes altres, sobretot en els anys més recents, van desaparèixer com a tals per manca de descendència, que és una seqüela d'aquesta tendència generalitzada a la solteria que s'observa a totes les poblacions d'alta muntanya.[12]

Tradicionalment, aquells propietaris d'edat avançada que no tenien descendència utilitzaven el recurs de l'acolliment com una forma de donar continuïtat a les seves cases, tal i com hem comentat més amunt. Avui és quasi impossible trobar joves disposats a integrar-se a la disciplina d'una casa aliena per a accedir així a la propietat, ja que durant els darrers anys han existit moltes alternatives per a prosperar econòmicament, sense detriment de la independència i sense sotmetre's a una nova autoritat. Els pocs casos d'acolliment registrats en les últimes dècades representen un canvi dràstic de sentit de la pràctica social. Ara és el propietari el que és «acollit» pels membres d'una altra casa, que se n'ocupen durant la seva vellesa, a canvi d'instituir-los hereus després de la mort. Això suposa, no la continuïtat de les cases, sinó la seva fusió amb les de les persones que en tenen cura per a heretar després.[13] També hi ha casos en què els vells propietaris no recorren a l'acolliment, sinó que testen a favor d'algun familiar que pertany a una altra casa i es dóna així mateix una fusió patrimonial.[14]

La invenció de noms és aquí, com hem vist, l'excepció i no la regla, a diferència del que succeeix en el cas dels renoms o malnoms individuals, en què qualsevol circumstància, per insignificant que sigui, desencadena la capacitat d'invenció onomàstica. La preservació dels antropònims es troba directament lligada a la capacitat de reproducció social i econòmica de les cases. Entre les famílies més pobres, sense propietats, obligades a viure del treball a jornal (i antigament com a pastors i criats de les cases més poderoses), existien grans incidències i capgiraments a cada trànsit generacional. La seva continuïtat en un mateix poble no era fàcil, ja que la seva supervivència

(9) Això succeeix a cases com *Achés, Baretoné, Burré, Camilo, Casero, Catibiela, Cazo, Culera, Cullito, Chil, Larraza, Manolero, Marcón, Morera, Negro, O Sastre Fraché, Rosa* i *Sabineta* d'Ansó, i també a *casa Corrucato, Francha, Lurbés, Marcos* i *Maurizia* d'Echo.

(10) Alguns exemples de cases llogades a Echo que mantenen el seu antic nom són *Abilia, Blanco, Caballero, Cardesa, Lo Calbo, Lo Miñón, Marina, Marraco, Migosancho, Piché, Roque, Belzuz, Corradasa, Del Sol, La Fraila, La Zibilia, Larraspa, Lo Bastero Barré, Lobera, L'Onset, Lorenza* i *Martes*. En el cas d'Ansó, tenim els exemples de casa *Baldragas, Berdadero, Casero, Collao, Chanmiguel, Molino, Maroto, Ospital, Peña, Piero, Ramoné, Salomón, Sarto, Terrén, Zarzalla* i *Zirilo*.

(11) Entre 1900 i 1960, el nombre de xesos i ansotans emigrats a Amèrica fou d'unes cent cinquanta persones, xifra similar a la dels que van marxar a viure a França i a Catalunya, sobretot després de la guerra civil (cf. Comas d'Argemir i Pujadas, 1985:72).

(12) En els cinquanta anys compresos entre 1924 i 1974, es van tancar setze cases a Echo i trenta-una a Ansó perquè l'hereu no s'havia casat. Avui, hi ha quaranta-tres i cinquanta-sis cases respectivament en què l'hereu és solter i té més de quaranta anys.

(13) A Ansó, *casa Maroseta* es va incorporar a *casa Taconera*, i *casa Melera* a *casa Tolún*. A Echo, *casa Lobera, Lo Cleigo* i *La Casera* s'han incorporat respectivament a *ca. a Zeferino, Malaño* i *Royé*.

(14) Veiem alguns exemples de cases que han passat per herència a d'altres, contribuint per tant a augmentar el seu patrimoni:

*Xanca* ... *Chunillas* (Ansó)
*Madrilejo* ... *Gratón* (Echo)
*Tierno* ... *Fonz* (Ansó)
*Lo Coixo* ... *Ubero* (Echo)

depenia de tenir feina assalariada i els canvis de residència eren freqüents. La memòria del pas pel poble d'algunes d'aquestes unitats familiars de condició humil s'ha perdut irremissiblement, junt amb les poques traces materials que van deixar darrera seu. No succeeix el mateix amb els grans patrimonis, el de les «grans cases». L'absentisme de la família *Changrosín* a Echo o de *Serrano* a Ansó no poden esborrar el record de la seva grandesa, que està sustentada en nombrosos testimonis materials: les seves cases, els seus camps, els seus pallers, les seves eres i, sobretot, la llegenda oral que acompanya la seva memòria.

## 3. ELS NOMS DE PERSONA

Al Pirineu aragonès el nom de pila es subordina als imperatius de la casa. A cada família hi havia una gamma reduïda de noms, que es repetien d'una generació a la següent, constituint fins i tot en alguns casos un signe emblemàtic del llinatge. A Ansó, el fet de dir-se Hermenegildo o *Chusé Ramón*, per exemple, estava estretament associat a la pertinença a certes *cases*, que qualsevol ansotà identificava fàcilment. Posar un determinat nom a un nounat marcava la seva forma d'inclusió a la família i a la comunitat, ritualitzada mitjançant la cerimònia de bateig.[15] En l'annex 1 presentem els noms de pila més corrents a Echo i a Ansó. A partir d'aquestes dades podem sintetitzar els principals trets de l'estructura i evolució del sistema onomàstic:

a) A començaments de segle, el corpus dels noms de pila era força limitat. Existia el predomini gairebé absolut dels noms simples; tant a Echo com a Ansó coincidien els tres noms més freqüents: Maria, *Chusé* i Antonio. La restricció del corpus onomàstic s'accentua si considerem que molts noms apareixen en la seva doble versió masculina i femenina. Si sumem, per exemple, els José (*Chusé*), les Josefa (*Chusefa*), més els noms compostos iniciats per *Chusé*, tots ells amb una elevada freqüència d'aparició, obtindrem un percentatge força elevat de persones que tenen un mateix nom. No és estrany que les celebracions onomàstiques fossin veritables esdeveniments familiars i també comunitaris. Esmentem com a detall curiós que a començaments de segle s'introduïren dos noms femenins, Carmen i Pilar, que van aconseguir gran acceptació. Abans de 1921 no hi havia a Ansó cap dona que es digués Carmen i l'aparició del nom Pilar la constatem cap a 1910. A Echo ambdós noms s'introdueixen abans, i ja des de 1920 es convertiran en els noms femenins majoritaris, arribant fins i tot a superar el que sempre havia estat el capdavanter dels noms femenins, Maria.

b) Els noms dobles o compostos comencen a generalitzar-se a partir de la dècada dels anys quaranta; abans eren molt minoritaris.[16] En el cas dels homes, es formen a partir de la presència d'un dels noms dominants (*Chusé* sobretot, però també *Chuan*, Luis i Miguel), al qual s'afegeix un altre nom d'ús comú. Comencen així a proliferar

*Berdoy* ... *Marconé* (Ansó)
*Lo Pipau* ... *Blascosanz* (Echo)
*Seta* ... *Ferreré* (Ansó)
*Manolón* ... *Pazenzia* (Echo)
*Coles* ... *Gollins* (Ansó).

(15) Observem que si un nen moria abans d'ésser batejat, no se li posava nom.

(16) A Ansó, exclusivament, hem pogut ja trobar alguns noms compostos masculins a començaments de segle. Constituïts sempre per Joan com a primer component, apareixen normalment en famílies de bona posició econòmica. Són els següents: Juan Ramón, Juan Antonio, Juan José i Juan Blas.

els *Chusé María*, *Chusé Luis*, *Chuan Carlos* o *Miguel Anchel*. En el cas de les dones, María passa a precedir una àmplia varietat d'altres noms (Mª Pilar, Mª Teresa, Mª Luisa, etc.). Fins a tal punt es generalitza aquesta pràctica que, en realitat, i com succeeix a molts altres llocs, el nom María funciona com *additiu*, quasi imperceptible davant el que serà el nom principal i recurs obligat quan s'utilitzen noms de format masculí (*Mª Chesús, Mª Chusé*) o bé monosil·làbics (Mª Cruz, Mª Luz, Mª O, Mª Mar). No es aquest el cas, en canvi, dels noms dobles masculins, veritables *noms compostos*, la gamma dels quals és més limitada i on el primer es considera com a principal. Quan María es generalitza com additiu deixa d'imposar-se com a nom simple i desapareix com a tal del repertori onomàstic. El mateix succeix, encara que de forma menys dràstica, amb els noms integrants de noms compostos.

c) En les dues últimes dècades, els noms que havien estat més comuns deixen d'imposar-se de forma gairebé absoluta. Fa més de quaranta anys que cap nou nat ha estat batejat amb el nom de Domingo, Mariano o Sebastiana, i en fa uns vint-i-cinc que ningú a Echo ni a Ansó posa el nom de María, Miguel, Pedro, Francisca, Joaquina (*Chuaquina*), Pascual, Juan (*Chuan*), Manuela, Orosia, Miguela o Santiago. En canvi, s'han anat introduint altres noms nous, sense cap arrelament local previ (Eva, Olga, Sandra, Rubén, César, Víctor Manuel...) En aquest fet ha influït sense cap dubte l'acció dels mitjans de comunicació i de les modes. És també un marcador del nou esperit individualista, que busca singularitzar i trenca així amb l'antiga cadena de transmissió onomàstica que es succeïa a cada *casa*. Aquest fenomen concideix amb el rebuig dels punts de referència tradicionals, juntament amb la crisi de l'antic sistema agropastoral, amb el canvi dels valors associats a la institució de la *casa* i amb l'emigració massiva cap a la ciutat.

d) A Ansó existeix una concentració de noms més gran que a Echo. Resulta lògic que les innovacions onomàstiques tinguin una menor incidència sobre la comunitat ansotana, amb un notable envelliment de la població i un escàs dinamisme demogràfic a l'actualitat.

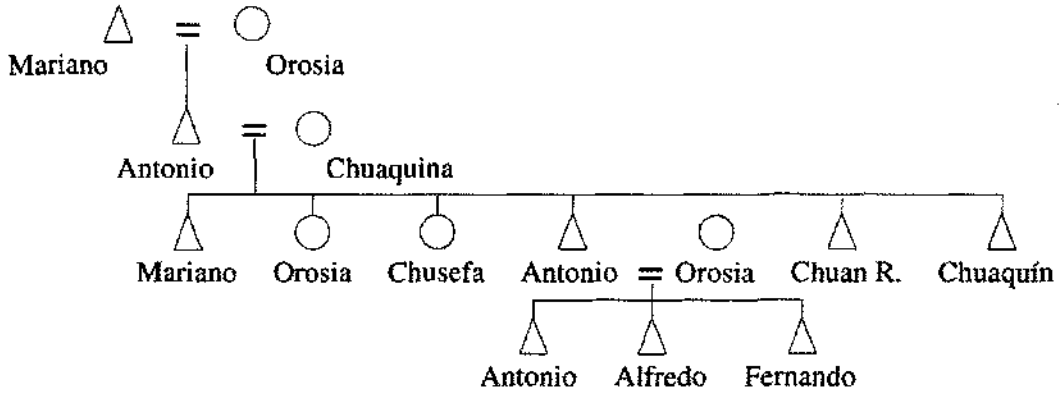
e) El repertori onomàstic és més reduït entre els homes que entre les dones. Sembla com si les normes de transmissió del nom fossin per elles més flexibles, donant lloc a un major nombre de noms singulars buscats per la seva bellesa estètica o pel seu poder evocador. Les causes cal buscar-les en la lògica social que governa la institució de la casa i la comunitat local, cosa que ens remet a l'anàlisi de les formes de transmissió que condicionen els repertoris onomàstics.

Quan naixia un nen no eren els seus pares els qui decidien el nom que hauria de dur, sinó els padrins que ells havien escollit. Ser padrí o padrina no era quelcom que es pogués eludir, ja que no es demanava a ningú si volia ser-ho: se'l nomenava (*se le feba*, segons l'expressió local). I tothom ja sabia quan li correspondria aquest rol, perquè

era quelcom previsible. L'elecció no es realitzava a l'atzar: el padrí i la padrina es buscaven entre els parents més propers, seguint un determinat ordre. Es donava prioritat als avis/àvies; a mida que naixien altres fills, s'anava fent padrins els diferents germans/es dels pares, per arribar després, i segons les circumstàncies, als germans/es més grans del nounat o als parents afins. Encara que l'ordre podia alterar-se en favor d'algun parent prestigiós o de bona posició econòmica.

Els padrins tenien la potestat d'imposar el seu propi nom al nen o nena. Encara que la branca paterna i materna estaven sempre representades en cadascun dels padrins, en l'atribució del nom es donava preferència als membres de la *casa*, i es seguia una troncalitat semblant a la transmissió dels béns. Era gairebé segur que el primer nat rebria el nom de l'avi de la casa, si era un nen, o de l'àvia, si era una nena. Els avis de l'altre llinatge solament imposarien el seu nom si el nen nat en segon lloc era del mateix sexe que el primer. A partir del tercer naixement actuaven ja com a padrins els germans dels pares, que també podien atorgar el seu propi nom, i així successivament.

Tal com assenyala Fine (1980), existien així dos mecanismes de transmissió onomàstica, un vertical, d'avis a néts, i un altre lateral, d'oncles a nebots. El primer afectava als hereus, el segon als no hereus; el primer era més rígid, el segon molt més flexible. Efectivament, el nen que es suposava que seria l'hereu rebia quasi invariablement el nom de l'avi de la casa on naixia, fins i tot si l'avi havia mort i era padrí un altre parent. Per als no hereus, en canvi, hi havia una més gran obertura a la variació: els padrins procuraven satisfer els gustos dels pares, escollint el nom d'ells, o introduint noms de moda. I si un nen naixia després de la mort d'un germà/na, freqüentment l'anomenaven igual que el difunt. Vegem un exemple:




Podem comprovar que per als hereus s'acompleix la transmissió alterna dels noms, dels avis a néts, tot seguint fidelment la línia patrimonial: com la política d'aliances

va conduir el primogènit de la tercera generació a casar-se en una altra casa, fou nomenat hereu el germà següent que, lògicament, no portava el nom del seu avi; però l'alternança es recupera mitjançant el fill gran, fins i tot a costa de l'homonímia que es va produir a la casa. A la tercera generació es pot analitzar millor el conjunt de mecanismes de transmissió onomàstica. El primogènit, que en néixer era considerat com el futur hereu, va rebre el nom del seu avi patern. La germana que el va seguir es va dir com la seva àvia paterna i la següent com la seva tieta i com l'àvia materna. Antonio tenia el mateix nom que el seu pare, i els dos últims germans es van dir com els seus oncles materns.

Mitjançant aquest exemple es pot apreciar com els mecanismes de transmissió onomàstica constitueixen una expressió de la lògica social que subjau al sistema de transmissió de béns. Els noms *classifiquen* els membres d'un mateix grup de germans, *identifiquen* el futur hereu, *simbolitzen* la continuïtat patrimonial de la casa. L'atribució del nom a un nounat no és, doncs, quelcom neutre, és un acte carregat de *significat*, és un missatge fàcilment descodificable pels qui es troben immergits en el sistema social. És, doncs, una forma de *comunicació*. I en termes de sistema de comunicació podem afirmar que la transmissió del nom d'avi a nét constitueix un senyal, o un indici, que té la seva equivalència en el sistema de transmissió de béns. Es rep el nom de forma exclusiva i troncal, de la mateixa forma que s'hereta el patrimoni de forma exclusiva i troncal. És una espècie de demostració davant de tots, i especialment dels no hereus, del dret indiscutible de l'hereu sobre la casa i el seu patrimoni (Vernier, 1980:65). Referint-se al Pirineu francès, Fine indica: «Donant el seu nom al nét, l'avi marca simbòlicament i materialment a qui es transmetran els béns i qui serà més tard el cap» (1980:120).

Una altra dimensió del missatge transmès pel sistema onomàstic és de caràcter simbòlic. El pas del nom d'avi a nét simbolitza la continuïtat de la casa i del llinatge, com una cadena sense fi. El nom té aquí el seu màxim poder evocador, ja que ressuscita els avantpassats i està carregat de valors morals i al·legòrics per l'hereu, que ha de continuar el treball d'aquells que foren els seus homònims. El nom és un bé simbòlic que s'afegeix als béns materials que rebrà l'hereu.

Però, a més a més, el simbolisme de la continuïtat del llinatge es reforça pel nom que s'atribueix als no hereus. Si hi ha hagut una bona política matrimonial, els parents que van actuar com a padrins i padrines hauran entrat en altres cases i no cal oblidar que el prestigi social i el poder local d'una casa depenia del seu èxit per a aconseguir una densa gamma de relacions que l'aliança matrimonial reforçava. La identificació i implicació dels no hereus amb els interessos de la casa natal s'expressa justament en la selecció de nom pels seus fillols, procurant assignar els prototípics del llinatge. Per això hem observat que la major concentració onomàstica es produeix entre les famílies més riques i millor relacionades socialment, ja que per elles el poder evocador del nom com a símbol assoleix la seva màxima cota. A les cases més pobres apareix,

en canvi, una major variabilitat en la gamma de noms escollits, precisament per la mateixa raó que acabem d'apuntar, però aplicada de forma diferent, ja que els referents no són tan clars, ni estan carregats del valor de demostració que assumeixen entre els rics.

A començaments de segle, la institució de la casa tenia un dels seus fonaments bàsics en la valoració de la continuïtat. Aquest valor forma part de l'*ethos* relacionat amb la protecció de la integritat patrimonial, materializat en la institució de l'hereu únic i simbolitzat per la transmissió del nom. La continuïtat s'expressa per la pervivència del passat en el present, pel valor absolut de l'antic i, per tant, a nivell de relacions domèstiques, per la supremacia dels vells. Però quan és la innovació o la «modernització» el que es valora com a positiu els termes s'inverteixen: agraden les novetats, es menysprea el que és antic, prevalen els joves, menys conservadors i millor connectats amb les modes de cada moment. I quan aquest fenomen es produeix al Pirineu, es reflecteix també en el sistema onomàstic, que s'altera per la introducció de nous noms, trencant-se la cadena de transmissió que hem mostrat. Però això no es produeix sense transició.

Efectivament, a partir de la dècada dels anys quaranta comença a sorgir un nou fenomen, i és la tendència a anomenar els nens com els seus pares i no com els seus avis. Sembla un signe clar del paper protagonista que desitgen tenir els cònjuges més joves, una reafirmació simbòlica de la seva posició de poder davant dels ancians. En ocasions, en canvi, sembla que es busca deliberadament un arranjament a gust de tots, cosa que s'aconsegueix mitjançant els noms compostos, que tant d'èxit tindran en aquests anys (és el cas dels nens que reben dos noms, el del pare i el del padrí, per exemple). Fins que irrompen els nous noms inspirats en cantants i actors cinematogràfics, els bíblics i els que agraden per la seva fonètica. Això representa ja una ruptura total, una demostració que avui la vida al Pirineu s'associa a valors diferents als que van governar la casa durant un procés secular, com el que acabem de descriure.

## 4. CONCLUSIONS

El cabal onomàstic forma part del *patrimoni cultural*, veritable capital simbòlic que representa metonímicament la història local, la memòria de la qual contribueix al manteniment de les arrels socials i de la identitat. La memòria col·lectiva no elimina del sistema comunicatiu onomàstic les unitats corresponents a famílies desaparegudes de la vida local: tant a Echo com a Ansó bona part de la toponímia menor està impregnada d'antics noms de casa, cosa que serveix de mnemotècnia per a no esborrar de la seva esfera cognitiva el record d'un passat encara present.[17]

Aquesta resistència a la desaparició dels noms de casa és una marca del paper essencial que aquesta institució ha jugat, i juga encara, com a entitat complexa, que con-

(17) Cf. Pujadas, 1976: 149 i ss.

densa tant la dimensió econòmica com les relacions socials dels seus membres. Avui, bona part de la vigència i significació sòcio-econòmica de les unitats domèstiques està en un procés de decadència irreversible, davant del qual poc poden fer les persones per a aturar-la. Mantenir els noms significa preservar d'alguna manera la memòria col·lectiva dels absents, tant dels que van morir com dels que van marxar. En aquesta persistència onomàstica hi ha molts trets comuns amb la fidelitat lingüística de la qual fan gala aquests pobles, i molt especialment Echo. En ambdós casos destaca una actitud de resistència davant la desaparició total d'uns trets d'identitat específics. Si ens aturem, pel contrari, en l'altra dimensió antroponímica, la dels noms de pila, comprovarem com el vell ordre onomàstic s'ha trencat i ha estat suplantat per criteris d'individualització i per un gust estètic a la moda, que ens mostra tant la crisi de la institució del padrinatge, com la crisi de l'orientació gerontocràtica del poder domèstic.

Aquesta petita incursió en la història i en l'etnografia recent dels antropònims xesos i ansotans ens exemplifica un procés de tipus més o menys general, segons el qual la crisi social i les transformacions culturals operen sempre dintre de contextos més generals. El Pirineu no és ja una terra de boscos i pastures, llunyana aprovisionadora de carn, fusta i llenya per als habitants urbans. Avui, moltes comunitats pirinenques, sense deixar de ser el que eren en el passat, s'han convertit en centres turístics i, com assenyala Gaston (1979:933), han esdevingut una espècie de museus etnològics. Són una representació emblemàtica d'unes formes de vida i d'un bagatge cultural que avui, superada l'etapa del desenvolupament urbà dels anys seixanta i setanta, són valorades de forma creixent. Els nostres veïns xesos i ansotans van canviant cada cop més les seves formes de vida, però el seu capital és doble, ja que ells mateixos, les seves cases, les seves persones, els seus records, fins i tot el seu nom i el de les seves cases, són un nou objecte d'atenció d'aquella societat més àmplia que els envolta.

JOAN J. PUJADAS MUÑOZ
DOLORS COMAS D'ARGEMIR
*Universitat Rovira i Virgili*

## REFERÈNCIES BIBLIOGRÀFIQUES

BARRETT, R. A. (1978) «Village modernization and changing nicknaming practices in Northern Spain». A: *Journal of Anthropological Research*, 34(1), 1978, pàgs. 92-108.

BRANDES, S. (1975) «The structural and demographic implications of nicknames in Navanogal, Spain». A: *American Ethnologist*, 2, 1975, pàgs. 139-148.

BURGUIÈRE, A. (1980) «Un nom pour soi. Le choix du nom de baptême en France sous l'Ancien Régime (XVIe-XVIIIe siècles)». A: *L'Homme*, XX (4), 1980, pàgs. 25-42.

COLLOMP, A. (1980) «Le nom gardé. La dénomination personnelle en Haute-Provence aux XVIIe et XVIIIe siècles». A: *L'Homme*, XX(4), 1980, pàgs. 43-61.

COMAS, D. (1980) «Sistema d'herència i estratificació social: les estratègies hereditàries al Pirineu aragonès». A: *Quaderns de l'Institut Català d'Antropologia*, 2, 1980, pàgs. 25-55.

— (1983) «Ganaderos, boyeros, pastores, obreros... Estrategias económicas en el Pirineo de Aragón». A: *Temas de Antropología aragonesa*, 1, 1983, pàgs. 63-83.

COMAS, D.; PUJADAS, J. J. (1985) *Aladradas y güellas. Trabajo, sociedad y cultura en el Pirineo aragonés*. Barcelona, Anthropos.

DUPÂQUIER, J. (1981) «Naming practices, godparenthood, and kinship in the Vexin, 1540-1900». A: *Journal of Family History*, 6(2), 1981, pàgs. 135-155.

DUPÂQUIER, J.; BIDEAU, A.; DUCREUX, M. E. (comps.) (1984) *Le prénom. Mode et histoire*. Paris, EHESS.

FINE, A. (1984) «Transmission des prénoms et parenté en Pays de Sault, 1740-1940». A Dupâquier et alii (comps.) *Le prénom. Mode et histoire*, 1984, pàgs. 109-126.

GASTÓN, E. (1979) «La cultura popular aragonesa». A: *Estado actual de los estudios sobre Aragón*, I(2), 1979, pàgs. 921-941.

GRACIA, L. (1978) *Juegos tradicionales aragoneses*. Zaragoza, Librería General.

HOYER, E. (1976) «Nicknames in Northern Spain». A: *Folk Dansk Etnografisk Tidsskrift*, 18, 1976, pàgs. 103-111.

ISZAEVICH, A. (1981) «Una mostra del sistema tradicional del moteig a Catalunya: els renoms de Barberà de la Conca». A: *Aplec de Treballs*, 3, 1981, pàgs. 5-17. Montblanc, Centre d'Estudis de la Conca de Barberà.

JULIANO, D. (1984) «Apellidos y 'renoms': dos lógicas de transmisión de la identidad». A: *Comentaris d'Antropologia Cultural*, 6, 1984, pàgs. 71-83.

LÉVI-STRAUSS, Cl. (1964) *El pensamiento salvaje*. México: FCE.

LYONS, J. (1971) *Introducción en la lingüística teórica*. Barcelona, Teide.

MOREU-REY, E. (1981) *Renoms, motius, malnoms i noms de casa*. Barcelona, Millà.

PUJADAS, J. J. (1976) «Ensayo etnolingüístico: el campo léxico de las denominaciones orográficas». A: *Ethnica*, 11, 1976, pàgs. 149-177.

— (1977) *Aspectos etnolingüísticos del Alto Aragón: etnociencia y taxonomías folk*. Tesi de doctorat inèdita (Universitat de Barcelona).

VERNIER, B. (1980) «La circulation des biens, de la main d'oeuvre et des prénoms à Karpathos: du bon usage des parents et de la parenté». A: *Actes de la Recherche en Sciences Sociales*, 31, 1980, pàgs. 63-92.

ZONABEND, F.(1979) «Jeux de noms. Les noms de personne à Minot». A: *Études Rurales*, 74, 1979, pàgs. 51-86.

— (1980) «Le nom de personne». A: *L'Homme*, 20(4), 1980, pàgs. 7-23.

ANNEX 1
EVOLUCIÓ DELS NOMS DE PILA MÉS CORRENTS. ECHO I ANSÓ
(persones nascudes entre 1861 i 1980)

| | PERÍODES | | | | | | |
|---|---|---|---|---|---|---|---|
| **NOMS** | **1861-80** | **1881-1900** | **1901-20** | **1921-40** | **1941-60** | **1961-80** | **OAL** |
| José | 23 | 29 | 50 | 33 | 12 | 3 | 150 |
| María | 30 | 40 | 49 | 21 | 3 | - | 143 |
| Antonio | 16 | 26 | 35 | 32 | 14 | 1 | 124 |
| Josefa (*Chusefa*) | 17 | 19 | 29 | 15 | 8 | 2 | 90 |
| Francisco | 17 | 17 | 27 | 14 | 3 | 2 | 80 |
| José + altres | - | 2 | - | 12 | 35 | 22 | 71 |
| Pilar | 1 | 6 | 19 | 18 | 20 | 6 | 70 |
| Carmen | 1 | 5 | 4 | 28 | 22 | 8 | 68 |
| Antonia | 7 | 12 | 20 | 20 | 5 | 2 | 66 |
| Miguel | 8 | 16 | 21 | 13 | 5 | - | 63 |
| Pedro | 12 | 15 | 18 | 8 | 5 | - | 58 |
| Joaquín (*Chuaquín*) | 7 | 9 | 26 | 10 | 3 | 1 | 56 |
| Francisca | 5 | 20 | 14 | 7 | 2 | - | 48 |
| Domingo | 11 | 10 | 7 | 19 | - | - | 47 |
| Joaquina (*Chuaquina*) | 12 | 8 | 16 | 8 | 2 | - | 46 |
| Pascual | 4 | 14 | 17 | 9 | 1 | - | 45 |
| Juan + altres | 4 | 10 | 11 | 4 | 6 | 9 | 44 |
| Juan (*Chuan*) | 14 | 7 | 12 | 8 | 2 | - | 43 |
| Sebastián | 4 | 7 | 13 | 8 | 6 | 2 | 40 |
| Manuel | 7 | 4 | 13 | 10 | 3 | 1 | 38 |
| Manuela | 6 | 10 | 12 | 6 | 1 | - | 35 |
| Orosia | 5 | 7 | 12 | 6 | 3 | - | 33 |
| Luis | 1 | 8 | 6 | 7 | 9 | 1 | 32 |
| Mariano | 7 | 10 | 11 | 4 | - | - | 32 |
| Sebastiana | 6 | 7 | 9 | 3 | - | - | 25 |
| Miguela | 7 | 5 | 7 | 3 | 2 | - | 24 |
| Santiago | 4 | 5 | 9 | 4 | 2 | - | 24 |

## ANNEX 2
## LLISTA DE NOMS DE CASA. ECHO I ANSÓ[18]

**ECHO**

**Cases existents entre 1924 i 1986**

Abila
Agustineta
Ainica (tb.Maurizia)
Alifonso (tb.Pericón)
Ansotana (a.Escultor)
Bartolillo
Batista (a.Nicot)
Belén
Belzuz (ara Popolín)
Benigno
Bernabé
Biec
Bisaltico
Bisqué (a.Arnaudé)
Bizente
Blanco
Blasa (ara Chuané)
Blascosanz
Blasquico
Bolanté
Borderas
Borruel (1)
Borruel (2)
Cabalero
Caballero
Cabana (a.Chaime)
Caberé
Cabero
Cachón
Canteré
Cañizo
Capeta
Cardesa
Cardeseta
Carlos (tb.Chandón)
Carlota
Carpintero
Catarecha (1)
Catarecha (2) (a.Jacobo)
Clara (tb.Díez)
Clareta
Coles(1) (ara Gollins)
Coles (2)
Colorau
Concheta (a.Chaniquet)
Corradasa
Corrutaco (tb.Costeros)
Costeros (tb.Corrutaco)
Cristineta
Chalmarraco
Chardedeu
Chandegarzía
Chandón
Chanferré (tb.Carlos)
Changrosín
Chantarrén (a.Pericón)
Cheto (tb.Nabarro)
Chicón
Chilburro
Chilica
Chilico
Chuané (1)
Chuané (2) (a.Blasa)
Chuané de Sta.Luzía
Del Sol
Díez (tb.Clara)
Don Felipe (1)
Don Felipe (2)
Don Pascual
Dotor
El Herrero
Farol
Farolé
Ferrero (1)
Ferrero (2)
Ferrero (3)
Florentina
Francha (1)
Francha (2)
Galindo (a.Jenaro)
Gallé (1)
Gallé (2) (a.Lorenza)
Garzía
Gollins
Grábiz
Graitón
Guallar
Ibor
Iguarte (a.Marcantón)
Isidro
Jenaro (a.Galindo)
Jerónimo
Juanantonié
L'ospital
La Bentera
La Bizenta
La Casera (ara Royé)
La Caseta
La Fraila
La Mala
La Mina(1) (a.Mingarro)
La Mina(2) (a.Brigeda)
La Pastora
La Pena (tb.Martina)
La Rosa
La Rufa (a.Alfaro)
La Rufa de Sta.Luzía
La Zibila
Larraspa (a.Molinero)

(18) La transcripció dels noms s'ha fet d'acord amb la seva pronunciació

Larraz (ara Blasquico)
Leneta
Lo Bastero
Lo Bastero Barré
(a. Loarre)
Lo Biejo
Lo Cacho (1)
Lo Cacho (2) (a.Roque)
Lo Calbo
Lo Cantero
Lo Cleigo
Lo Coixo
Lo Chapi
Lo Lobé
Lo Maestro
Lo Miñón
Lo Nabarro Galdé
Lo Pallar
Lo Pelaire
Lo Pico
Lo Pipau
Lo Piquero
Lo Royo
Lo Sastre
Lo Sillero
Lo Tío
Lobera (1)
Lobera(2) (a.Maután)
L'Onsé
Lorenz
Lorenza (ara Gallé)
Lurbes
Machín
Madrilejo
Maestro Música
Malaño
Manché (1) (a.Candila)
Manché (2)
Manolón
Marcos
Maribarba
Marica (1)
Marica(2) (a.Tiroliro)
Mariferrera
Marín
Marta
Martes
Martín d'Ubero
Martina (tb.La Pena)
Marraco
Maurizia (tb.Ainica)
Menescal (a.La Mala)
Mercadé
Migosancho
Migosancho de Sta. Luzía
Mingolarraz
Miquetas
Miral
Murillo
Nabarro (tb. Cheto)
Nadal (1)
Nadal (2) (a.Chaniqué)
Narziso
Panadera
Panadería
Payán
Pazenzia
Pedraz
Pedro Marica
Pelaire
Peluca
Pepa
Pepe
Pericón (tb.Alfonso)
Pesorau
Petré (a.Saturnino)
Pexenca
Piché
Pierlis
Pimpán
Poli (ara Ansotana)
Ponz
Popolín (1)
Popolín (2) (a.Belzuz)
Prigoy
Ramoné
Reluchero
Roque
Rosalía
Royé
Sanclemente
Soto
Tracolillo
Ubero (a.Baretas)
Ubero de Sta.Luzía
Urbana
Zapatero
Zebadero
Zeferino

**Cases desaparegudes (abans 1924)**

Adolfo
Alfaro
Ardeu
Arnaudé
Baretas
Blasa
Brigeda
Bruno
Candila
Chaime
Chanchiqué
Chaniqué
Childópez
Escultor
Garín
Gastón
Gralao
Jacobo
L'Andaluz
La Cope
La Jota
La Molinera
La Mota
Lo Sastre
Lo Serrano
Loarre
Macacas
Marcantón
Maután
Mazo

Mingarro
Molinero
Mosín Chaime
Napón
Nicot
Pelegrín
Pericó
Peritón
Pichorro
Pipau
Rescate
Rin
Roque
Saturnino
Sebastián
Sorda
Tiroliro
Tondebarbero
Tracolillo
Zenarbe

**ANSÓ**

**Cases existents entre 1924 i 1986**

Abau
Abillanas
Achés
Adrián
Aguada (a. Capitán)
Albeita (a. Aguau)
Alberro
Alterao
Ambrosio (tb.Sampedro)
Andreba
Animetas
Anizeto
Antón
Añeta
Atanasio (a. Petra)
Aturra
Azná
Anarillo
Baldragas
Balera
Balero (a.Benabides)
Baretón
Baretoné (a.Escribano)
Bartola
Barrena
Barreneta
Basteré
Bastero
Belzuz
Benabides (a.Balero)
Bernarda (a.Pajalarga)
Berdadero
Billanúa
Bizén
Blanquis
Boque
Botero
Broxeta
Brullo
Bufés de Marcón
Burré
Burro
Caballé
Calba (tb.Chuané)
Camarón
Camilo (tb.Marilarrucha)
Camín
Capuch
Casero
Catalán
Catalina
Catibiela
Cazo
Clara
Clica
Cocorro
Collao (tb.José Pastor)
Conzezión (a.Manuela)
Cortés
Corradasa
Correo
Cuartón
Culera
Culerón
Cullito
Chacona
Champuyó (1)
(tb. Mañas Menor)
Champuyó (2)
(tb. Mendiareta)
Chanfuertes
Changalé
Chanmiguel
Chantomás
Chenero
Cheso
Chil
Chiquín
Chonillero
(a.Mariachunillas)
Chorchis
Chudas
Chullán
Chunillas
Chusepón
De Paco
Dionisio
Domén (tb.Gusanillo)
Donaz
Fabián
Fanal
Fernanda
Fernandé
Ferreré
Ferrero (a.Beatriz)
Fifaras
Fonz
Forastero
Foria
Foriereta
Franché
Fuertes
Galachera
Galero
Garay
Gargallé
Garós

Garzía
Gastón
Gato
Gorría
Gorrillón (tb.Latre)
Granero
Grazián
Güellos
Gusanillo (tb.Domén)
Jaqué
Jazinto
Jenaro
Jorge
José Pastor (tb.Collao)
Juana
Juanblas
Juandomingo
Juané (tb.Calba)
Juanera
Juanico
Juaniqueta
(a. Alejandro)
Julepe
Juliana (a.Manoleta)
Labarta (a.Corradasa)
La Mina
Larraza
Larrosa
Lasola
Latín
Latre (tb.Gorrillón)
Licao (a.Estrela)
Liforas
Litaco
Lobo (a.Antonio Pérez)
López (tb.Notario)
Lopico
Lorenta
Luzía
Llerta
Malcarao
Malojo
Manata
Manato
Manolero
Mañas
Mañas Capador
Mañas Menor
(tb.Champuyó)
Marcón
Marconé
Marcos (a.Molinero)
Marichunillas
(tb.Chunillero)
Marialarrucha
(tb.Camilo)
Marianico (1)
(a.Bonafonte)
Marianico (2)
Marichatis
Marión
Marosa
Maroseta
Marzelino Alto
Marzelino Baxo
Marrabil
Mayaguerra
Melera
Mendiara
Mendiareta
(tb.Champuyó)
Menitecla
Menuda
Menudo
Merejildo
Merín
Migalón
Mingurría
Mitierra
Molinero (a.Chesé)
Molino
Morené
Moreno
Morera
Moroto
Munrial Dos Ricos
Munrial Dos Pobres
Nabarré (a.Estudiante)
Nadal
Negro
Notario (tb.López)
O Ferrero Gusepillo
O Sastre Franché
Ojalatero
Ornà
Ornata
Orosia
Ospital
Ostias
Pabirrio
Pablo
Panchané
Pastor
Pataduro
Paulín
Payareso
Pedromiguel
Peña
Perromeu
Petra
Petré (1)
Petré (2) (a.Poli)
Pellejero (a.Marta)
Piero
Pierra
Piloré
Primo
Prisca
Pucho
Puro
Puyó (a.Montescuro)
Rafa (a.Antonio de Tío Bárbara)
Ramoné
Requinto
Rezio
Rogante
Romo
Rosa
Roscadiau
Royo
Sabineta
Salas
Saletas
Salomón

Sampedro (tb.Ambrosio)
San Bartolo (tb.Xierpe)
Santos
Sarto
Sebas
Serena
Serené
Serrano
Seta
Socorro
Soldada
Soro
Taconera
Talecón
Terrén
Tiadora
Tiarramonabau
Tierno
Tixinero
Tocata
Tolún
Tomás
Trinqué
Tuerto Nabarré
Tunante
Xanca
Xarnau
Xera
Xerón
Xierpe (tb.San Bartolo)
Ximena
Zapatero (a.Simón)
Zarzalla
Zezilio
Zirilo
Zurdo

**Cases desaparegudes (abans 1924)**

Aguau
Alfama
Agarra
Ana
Antonio de Tío Bárbara
Antonio Pérez
Arregui
Bartibax
Bastero Miguel
Beatriz
Bela
Bellote
Berdoy
Bergüés
Berruda
Bizenta
Bonafonte
Capitán
Correa
Corruscos
Curreta
Chandomingo
Chandomén
Chanretor
Chapis
Charnadé
Chartica
Chesé
Chilbarne
Chicurro
Chirón
Chucheta
Escribano
Estrela
Esquilero
Estudiante
Galé
Garzés
Gil de Cazo
Güellos Menor
Jerónimo
José Requinto
Juandemanuela
Juanillo
Julepe Menor
Jusepe
La Torreta
Lacruz
Larje
Lescuin
Luzio
Manate
Manoleta
Marta
Matón
Minilla
Micaela
Minguillo
Minilla
Monteoscuro
Norea
Pabloantonio
Pablero
Pantaleón
Paula
Pelao
Penetra
Periz
Perpetua
Petrazo
Poli
Polido
Roi
Salabé
Sancho
Tamborina
Tía Ramona Ornata
Tomasé
Xorila
Zibil

RAMON CERDÀ

# LEXICOGRAFIA I INTEL·LIGÈNCIA ARTIFICIAL

## 0. INTRODUCCIÓ

El contingut d'aquest treball, d'encàrrec, és extraordinàriament extens encara que només comprengui la intersecció de la lexicografia i la intel·ligència artificial. Havent de seleccionar, per raons d'adequació i realisme he triat adreçar-me primordialment a lectors amb formació lingüística. L'exposició està concebuda en bloc com una tipologia de la complexitat de les modalitats lexicogràfiques a partir d'un parell ordenat d'eixos: 1) el caràcter tancat o obert del text i, dintre del segon, 2) l'adopció de criteris sintàctics, semàntics o conceptuals de processament. De passada m'he esforçat a fer veure els vincles i les implicacions de tota mena que hi ha entre la lingüística computacional i la teòrica. No m'he pogut estar d'utilitzar sovint l'enutjós recurs a les notes al peu com un text en paral·lel tot assegurant-me que, en una lògica compensació, la lectura del text pelat en pugui prescindir sense dificultats.

La bibliografia queda reservada a unes recomanacions orientatives al final.

## 1. L'ESCENARI DE LA INTEL·LIGÈNCIA ARTIFICIAL

1.1 Simular i, si pot ser, emular tot el comportament humà no ha deixat mai de ser la fita prioritària de la informàtica en general. S'hi oposen dues grans limitacions correlatives: 1) la complexitat del comportament i 2) la seva integració. Ara per ara

*Caplletra 17* (Tardor 1994), pp. 81-108

és impossible reunir en una sola programació global totes les fonts de decisió i de processament que vénen a confegir actes tan senzills en aparença com, per exemple, reconèixer i saludar un amic pel carrer. La simulació s'ha hagut de conformar a actuar sobre unes quantes pautes regulars i a extreure'n models no sols simplificats, sinó també molt més dissociats del que hom voldria. D'aquí ve que la simulació humana global s'hagi fraccionat en sengles disciplines particulars:

(1) (a) la robòtica, que simula el comportament físic,[1]
(b) el processament del llenguatge natural (PLN), que simula el comportament lingüístic, i
(c) la intel·ligència artificial (IA), que simula el comportament intel·lectual.

Seguint un model humà, és lògic considerar que la IA governa les altres disciplines i en crea el marc natural d'integració. D'aquest angle estant podem considerar, doncs, el PLN com una part de la IA i aquesta com una part de la informàtica. També convindrem que si a un sistema se li demanen decisions estrictament lingüístiques –com ara verificar l'ortografia o l'estil d'un text, recuperar cadenes lèxiques o traduir–, el sistema en qüestió pertany a l'àmbit del PLN. Si se li demanen, en canvi, decisions 'pràctiques' –com calcular la direcció i la velocitat d'un mòbil i predir-ne la posició exacta en un temps donat o bé optimitzar la producció i els costos d'una fàbrica d'electrodomèstics–, el sistema pertany a la IA.[2]

Deixant de banda la robòtica i pel que fa a la IA i al PLN, si haguéssim de fer una caracterització molt sumària del decurs històric més recent, valdria a assenyalar unes quantes etapes:

(2) (a) Els informàtics primitius, als anys 50, preveien programes relativament senzills i, com a molt, llargs per construir sistemes capaços de raonar, decidir, sostenir converses i traduir entre llengües. Els fracassos foren espectaculars.
(b) Cap a finals de la dècada del 60, es va comprovar que calia canviar radicalment les concepcions bàsiques i es va demanar l'ajut de psicòlegs i lingüistes per construir programes no trivials d'inferències, sobretot en traducció automàtica (TA). Al mateix temps s'aconseguien avenços substancials i convergents en els camps del hardware, del software i de la representació lingüística.[3] A l'entorn de 1970 es desenvoluparen els primers sistemes experts (3.3) de gran abast enmig d'unes expectatives gairebé utòpiques sobre les possibilitats de la IA.
(c) Vint anys més tard es pot dir que s'han acumulat grans quantitats de coneixement efectiu sobre molts fenòmens del comportament humà tant racional com lingüístic, però manca encara una teoria global psicolingüística i sobretot resultats pràctics d'alt nivell prou atractius per propiciar inversions encara més grans en la recerca bàsica. Es renuncia, si més no momentàniament, a la integració teòrica i es produeix una concentració d'interessos a la bus-

(1) Encara que sol dur incorporat algun tipus de sistema 'intel·ligent' avançat (per exemple, per al reconeixement de formes), la seva característica pròpia és la precisió cinètica i motriu.

(2) En general, tot i que avui dia s'han anat incorporant especialitats originades en disciplines tradicionalment autònomes (lògica, psicologia i lingüística sobretot), la inspiració fonamental i la terminologia ('hardware', 'llenguatge natural', 'implementació', etc.) provenen de la informàtica (vegeu també nota 11). No seria raonable resistir-nos-hi aquí.

(3) Durant aquells anys crucials aparegueren el xip com a substitut del transistor, en el camp del hardware; en el camp del software, els llenguatges de programació –com ara el COBOL, BASIC i FORTRAN, que venien a sistematitzar i, per tant, a ampliar i simplificar enormement la comunicació amb la màquina– i, en el camp de la representació lingüística, la primera versió de la teoria estàndard de Chomsky, que acostava la descripció gramatical als requisits de formalització de la programació informàtica.

ca d'aplicacions pràctiques més realistes.[4] Mentrestant, es forma una immensa xarxa d'activitat en l'àmbit de la lingüística computacional (LC) concentrada sobretot als EUA, el Japó i alguns estats de l'Europa Comunitària. Tot això dintre d'un marc tecnològic extraordinàriament dinàmic on les innovacions poden crear de sobte qualsevol situació sense precedents.

Si s'ha perdut el sentit que tenia deu anys enrera discutir la presumpta correspondència psicològica d'un sistema informàtic no sols és perquè encara no se sap res de substancial sobre els processos psicològics que governen el llenguatge i el raonament en general, sinó també perquè se cerquen resultats més que no pas explicacions; es pretén tot plegat que un sistema tradueixi i en general que una aplicació funcioni, i prou. De fora estant, si hi ha res que distingeixi més clarament la lingüística teòrica (LT) de la computacional és la malfiança d'aquesta última envers les especulacions i l'orientació cap a l'obtenció de productes industrialment apetibles. Aquest caràcter aplicat es fa lògicament substancial en tots els apartats de la LC, i la lexicografia, com veurem, no és cap excepció.

1.2 Com és natural, també hi ha aspectes teòrics dintre la LC, que els tècnics anomenen de 'recerca bàsica' i que no sempre es diferencien nítidament dels que crea la LT. No està gens descartat que cadascuna es beneficiï de qualsevol troballa de l'altra i que fins i tot comparteixin models teòrics sencers. De fet, s'ha especulat bastant sobre la interrelació entre nivells metateòrics, teòrics, formals i procedurals. Així, per exemple, la Gramàtica lèxico-funcional (*LFG, Lexical-Functional Grammar*) de R. Kaplan i J. Bresnan fou concebuda per ser implementada com a mecanisme computacional, però el seu desenvolupament ha estat igualment remarcable com a teoria de la gramàtica. També es pot dir el mateix en alguna mesura d'altres models com ara la Gramàtica d'estructura sintagmàtica generalitzada (*GPSG, Generalized Phrase Structure Grammar*) de G. Gazdar, E. Klein, G. K. Pullum i I. A. Sag, o la seva successora, la Gramàtica d'estructura sintagmàtica orientada al nucli (*HDPSG, Head-Driven Phrase-Structure Grammar*) de C. Pollard i col·laboradors. D'altres models teòrics, en canvi, no tenen rellevància estricta per a la LC, sigui per manca de formalització (per exemple, la glossemàtica de L. Hjelmslev) sigui perquè no és computacionalment accessible (per exemple, la teoria de l'adquisició lingüística de J. Piaget) sigui pel seu abast (per exemple, una descripció interdialectal), etc., etc. Com he dit, però, la LC s'orienta cap a l'obtenció de principis i procediments lingüísticament adequats, explícits i computacionalment efectius. D'aquests principis i procediments se'n diu tècnicament 'formalismes', per oposició a les teories gramaticals de la LT, que cerquen una concepció abstracta o, com a mínim, general del llenguatge. Entre els principals formalismes cal citar, si més no, la Gramàtica d'unificació funcional (*FUG, Functional Unification Grammar*) de M. Kay i la sèrie derivada dels sistemes PATR de S. M. Shieber, L. Karttunen i F. C. N. Pereira.

(4) Pel que fa al PLN aplicat a la lexicografia, hi ha una remarcable concentració d'esforços encaminats a la creació de grans corpus de referència i de diccionaris electrònics de diversos nivells, a la reutilització de recursos lexicogràfics, és a dir a la creació d'estris que puguin recuperar automàticament, o almenys semiautomàticament, la informació continguda en diccionaris convencionals per a diversos sistemes de processament, i a establir les possibilitats de migració de formalismes gramaticals per tal d'integrar i optimitzar models en ús més o menys afins i crear-ne d'estàndards d'àmplia acceptació.

## 2. LES MANIFESTACIONS BÀSIQUES DE LA LEXICOGRAFIA

2.1 Fins fa relativament poc temps, la lexicografia es limitava gairebé a confegir diccionaris que els parlants nadius feien servir per millorar l'expressió i aprendre el significat d'algun mot desconegut i els estudiants de llengües, utilitzant-ne de bilingües, per traduir. Els diccionaris eren estris auxiliars físicament i temàticament independents de les gramàtiques de les llengües respectives. O, almenys, a les classes de sintaxi o de morfologia ni es miraven, com si fossin camps inconnexos. I tot i que els lexicògrafs s'escarrassaven a reclamar que la feina que feien era veritablement científica i lingüística, el cert és que la seva visió de la llengua era la d'un cos esbaconat amb els òrgans distribuïts pel terra en rigorós ordre alfabètic, cosa que difícilment podia reproduir o suggerir cap activitat vital (compareu a 4.3.2 les figures (19) i (20)).

En la LT dels anys 60 ençà, amb la formalització de la gramàtica, la situació ha canviat dràsticament. El diccionari s'ha convertit en un element cada cop més integrat dins el sistema general, fins al punt que sovint ni tan sols no es pot distingir o aïllar com a mecanisme independent. Prenent denominacions i aparences més o menys diferenciades, el lexicó, el sector que conté la informació lèxico-semàntica de la gramàtica, sol ocupar una posició de partença i de control en tots els models, a vegades amb una preeminència explícita, com en la Gramàtica lèxico-funcional, que he mencionat abans.

En aquest aspecte és força simptomàtica la derivació que N. Chomsky fa de la teoria anomenada de Recció i lligament (*GB, Government and Binding*) on el mecanisme generatiu d'estructures oracionals ja no correspon a regles del component de base, sinó a principis de projecció que provenen de les entrades lèxiques subcategoritzades, és a dir amb estipulacions necessàries per a establir els esquemes sintàctics locals (vegeu 2.3 i sobretot el capítol sobre Lexicografia i Models Lingüístics). Així s'ha optat per una concepció lexicografista de la gramàtica pròxima en certs aspectes a la de molts models computacionals, com veurem successivament.[5]

2.2 Dintre del PLN la prioritat correspon a la informació lexicogràfica. Des del principi, la visió *naïve* característica de les primitives experiències en TA (l'activitat més complexa dintre el PLN) ja era essencialment lexicografista, en el sentit que concebia les llengües en primer lloc com a grans diccionaris complementats per unes quantes instruccions combinatòries, és a dir sintàctiques. Traduir es limitaria, segons això, a establir la correspondència biunívoca entre els mots de dues llengües que componen un text i a arranjar el resultat d'acord amb la sintaxi de la llengua de destinació. Els productes d'aquest procediment mot a mot, però, sovint no deixen ni copsar de què tracta el text així "traduït" per la senzilla raó que la massa lèxica d'una llengua no té estructura discreta (en el sentit matemàtic) ni estable i molt menys guarda una relació isomòrfica amb la massa lèxica d'una altra llengua. Qualsevol diccionari bilingüe de butxaca ho mostra prou bé; més encara, la qualitat dels diccionaris bilin-

(5) Curiosament, en suprimir les regles sintàctiques s'invalida el concepte primigeni de gramàtica generativa introduït precisament pel propi Chomsky. Cal afegir, arran d'aquesta aproximació estratègica, que la teoria estàndard es va mostrar molt poc apta per a la implementació computacional i que el mateix Chomsky sempre ha sostingut no haver tingut mai en compte concepcions o fites computacionals.

gües es basa precisament a magnificar i descriure escrupolosament la incongruència lèxica entre dues llengües qualsevol.

Cal advertir, però, que la gran majoria d'aplicacions més perfeccionades no han canviat pròpiament el disseny lexicografista, sinó que n'han identificat les principals mancances i han incorporat un criteri composicional, oracional o textual. En tot cas, la diferència és força transcendent, com veurem més avall, ja que aconseguir que una expressió sencera en la llengua A correspongui d'una manera plausible o útil a una altra expressió sencera de la llengua B requereix un procés de formalització analítica i sintètica d'una complexitat molt variable.

I trobar el camí més curt o, si es vol, el procediment més barat i efectiu representa tot el que es proposa en definitiva el PLN aplicat a la TA i en general tota aplicació lingüística de la IA.

## 3. TIPOLOGIA DE LA COMPLEXITAT

3.1 El concepte d'optimització o de solució minimista –equivalent a la 'navalla d'Ockham' clàssica– és crucial en informàtica. Aplicat al PLN, representa l'afany de salvar amb el mínim esforç possible la complexitat provinent de tres paràmetres bàsics i teòricament independents: 1) l'abast temàtic del text, 2) la indeterminació del flux de la informació i 3), quan hi ha algun procés de transferència interlingüística, la distància tipològica que separa les llengües implicades (vegeu 4.2 (10)). Aplicat a la IA, representa enfrontar-se a la complexitat 1) de la interacció entre l'usuari i el sistema i 2) de l'execució dels requeriments implicats.

Com dèiem abans, els models en PLN són lexicografistes almenys en dos sentits: 1) perquè sempre es parteix i/o s'acaba en un lexicó, i 2) perquè, sigui quina sigui la finalitat del producte, el procés s'inicia en un text escrit o oral i la primera operació 'intel·ligent' que s'executa sobre ell és de naturalesa lexicogràfica (tot i que el primer escorcoll pot ser purament gràfic o fonètic). Podem convenir que el grau de complexitat d'un sistema computacional determinat pels paràmetres del paràgraf anterior promou la necessitat (no necessàriament correlativa amb cadascun d'ells) de recórrer a mecanismes addicionals més o menys distingibles, a part dels lexicogràfics: el més immediat és essencialment sintàctic, el següent, semàntic, un tercer lògic i finalment un de 'coneixement del món'. Ara bé, tot i que la incorporació creixent d'aquests mecanismes augmenta el grau de complexitat i de poder executiu dels sistemes, el principi de l'optimització fa que es tendeixi a utilitzar sistemes adequats a cada aplicació, vetllant que el preu de l'esforç no ultrapassi el valor del resultat. No té sentit, per exemple, incorporar un mòdul sintàctic a una aplicació que hagi de distingir només entre una dotzena d'opcions significatives.

Cal afegir, d'altra banda, que en la pràctica els models que exemplifico a continuació com a il·lustració dels graus i modalitats de complexitat computacional, en especial els més elaborats, sovint no es donen en forma pura o absoluta en la realitat, sinó en alguna combinació entre ells.

3.2 La producció computacional d'estris lingüístics més simple de totes és purament lèxica i es troba en la presentació d'alternatives binàries de tipus *Sortir? (S/N)* o bé *Vol resguard? (S/N)*, força típiques en els 'menús' de programes d'ordinador i de caixers automàtics, respectivament, on s'insta l'usuari a pitjar la tecla adient representativa de "sí" o "no". L'abast temàtic del text es redueix aquí a tres mots, comptant la manca de resposta. El sistema es limita a comprovar, a través d'algun codi intern de tipus ASCII, amb quina de les tres possibilitats previstes al seu lexicó coincideix la tria de l'usuari. En general, aquesta comprovació lèxica sense anàlisis prèvies ni ulteriors s'anomena 'correspondència de patrons' (*pattern matching*) i s'aplica en textos tancats, és a dir sempre que hi ha un abast temàtic limitat pel propi sistema, en el sentit que la comunicació entre l'usuari i el sistema segueix un flux preestablert –però no necessàriament reduït, ja que pot comprendre un conjunt de centenars de menús i submenús jerarquitzats (com, per exemple, en el WordPerfect)–. No cal que el patró sigui un mot realment existent en una llengua natural. En els exemples de més amunt el patró es redueix a [ϕ['s'/'n']], com hem vist, i en d'altres pot equivaler a una seqüència complexa com ara el nom i cognoms de l'usuari o l'ordre d'imprimir un tros seleccionat d'un text. Les aplicacions de la correspondència de patrons són força populars i diversificades, ja que, a part dels menús informàtics a què hem al·ludit, es troben en productes com ara els traductors de butxaca, capaços d'establir correlacions entre milers de mots o frases fetes d'un bon grapat de llengües (aquest límit és purament comercial) amb o sense síntesi de veu incorporada per simular la pronúncia figurada del resultat, en els reconeixedors de veu per telèfons sense mans o per consultes bancàries o en els 'correctors ortogràfics' de molts processadors de text. I encara, dintre de sistemes molt perfeccionats de TA (vegeu 4.3), s'utilitzen també procediments anàlegs de correspondència lèxica entre termes tècnics gràcies a la seva equivalència interlingüística estricta.

3.3 Una família més complexa de productes computacionals en PLN essencialment lexicogràfics són les interfícies home-màquina en llenguatge natural per a moltes bases de dades i alguns sistemes experts senzills. Tant la base de dades com el sistema expert consten d'un fitxer amb informació ordenada i un mecanisme incorporat per establir relacions específiques d'informació. Com veurem després, la diferència fonamental està en la simplicitat, la integració i l'orientació de l'algorisme relacional del mecanisme. En una base de dades convencional es pretén recuperar informació factual entre les dades ja contingudes en la base; aquesta informació és bàsicament lingüística i

sovint consisteix en la localització i recuperació de cadenes específiques de caràcters, tècnicament mots. Entre els exemples més típics hi ha la guia telefònica, el catàleg d'una biblioteca o el Diari Oficial de la Comunitat Europea. En un sistema expert es pretén, al seu torn, obtenir o, millor, simular informació nova conjuminant les dades ja contingudes en el sistema amb les que li forneix l'usuari. Es tracta ara d'informació conceptual, que requereix la creació d'un cert món possible amb objectes coherents (vegeu 5.2). Els sistemes experts es construeixen a partir de dades, anomenades bases de coneixement, jeràrquicament emmagatzemades per especialistes, o experts, en una matèria (medicina, geologia, enginyeria, finances...) juntament amb un algorisme relacional, anomenat motor d'inferència, capaç d'establir un raonament dialèctic amb l'usuari i finalment de formar un criteri, posem per cas, sobre un diagnòstic mèdic, la probabilitat de trobar petroli en un determinat sòl, les expectatives d'èxit d'una certa inversió o les continuacions imprevisibles d'un conflicte armat a gran escala.

La principal utilitat d'aquelles interfícies, genèricament anomenades 'gramàtiques semàntiques' rau en evitar que s'hagin d'aprendre o memoritzar codis o llenguatges només intel·ligibles per la màquina i en canvi permetin a l'usuari recuperar o obtenir informació a través d'un diàleg interactiu, és a dir de consultes bàsicament no limitades en la seva formulació ni ordenació. La gramàtica semàntica comporta un cert tipus d'anàlisi essencialment orientat cap a la descoberta de 'mots claus' enmig d'estructures més o menys diverses i complexes. El principi es basa en el supòsit que l'usuari pot triar entre una munió d'estils lingüístics segons les seves habilitats o preferències, però que, en qualsevol cas, farà servir un grapat força reduït de mots precisos per descriure l'objecte central de la seva consulta. L'escassa variabilitat d'aquests mots permet confegir un sistema que els localitzi i arbitri una interpretació de la petició global només a partir de la seva co-aparició.

Imaginem que un lector d'una biblioteca general vol saber què hi ha publicat sobre els índexs d'escolarització a Espanya durant els últims deu anys. Com serà una aplicació teòricament mínima d'una gramàtica semàntica? Un sistema així es compon bàsicament de quatre mòduls:

(3) (a) un conjunt estrictament jerarquitzat de camps de contingut (inspirats, per exemple, en la classificació bibliogràfica decimal) representats per sengles mots claus
(b) un conjunt de possibles alternatives lèxiques i morfològiques a cada mot clau de (a)
(c) un conjunt de regles de detecció de mots claus i de relació entre ells
(d) el conjunt de publicacions catalogades amb indicació acurada del contingut mitjançant els camps de la llista (a)

La tercera llista és, de fet, una gramàtica i consisteix en una col·lecció de regles que executen tres funcions específiques: 1) identifiquen els mots claus a través de les seves

possibles alternances expressives previstes en (b) i eliminen la resta de mots, 2) estableixen les possibles relacions de sentit entre els mots claus i els camps de contingut adients i 3) componen una llista de les publicacions amb la reunió o la intersecció dels camps de contingut seleccionats.[6]

Vejam, doncs, com funciona. En primer lloc, el sistema negligirà tot allò que no faci referència als mots claus o a les seves alternatives. Coses com, per exemple:

(4) (a) Quines obres hi ha sobre ...?
(b) Què hi ha de ...?
(c) Voldria saber si s'ha editat res sobre ...
etc.

Els mots claus de la consulta són, en canvi, 'escolarització', 'Espanya' i 'període d'anys' en les seves possibles modalitats expressives. Respectivament:

(5) (b) escolaritat, escolar, escolaritzat, estudiant, estudiantil, que estudia, que rep educació, discent, alumne ...
(c) Estat, estat espanyol, espanyol, estatal ...
(d) deu anys, decenni, dècada, període ...

tot tenint en compte les possibles combinacions i alteracions morfològiques que en definitiva poden donar lloc a construccions tan diverses com les següents:

(6) (a) Desitjaria obtenir dades sobre la [població estudiantil] [espanyola] [de 1983 ençà]
(b) Quines publicacions teniu sobre l'[educació] a l'[Estat] dels [últims deu anys]?
(c) Durant el [present decenni], què hi ha sobre la [gent que estudia] a [Espanya]?
etc.

La intersecció dels camps 'escolarització', 'Espanya' i '1983-1993' donarà presumiblement una sola interpretació força congruent amb el que cercava l'usuari. Noteu de pas que, tal com està dissenyada aquesta gramàtica semàntica, en teoria almenys també s'obtindrien resultats anàlegs als de (5) si s'escrivia la consulta a base de:

(7) (a) L'espanyola, quan estudia, estudia de debò, de 1983 ençà
(b) Espanyols i estudiants, molts i grans l'últim decenni
(c) Decenni últim el grans i molts estudiants i espanyols
etc.

és a dir sempre que s'hi incloguessin rèpliques adients dels mots claus.

(6) En el nostre exemple, assignaran la referència als últims deu anys al tema (a l'escolarització espanyola) i no pas a les publicacions mateixes i establiran si la relació entre determinats camps és d'unió o d'intersecció.

Podem concloure que, al marge dels mecanismes morfològics i combinatoris (i per tant sintàctics) auxiliars, el sistema funciona amb principis essencialment lexicogràfics semblants als dels diccionaris electrònics de baix nivell, on els valors de les entrades són fixos. Seguint un símil convencional, els mots claus fan de morfemes biunívocament relacionats amb els camps de contingut, equivalents a semes. I el diccionari més extens (la llista (3) (d) de més amunt), el catàleg de publicacions, consta d'entrades lèxiques subcategoritzades per un conjunt ordenat de semes o camps de contingut.[7]

3.4 El següent grau de complexitat es presenta en el text obert, quan l'abast temàtic de l'aplicació queda indeterminat. No es tracta pròpiament d'una mera ampliació lineal, com si s'afegissin volums a una obra, tot i que això també afavoreix que les gramàtiques semàntiques siguin insuficients. La indeterminació temàtica es dóna quan l'usuari pot triar entre una sèrie oberta –o prou ampla– d'expressions i els valors lèxics dels 'mots', que en els models anteriors eren fixos, passen a dependre del context. En no poder negligir cap segment del text, cal analitzar-ne l'estructura sencera i produir així una nova dimensió composicional, sintàctica i/o semàntica, si més no. Això comporta afegir al lexicó una gramàtica i diversos mecanismes que creïn alguna representació canònica del text inicial capaç de ser utilitzada adequadament per intercanviar informació amb usuaris humans o amb altres sistemes o convertida en text d'una altra llengua. En el primer cas, podria tractar-se de bases de dades i sistemes experts amb moltes variables, correctors ortogràfics avançats i correctors d'estil –que requereixen analitzar, com a poc sintàcticament, l'estructura de textos no predeterminats– i sobretot en la traducció automàtica.[8] En condicions normals això faria, per exemple, que les distintes versions que hem assenyalat a (6) i (7) fossin considerades no equivalents i en particular que (7) (c) quedés rebutjada per agramatical.

Un cop accedim a aquest tipus màxim de complexitat en PLN ja es planteja obertament la simulació sencera del parlant humà en totes o almenys en algunes de les seves funcions lingüístiques. Tothom accepta que hi ha diferències fonamentals –possiblement insalvables– entre el processament lingüístic humà i el computacional. El parlant humà utilitza en paral·lel, és a dir alhora, fonts d'informació tant lingüístiques (coneixement d'estructures fonològiques, morfològiques, sintàctiques, lèxiques, etc.) com extra-lingüístiques (coneixement propi i compartit del món ambiental, del tema que es tracta, la seva coherència i implicacions, etc.). Noti's que aquesta integració informacional representa un mitjà potentíssim per recuperar moltes de les mancances i deficiències que es produeixen en el curs de tota comunicació. Al seu torn, la gran majoria de sistemes computacionals d'avui dia 1) només poden comptar amb informació estrictament lingüística, i 2) incorporen aquesta informació en forma seqüencial o modular, és a dir per un costat fonològica, per un altre sintàctica, per un altre lèxica... seguint algun ordre preestablert (vegeu 4.3.2 (21)). A més a més, el parlant humà

(7) Avui en dia s'han desenvolupat formalismes dintre d'aquesta classe, intermèdia entre la correspondència de patrons i els *parsers* (vegeu nota 11), del tipus anomenat 'SQL' (*Structured Query Language*) que vénen a optimitzar el diàleg interactiu entre usuaris i màquines a base de llenguatges reduïts per selecció de necessitats i que són capaços de processar quantificacions i variables contextuals força més complexes de les que he exemplificat abans (per exemple, ordres que demanen classes d'objectes diferents en alguna característica a les que s'han obtingut en una consulta anterior).

(8) Noteu que, sense control sintàctic, un corrector ortogràfic no detectarà cap raresa en una seqüència com ara *el flors blaus*.

combina les distincions discretes (o digitals) de la fonologia i la morfologia amb distincions contínues (o analògiques) en multitud d'aspectes primordials com ara la modulació de la veu i la semàntica.[9] En canvi, tot i la insistència en l'adaptació de tècniques de raonament aproximat en sistemes experts, els resultats encara no s'acosten significativament als humans.

Aquestes diferències –que desemboquen en la manipulació d'informació redundant, per part dels humans, enfront d'una informació fortament deficitària, per part dels sistemes computacionals– han suscitat la necessitat de cercar tècniques imitatives o bé alternatives. La possibilitat d'incorporar coneixements del món a un sistema depèn molt dels nostres coneguts factors de complexitat (sobretot l'abast temàtic i la indeterminació), però per a molts especialistes és una tasca purament utòpica, ara per ara, al marge d'àmbits d'aplicació molt restringits. Hi ha qui dubta fins i tot de la possibilitat de construir diccionaris electrònics coherents i efectius, tant multilingües com monolingües, que ultrapassin el vocabulari de terminologies tècniques i algun grupet de mots d'una sola accepció (si és que n'hi ha cap). Tanmateix gairebé tots els sistemes més vàlids en IA i PLN es fonamenten en el supòsit que tot això no sols és possible, almenys en teoria, sinó que els formalismes del futur seran essencialment diccionaris integrats (amb informació morfològica, sintàctica, semàntica, etc.) universals.

A continuació ens concentrarem en les modalitats i tècniques lexicogràfiques que s'utilitzen en aquest últim grup més complex i avançat de sistemes computacionals.

## 4. LEXICOGRAFIA I FORMALISMES EN PLN

4.1 Com dic, la necessitat de recórrer a una anàlisi combinatòria de les expressions assoleix un nivell de complexitat diferent. Des de l'angle de la lexicografia, es caracteritza perquè desapareix el supòsit que els mots tenen un valor fix i en prenen un de determinat pel context, tal com s'esdevé precisament en la parla real. És el fenomen genèricament conegut per 'ambigüitat', que cobreix un seguit ben diferenciat de manifestacions i graus possibles:

(8) (a) **ambigüitat lèxica:** *Els lingüistes computacionals són decididament **rars*** (pot ser en el sentit d'"escassos" o d'"estranys")

(b) **ambigüitat sintàctica:** *La va veure **tot passejant*** (no se sap qui passejava, si el subjecte o l'objecte de *veure*)

(c) **ambigüitat fonològica:** *Allà hi ha **les ones/les zones** més perilloses* (en ambdós casos amb una mateixa pronúncia)

(d) **ambigüitat referencial:** *Va treure la mosca de la truita i se **la** va menjar* (no se sap què va menjar exactament, si la mosca, la truita o alguna altra cosa referida prèviament)

(e) **ambigüitat pragmàtica:** *plat d'arròs, plat de test, plat de postres* (només

(9) L'extraordinària potència semiòtica del parlant li permet reconstruir o, si més no, establir hipòtesis sobre el sentit i la significància empírica de textos (orals, escrits...) encara que siguin molt fragmentaris, incorrectes o anòmals, des d'un simple gargot fins a un missatge críptic, passant per les més alambinades figures de dicció, entre elles, per exemple, la ironia. Sense esforç aparent, el parlant pot canviar radicalment de registre i de tema (o totes dues coses alhora), utilitzar més d'un codi en paral·lel i construir interpretacions diverses sobre un mateix text. El parlant humà, a més a més, no fa un aprenentatge procedural de la llengua materna (assimilant prèviament una col·lecció de regles), sinó declaratiu, en el sentit que infereix, a partir d'un nombre limitat de fets de parla, les estratègies generals que governen la creació de nous fets de parla que mai no ha sentit abans. Com veurem, tot això no fa sinó il·lustrar la indeterminació o, per ser més exactes, la virtualitat essencial dels significats lèxics en l'ús real de les llengües naturals.

el coneixement del món ens fa interpretar que l'arròs és el contingut, el test la matèria i les postres la funció del plat, però res d'això no es dedueix de la sola estructura gramatical)
etc., etc.

Cal comprendre que l'ambigüitat es presenta sempre, per la pròpia natura de les llengües, en el sentit que els significats virtuals dels mots només prenen valors específics, en rigor irrepetibles, quan compareixen dintre de determinades expressions. En part ho reflecteixen els diccionaris monolingües convencionals en distingir entre 'entrades' i 'accepcions'.

Això es posa especialment de relleu en la TA, l'activitat sens dubte més completa del PLN pel fet d'haver de fer conjuminar estructures de dues llengües, si més no. També, en els diccionaris convencionals bilingües cada mot de la llengua d'origen consta almenys d'una entrada lèxica amb un conjunt d'accepcions que equivalen a altres tantes entrades de la llengua de destinació. Per evitar que l'expressió bíblica en llatí *Spiritus quidem promptus est caro vero infirma* es verteixi com a *El licor és al punt, però el bistec està fofo* (en lloc de *L'esperit està disposat, però la carn és feble*) cal establir certament les condicions per les quals les entrades lèxiques, els mots, del text d'origen, passant per les accepcions, donin lloc a les entrades lèxiques adients del text de destinació.[10] Això només es pot aconseguir obrint inevitablement la via a l'anàlisi sintàctica i és en aquest sentit com s'estableix la relació intrínseca entre el diccionari i els mecanismes d'anàlisi, i síntesi sintàctica, els anomenats 'parsers' i generadors en l'àmbit del PLN.[11] A partir d'aquí comencen les diferències.

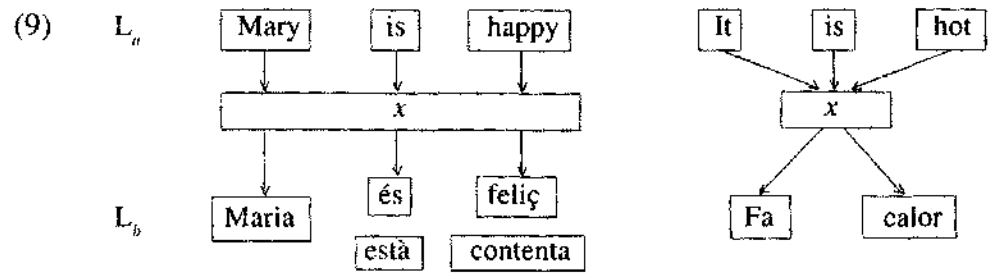
4.2 Ara bé, en comptes de penetrar en una casuística sobre els aspectes diferencials dels formalismes –que omplirien una pila de tractats–, optarem per presentar un panorama general tot adoptant, a més a més, la perspectiva més favorable de la lexicografia i la TA. La majoria d'analitzadors sintàctics actuals, entre ells els que hem citat a 1.2, tenen en comú que procedeixen de models coneguts per 'gramàtiques d'estats finits', especialment adequats per a la implementació computacional en forma de les anomenades 'xarxes de transició augmentada' (*ATN, Augmented Transition Nets*) impulsades sobretot per W. A. Woods cap a 1970. En desenvolupaments ja clàssics, a més de resoldre diversos problemes de procediment sintàctic (en especial, la recuperació d'informació cap enrere, la supressió de transformacions i l'aplicació sistemàtica de la unificació), havien anat incorporant una diversitat d'aspectes lèxics que analitzarem tot seguit.[12]

Per això cal aprofitar la distinció, de gran abast teòric i empíric en la TA, que es fa entre la 'transferència' i la 'interlingua'. Considerem-la a través dels següents exemples. Si comparem la traducció de l'anglès al català d'expressions com ara *Mary is happy* i *It's hot*, convindrem que ocupen gairebé posicions extremes en l'escala de la literalitat i la no literalitat. Gràficament ho podem representar així:

(10) L'exemple es refereix a una coneguda facècia que ridiculitzava les primeres experiències en TA mot a mot a partir d'aquell passatge, extret de la *Vulgata*. El mot *spiritus* fou traduït a l'anglès per l'equivalent a "licor", en comptes d'"esperit", i *caro* ho va ser com a "vianda", en lloc de "carn [humana]". El presumpte sistema va escollir accepcions possibles que un traductor humà segurament ni hauria advertit per l'oposició que el propi text estableix entre l'esperit i el cos.

(11) El 'parser', o analitzador sintàctic, és una composició terminològica confegida a partir del llatí *pars orationis* i s'aplica a l'operació de determinar computacionalment les derivacions d'una oració d'acord amb uns principis gramaticals. Més que una excepcional condescendència cap al llatí i als hàbits terminològics de la tradició, el barbarisme *parser* sembla una proclamació de l'actual hegemonia de l'anglès i d'altres tics concomitants (ús d'acrònims, etc.). Quant als termes 'generador', 'generació', provenen de la matemàtica i més en especial de N. Chomsky (vegeu 4.3).

(12) El procediment de la 'unificació', inicialment adduït per M. Kay, és considerat una aportació cabdal en el desenvolupament de formalismes computacionals. Consta bàsicament d'un mecanisme construccional que compara els trets gramaticals de les estructures (arbòries) locals i n'arreplega –unifica– els compatibles per tal de bastir estructures cada vegada més grans formalment definides (vegeu-ne un exemple a 4.3.1).




En general tota traducció comporta 1) la interpretació d'una estructura d'entrada en una llengua $L_a$ i 2) la seva conversió en una estructura de sortida en una llengua $L_b$ que tingui la mateixa interpretació. (Algú ho ha expressat sintèticament dient que es tradueixen interpretacions i no mots.) En l'esquema representem aquest doble processament amb un rectangle intermedi amb una *x* al mig. En el primer cas de (9) sembla que el processament manté l'estructura gramatical i lèxica del text inicial, tot i que la interpretació queda oberta a dues possibles traduccions en el text de destinació. En el segon cas, en canvi, la interpretació comporta una reorganització total d'estructura gramatical i lèxica. Del pronom *it*, la còpula *is* i l'adjectiu *hot* del text inicial no en queda cap rastre en el pro-verb *fa* i el nom *calor* del text de destinació. Si, malgrat això, la interpretació d'ambdós textos coincideix, caldrà convenir que el processament ha passat per una representació intermèdia neutra o, en qualsevol cas, diferent de les estructures gramaticals de tots dos extrems.

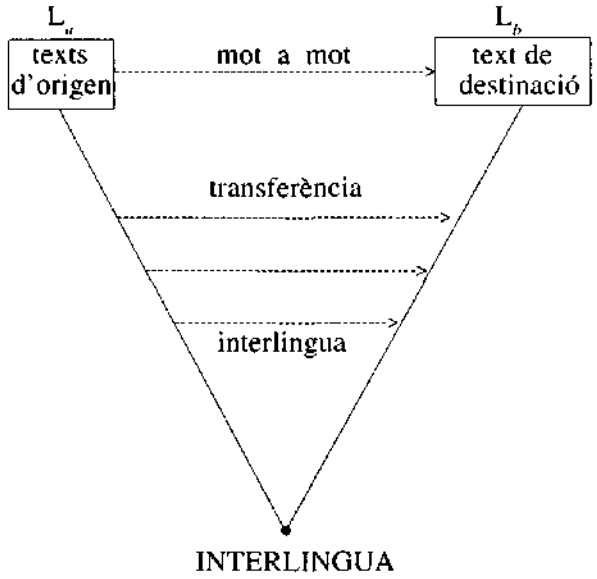
De quina naturalesa és aquesta representació? Molts filòsofs i lingüistes teòrics hi han dedicat grans esforços en les discussions seculars sobre els universals lògics i lingüístics. És la 'gramàtica universal' dels generativistes i la 'interlingua' de la TA, a la qual dóna un innegable suport l'evidència psicolingüística quotidiana dels traductors humans que verteixen sense cap restricció expressions entre llengües il·limitadament allunyades en disseny tipològic. Durant un temps es va pensar que el procediment per excel·lència de la TA consistiria a crear 1) una representació universal, una interlingua, per a tots els continguts lingüístics possibles i 2) les vies d'accés per a cada llengua natural. És, notem-ho bé, la hipòtesi justament contrària a la TA mot a mot. Encara que podia semblar que això complicava l'operació habitual de traduir, en interposar dues passes noves, s'entenia que el model era màximament 'realista' comparant-ho amb l'activitat del parlant humà. El mal és que ningú no ha aconseguit mai ni tan sols albirar de debò com poden ser cap d'aquelles dues fites...

La idea ha estat després replantejada en termes empírics (i amb canvis terminològics, com veurem), a partir d'una altra evidència intuïtiva i estadística igualment incontrovertible: l'afinitat tipològica entre dues o més llengües és directament proporcional al nombre de possibles traduccions literals o quasi-literals correctes entre textos respectius. En rigor, la similitud de les llengües és precisament això. Realistes

o no, en el sentit d'abans, el cert és que les traduccions literals són molt més barates, des del punt de vista procedural, que no pas les no literals. En canvi, les traduccions del segon cas de més amunt junt amb modismes, frases fetes, etc.– requereixen un esforç computacional màxim i molt sovint encara insuficient.[13]

En la mesura, doncs, en què la TA no és profitosa mot a mot ni accessible en la interlingua (en el sentit que hem descrit), s'ha optat per algunes solucions intermèdies, que els especialistes han denominat genèricament de 'transferència'. Gràficament:

(10)




Per anar del text d'origen en la llengua $L_a$ al text de destinació en la llengua $L_b$ es pot prendre la drecera, en la traducció mot a mot, o seguir el trajecte més remot, el que passa per la INTERLINGUA. La distància recorreguda queda compensada amb escreix pel fet que en la drecera les llengües implicades són màximament diferents, mentre que en la INTERLINGUA són una mateixa cosa.

Però no és aquesta l'única consideració important aquí. Com hem vist a 3.1, els sistemes de PLN que actuen sobre textos tancats, és a dir d'abast temàtic limitat i flux de la informació controlat, tenen la característica general d'atribuir valors fixos als elements lèxics utilitzats. I en la mesura en què això és vàlid, el procediment mot a mot o de correspondència de patrons (3.2) és justament l'indicat. Al seu torn, els sistemes que actuen sobre textos oberts, és a dir quan hi ha implicacions contextuals decisives en la determinació del valors lèxics, cal habilitar algun procediment construccional interpretatiu. I en la mesura en què se n'habiliten, el trajecte es fa certament més llarg i complex, però en acostar-se les representacions obtingudes de les llengües implicades la traducció és paradoxalment més directa i acurada. L'esquema (10) assenyala diverses possibilitats en aquests trajectes indirectes, unes més 'profundes' que altres. De menys

(13) Tampoc no han faltat alternatives a favor de llengües pivot preexistents considerades 'transparents' per la seva regularitat. Fa anys que I. Guzmán de Rojas, un informàtic bolivià, ve proposant l'adopció de l'aimarà, llengua ameríndia dotada, entre altres característiques, d'un sistema de sufixos òptim per a la lògica trivalent. L'empresa holandesa DLT utilitza l'esperanto per al seu sistema de traducció, que ja es comercialitza. Evidentment, són només aproximacions relatives a la noció d'interlingua que, com a molt, tenen un abast limitat a certs fenòmens i/o a un grup restringit de llengües naturals.

a més profunditat, en general l'escala estableix anàlisis successives sovint subdividides d'estructura morfològica, sintàctica, semàntica i lògico-conceptual.

Aquestes diferències no són tampoc aquí estrictament quantitatives, sinó que radiquen precisament en la complexitat del tractament lexicològic. Resumint i simplificant (passant per alt tota interferència): abans del nivell sintàctic, es treballa sobre mots o entrades lexicogràfiques predeterminades. Entre els nivells morfològic i sintàctic, es treballa sobre equivalents d'accepcions, és a dir subentrades lexicogràfiques en el sentit que precisarem a 4.3.2. A partir del nivell semàntic, es treballa sobre trets sèmics més abstractes.

Aquí hi ha, com dèiem, una impropietat terminològica, en el sentit que s'acostuma a aplicar la denominació d''interlingüístic' a tot sistema que actua més enllà dels nivells estrictament sintàctics, per molt lluny que estiguin encara d'arribar a la INTERLINGUA tal com l'hem definida abans.

Adoptant una perspectiva lexicològica, analitzarem per separat les principals característiques dels sistemes així caracteritzats, d'accepcions i de trets sèmics oracionals, tot recordant que en TA la complexitat del procediment ve molt determinada per la distància tipològica que separa les llengües implicades. Altrament dit, si les llengües incurses són prou semblants –típicament quan pertanyen a una mateixa família genètica, com portuguès i italià, neerlandès i alemany o mandarí i cantonès–, es podrà utilitzar un sistema de transferència. Per a llengües prou allunyades (entre japonès i hausa o basc, o entre elles i qualsevol de les anteriors), el sistema de transferència serà insuficient o, com a poc, extremament complex. I és a partir d'un cert grau de complexitat –especialment en el pas de solucions lèxico-sintàctiques a solucions semàntiques– quan se sol parlar d'interlingua.

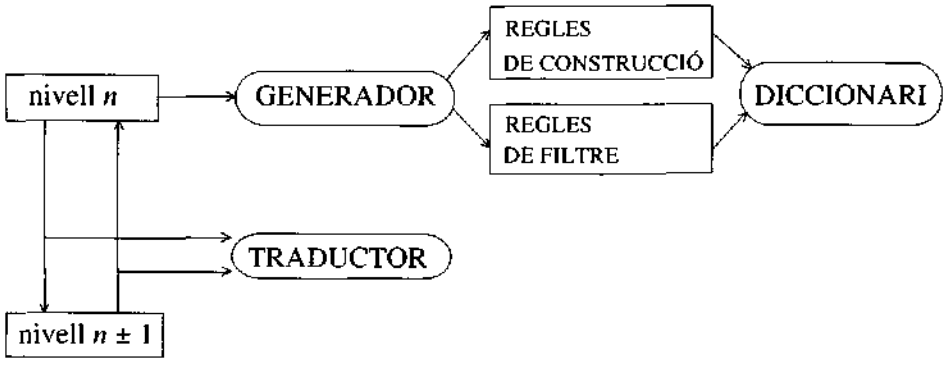
Nosaltres no sols evitarem aquesta confusió terminològica, sinó que aprofitarem la distinció entre 'interlingua' i 'INTERLINGUA' per a al·ludir, respectivament, al processament lingüístic semàntic i conceptual i establir així el límit entre el PLN i la IA. Al primer li dedicarem el pròxims quatre subapartats, i a la segona, un últim apartat sencer.

Un parell d'observacions encara, tornant a l'esquema (10). La part esquerra, la diagonal que va del text d'origen fins a un cert punt en direcció a la INTERLINGUA, indica el procés, denominat d'anàlisi, pel qual el text en qüestió queda interpretat o desambiguat, tècnicament convertit en una certa representació formal o canònica (morfològica, sintàctica...) del seu contingut (vegeu-ne una aplicació a 4.3.2 (21)). La diagonal dreta indica el procés invers, denominat de síntesi o generació, pel qual la representació formal del contingut provinent de l'anàlisi, un cop procesada en alguna mesura pel nexe d'enllaç, de transferència o interlingua, rep una materialització lèxica i és convertida en el text de destinació. Són, respectivament, els dominis del "parser" i del generador. Estrictament parlant, els sistemes de TA es caracteritzen només pel nexe que enllaça l'anàlisi amb la generació de textos, cadascun en llengües diferents. La resta d'elements, és a dir l'anàlisi i la generació són per ells mateixos propis del

PLN en general (per bé que en la TA s'orienten específicament cap al nexe d'enllaç o provenen d'ell, també respectivament).

4.3.1 Els sistemes que treballen sobre accepcions són, en realitat, mecanismes lexicogràfics no trivials que interactuen modularment amb gramàtiques. El "parser" típic d'un sistema així consta d'una sèrie ordenada de *n* nivells o mòduls que confereixen una o més representacions formals a partir de les representacions del nivell anterior o d'un text d'origen ja preeditat (del tipus que veurem a 4.3.2 (21)).[14] Cada nivell està constituït per una gramàtica, que a vegades s'anomena (no massa feliçment) 'generador', composta d'una col·lecció de regles de diversos tipus, i un diccionari associat. Entre els nivells hi ha, com és lògic, nexes interiors, que consten igualment d'una col·lecció de regles, que anomenarem 'traductor', per projectar o transferir les estructures d'un nivell al següent en la direcció que convingui a l'anàlisi o a la generació. Gràficament seria:

(11)




Examinem més de prop el funcionament de les gramàtiques i els diccionaris en un formalisme d'unificació amb un exemple força sumari. Suposem que el text d'origen sigui una oració com *La cartera porta un sobre*. La primera operació verifica si les formes de mot són reconegudes pel diccionari morfològic. Diguem que reconeix efectivament cinc formes, a les quals assigna categories sintàctiques corresponents (aquí simplificades):

(12)

| | | | |
|---|---|---|---|
| (1) | [la] ..... | *Pronom* [3 fs] / *Determinat* [3 fs] |
| (2) | [cartera] ..... | *Nom* [3 fs] |
| (3) | [porta] ..... | *Nom* [3 fs] / *Verb* [3 spr] |
| (4) | [un] ..... | *Pronom* [3 ms] / *Determinat* [3 ms] |
| (5) | [sobre] ..... | *Nom* [3 ms] / *Preposició* |

(14) Per 'preeditar' s'entén la indicació manual en un text de les expressions que cal no processar –noms propis, xifres, fórmules químiques...– o que han de rebre alguna consideració especial –cursiva, nota al peu...–. Es correspon amb la 'postedició', que comprèn el conjunt d'operacions igualment manuals que s'introdueixen en el text que resulta de la generació per tal de fer-lo més correcte o intel·ligible als usuaris humans.

Pel que fa als trets morfològics, totes les formes són singulars, les tres primeres femenines i les altres dues masculines. I pel resultat de l'escorcoll amb els cinc mots es poden produir setze combinacions i, per tant, altres tantes interpretacions morfològiques possibles:

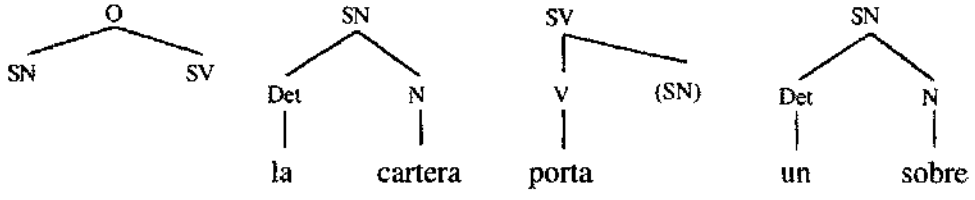
(13) (1) Pronom Nom Nom Pronom Nom

...

(16) Determinant Nom Verb Determinant Preposició

En l'anàlisi sintàctica 'superficial', les regles de construcció componen totes les alternatives combinatòries i les de filtre consulten primer quines de les possibles seqüències són gramaticals o no i després quines, entre les gramaticals, són prioritàries. La primera comprovació consisteix a comparar les cadenes potencials amb un catàleg d'estructures gramaticals oracionals i sintagmàtiques possibles. En el nostre exemple, la seqüència (1) és clarament agramatical, ja que només consta de sintagmes nominals i no correspon a cap estrucura oracional catalogada. Com que cal almenys un sintagma verbal perquè hi hagi una oració, s'ha d'assignar la categoria {verb} a la forma *porta* (l'única que l'accepta, segons l'anàlisi anterior) i descartar les altres combinacions. La (16) també és agramatical, tot i que duu un verb, perquè l'assignació de la categoria {preposició} a *sobre* requereix un sintagma nominal regit, que falta. Al seu torn, cal donar prioritat a l'assignació de categoria {determinant} (i no {pronom}) a les formes *la* i *un* perquè la seva unificació de trets amb els noms *cartera* i *sobre*, respectivament, produeix sintagmes nominals ben formats i més complets, en el sentit que componen una estructura global més compacta, sense formes soltes. De tot això en resulta una estructura oracional consolidada i tres de sintagmàtiques locals, dues nominals i una verbal:[15]

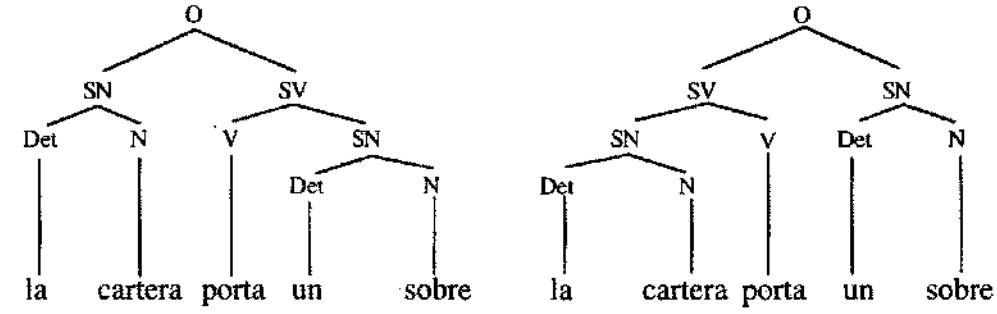
(14)




Comptant que el sintagma verbal tingui o no un sintagma nominal regit (dit d'altra manera, que el verb *portar* funcioni transitivament o no), hi ha quatre possibles estructures oracionals, segons que el primer sintagma nominal sigui ocupat per *la cartera* o per *un sobre* i que *porta* sigui transitiu o no. Les regles de construcció corresponents compondran totes aquestes possibilitats. Les de filtre establiran una prioritat a favor de l'ús transitiu de *porta*, perquè és l'únic que deixa ocupats tots els nusos de l'arbre general. Això produeix, encara, dues possibilitats oracionals:

(15) Els diagrames que segueixen a títol purament il·lustratiu ometen tota formulació de sintaxi d'X amb barres.

(15)




La consulta al diccionari d'aquest nivell d'anàlisi fornirà la subcategorització de *portar*, és a dir el nombre i l'estructura construccional dels seus arguments i els trets dels nuclis corresponents. Una cosa així com:

(16) PORTAR categoria verb; arg1 nom[+A]; arg2 nom[_A]; ...

L'entrada estableix que *portar* és verb que es construeix amb un subjecte nominal (argument 1) marcat amb el tret [+animat] i amb un objecte també nominal (argument 2) neutre per a l'animació.[16] Aquesta dada pot servir per establir, per via unificacional, la funció dels dos sintagmes nominals candidats si es dóna el cas que només un pot dur el tret [+animat]. Consultant les entrades dels nuclis respectius, trobem efectivament:

(17) (a) CARTERA_1 categoria nom; [+A, +H]
(b) CARTERA_2 categoria nom; [-A]

(18) (a) SOBRE categoria nom; [-A]

(aquí ja no compta l'opció de *sobre* com a preposició). L'entrada (17) (a) (equivalent, en certa manera, a una accepció del diccionari convencional) es refereix òbviament a la persona del sexe femení que treballa de repartidora a correus; la segona, (17) (b) als contenidors de mà per bitllets o documents, per anar a col·legi, etcètera. Amb això, arribem finalment a consolidar una estructura oracional única que coincideix amb la primera que hem vist a (15).[17]

4.3.2 Fem una nova aproximació. Comparem ara una entrada lèxica real, la del verb *elevar*, en un diccionari convencional:

(16) La subcategorització real de *portar* inclouria altres arguments (de destinació, etc.), que aquí ometem per simplificar.

(17) Potser que recapitulem algunes qüestions. El tret [+animat] es refereix a la capacitat d'un ésser de moure's per ell mateix i és característica pròpia dels 'animals' (humans o no). A més a més, la nostra gramàtica estableix que *portar* duu subjecte animat i descarta, per tant, no sols la combinació equivalent a **Un sobre porta la cartera*, sinó també **La cartera porta el sobre* entenent *cartera* com a [-animat]. Això obre un seguit de temes teòrics que miraré de resumir tan de pressa com pugui. En primer lloc, no hi ha d'haver cap impediment per sostenir la agramaticalitat de l'última construcció considerada en un registre lingüístic bàsic (on, en tot cas, s'hauria de canviar de verb i dir, per exemple, *La cartera conté un sobre* per a entendre *cartera* en sentit [-animat]). Computacionalment parlant, si més no, és indispensable establir amb claredat aquests tipus de restriccions, que marquen els límits de la gramaticalitat, encara que la gent pugui ampliar-los en altres registres (literaris, humorístics...). D'aquí a una mica ens referirem a una qüestió connexa, la caracterització lexicogràfica dels textos a propòsit de la terminologia. En segon lloc, noteu que aquesta disquisició sobre el caràcter animat o inanimat de *cartera* queda inèdita en el resultat final que forneix el sistema. Si sabem que el sistema entén *cartera* com a [+animat] només és perquè n'hem seguit el procés intern. I si haguéssim de traduir aquest resultat a l'anglès, al francès i a l'alemany, trobaríem en aquestes llengües sengles distincions lèxiques per al tret [-animat]. Molt per sobre:

**elevar** *v* **1** ***a*** *tr* Pujar, transportar a un nivell més alt. *Elevar l'aigua d'un pou mitjançant una bomba.* ***b*** *pron* Pujar, anar cap amunt. *Si no fa vent, el fum s'eleva ben dret.* ***c*** *tr fig* Portar a un càrrec, dignitat, grau, etc, superior. *Elevar un príncep al tron.* ***d*** *pron fig S'ha ele-vat a la presidència il·legalment.* ***e*** *tr fig* Posar més amunt, molt amunt. *Elevar els sentiments del poble.* ***f*** *pron fig El seu pensament s'eleva en especulacions filosòfiques.* ***g*** *tr fig* Transportar a un alt grau de contemplació, superior als sentits. *Elevar l'ànima envers Déu.* ***h*** *pron fig S'eleva en contemplacions místiques.* ***i*** *tr* Dirigir un escrit oficial a una autoritat, persona de categoria superior, etc. *Elevar una petició al govern.* **2** ***a*** *tr* Fer més alt, estendre fins a una major alçària. *La fosa de la neu ha elevat el nivell del riu.* ***b*** *pron En aquest barri industrial, el nivell de la pol·lució atmosfèrica s'eleva un 20%.* ***c*** *tr* Tenir a una certa alçària. *L'alzina elevava les seves branques per sobre dels altres arbres.* ***d*** *pron Aquell cim s'eleva molt més que els altres.* ***e*** *tr* Erigir, construir (un temple, un monument, etc) en honor d'algú. ***f*** *pron* Estar situat, existir en un lloc determinat, un edifici d'una notable alçària, una muntanya, etc. *Al costat del riu s'eleva un gran monestir.* **3** ***a*** *tr* Augmentar en intensitat, grau, quantitat, etc. *La calor eleva el volum dels gasos.* ***b*** *pron Aquesta setmana han elevat els preus dels llegums.* **4 elevar a una potència** *mat* Obtenir la potència d'un nombre. **5 elevar una perpendicular** *geom* Traçar una perpendicular a una recta en un punt determinat.

19. Entrada **elevar** del *Diccionari de la Llengua Catalana* Enciclopèdia Catalana, 1982.[18]

amb les entrades lèxiques que el mateix verb té en el sistema EUROTRA:[19]

**elevar_1=**{cat=v,e_lu=elevar,string=elev,thcat=string,infl=stem,pres_s1=t1,pres_s23p3=t1,pres_p12_inf=t1,impf_type=t1,past_s1=t1,past=s2p12=t1,past_s3=t1,past_p3=t1,futcond_type=t1,pressubj_s123p3=t1,pressubj_p12=t1,impfsubj_type=t1,ger_type=t1,part_type=t1,imper_s2=t1,imper_s3_p3=t1,imper_p1=t1,imper_p2=t1,pron=nil,pronlu=nil}.

**elevar_1=**{cat=v,e_lu=elevar,recform=nil,recpart=nil,fpformcomp=nil,fpformcomp2=nil,fp1type=dest,fp2type=nil,fatr=no,fpas=yes,pron=no}.

**elevar_1=**{cat=v,e_lu=elevar,ersfr=acnpronpas,e_pformarg1=nil,e_pformarg2=nil,e_pformarg3=nil,e_pformarg4=nil,p1type=dest,p2type=nil,control=nil}.

**elevar_1=**{cat=v,e_lu=elevar,e_isrno_'1',eisframe=arg1_2_goal,e_pformarg1=nil,e_pformarg2=nil,e_pformarg3=nil,e_pformarg4=nil,p1type=dest,p2type=nil,semarg1=conc,semarg2=ent,semarg3=ent,semarg4=nil,e_vtype=main,vfeat=nstat,attype=nil,instrumental=yes,term='0',erg=yes}.

20. Entrada **elevar** als diccionaris d'EUROTRA[20]

que corresponen respectivament a quatre nivells de representació o 'estructures', ordenades aquí per a l'anàlisi:

(21) (1) **estructura morfològica**, amb indicació precisa del model pertinent de flexió
(2) **estructura configuracional**, anàlisi sintàctica superficial a base de possibles agrupacions de constituents ordenats tal com estan en el text

| | [+animat] | [-animat] |
|---|---|---|
| **català** | cartera | cartera |
| **francès** | facteuse | portefeuille<br>porte-documents<br>pochette<br>cartable |
| **anglès** | postwoman | wallet<br>briefcase<br>portfolio<br>pocketbook |
| **alemany** | Briefträgerin | Brieftasche<br>Aktentasche<br>Bestand<br>Schulmappe |

La traducció des del català donaria solucions lèxiques úniques en tots els casos, però, tal com està, resultaria inevitablement ambigua si *cartera* quedés caracteritzat només amb el tret [-animat].

(18) Ben mirat, no sempre és fàcil endevinar el raonament lexicològic dels autors, atesa la barreja de sentits i de propietats sintàctiques dels exemples.

(19) EUROTRA és un sistema de traducció automàtica de disseny avançat auspiciat per la Comunitat Europea i desenvolupat entre 1983 i 1992. Comprèn alhora nou llengües, totes pertanyents a la branca occidental indoeuropea: quatre de germàniques (danès, alemany, neerlandès i anglès), quatre de romàniques (portuguès, espanyol, francès i italià) i el grec. Pel seu abast científic, lingüístic i econòmic ha estat, durant aquell període, el més extens entre els seus coetanis. Després de 1992 se segueix desenvolupant amb altres denominacions, formalismes (ALEP) i estructura organitzativa.

(20) En rigor es refereix al verb castellà *elevar* en una accepció (l'entrada '1') equivalent a **1-*i*** de (19).

(3) **estructura relacional**, anàlisi sintàctica amb assumpció (i eliminació visual) d'auxiliars i altres partícules, assignació de funcions oracionals (subjecte, objecte...) i canonització de l'ordre de mots[21]

(4) **estructura d'interfície**, assumpció i eliminació de determinants, preposicions i trets verbals com temps i aspecte, assignació de 'rols' d'estructura profunda (nuclis, arguments...)

El sistema operacional d'EUROTRA és del tipus anomenat de transferència (atesa la relativa similitud tipològica de les nou llengües amb què treballa) i està construït de tal manera que la transferència al mòdul de generació en l'altra llengua sigui el més simple possible i, si pot ser, inexistent. Altrament dit, l'estructura d'interfície que culmina l'anàlisi del text d'origen, escrit en la llengua $L_a$, tendeix a ser tot l'idèntica que es pot a l'estructura d'interfície amb què comença la generació del text de destinació, escrit en la llengua $L_b$. I quan diem idèntica, volem dir que, en aquell punt del procés, només cal substituir les entrades lèxiques de $L_a$ per les corresponents entrades lèxiques de $L_b$ en la mateixa representació arbòria.[22] Això és, tot plegat, el que justifica tantes i tan costoses operacions i el que ens duu amb naturalitat a la definició acurada d''accepció' dintre del PLN.

Acabem de veure a (20) que el diccionari és una col·lecció d'entrades, cadascuna d'elles caracteritzada per un conjunt de trets, el qual no es limita, com en l'exemple de *cartera*, a [±animat], sinó que en comprèn molts altres, segons el sistema, com ara [±humà], [±part_del_cos], [±vegetal], [±comestible], [±objecte], [±massa], [±líquid], [±continent], [±moble], [±localització], [±instrument], [±institució], etc., etc. Al seu torn, la gramàtica és un mecanisme que 1) subministra un catàleg d'estructures arbòries possibles, 2) localitza en el diccionari els mots del text d'origen i els assigna els trets de subcategorització que hi troba, 3) compon totes les estructures arbòries que pot amb els nusos ocupats pels mots del text definits per sengles conjunts de trets de subcategorització, i 4) mitjançant la unificació, acobla els conjunts de trets compatibles i filtra successivament les estructures arbòries obtingudes en virtut de la seva coherència amb les estructures catalogades. Aleshores, cada entrada del diccionari, delimitada per un subconjunt de trets, constitueix un possible nus d'una estructura arbòria o, si es vol, un tipus definit de comportament composicional formalment recognoscible pel formalisme d'unificació. Comparant l'entrada d'aquest diccionari amb la d'un de convencional, com el de (19), s'hi troba una correspondència, almenys tendencial, de la primera amb aquelles (sub)accepcions de la segona que presenten alhora propietats semàntiques i sintàctiques ben definides. Dintre del que els constructors del sistema han decidit o han pogut delimitar, es pot dir que tantes accepcions així com hi hagi, tantes entrades hi haurà en el diccionari computacional. Cal advertir finalment que en PLN la noció d'accepció equival a la d'interpretació, ambdues en els sentits descrits.

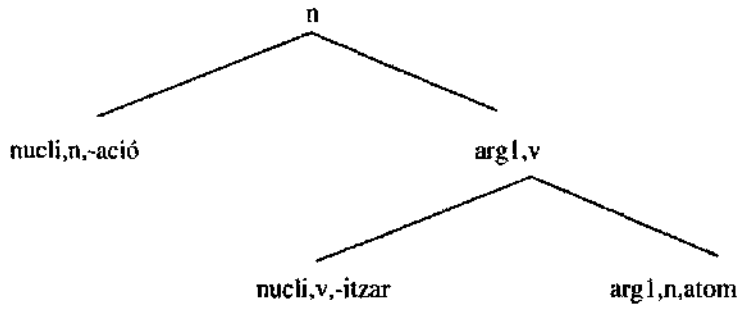
4.3.3 A part d'entrades generals com la que hem considerat a (20), el diccionari

(21) Per assumpció (tècnicament 'elevació') de trets o elements s'entén la incorporació del seu contingut en la matriu estructural que no es visualitza en el diagrama, però que acompanya els arbres locals o generals. Els continguts així assumits poden ser visualitzats de nou en la representació simètrica de la generació, si la llengua $L_b$ els lexicalitza d'alguna forma. I per 'canonització d'ordre' s'entén l'adaptació que el sistema fa a una seqüència fixa d'elements en les estructures locals: primer el nucli i després els arguments, igualment ordenats.

(22) En aquest cas, la transferència consta únicament d'un diccionari de correspondències lèxiques. En circumstàncies menys favorables conté també regles de reajustament.

computacional en conté d'altres que tampoc no coincideixen amb les rèpliques convencionals. Es tracta, d'una banda, de morfemes, clítics i partícules derivacionals que entren en la composició ortogràfica (o fònica, si s'escau) de les entrades generals d'acord amb les corresponents descripcions morfològiques. Mitjançant especificacions del tipus que hem vist a la primera subentrada de (20), corresponent a l'estructura morfològica, es poden descompondre i recompondre les formes superficials dels mots en procés. Aquesta funció la realitza el mòdul fonològic, sovint constituït per un conjunt de regles de dependència no lineals que estableixen no sols l'ortografia (o la pronúncia, en el seu cas) correcta dels mots, sinó també les possibles derivacions composicionals. Les regles fonològiques solen aprofitar l'estructura i la informació que forneix el mòdul sintàctic superficial fins al punt que n'identifiquen els límits en un grau considerable. Així, per exemple, ho reflecteix la composició d'*atomització* i *atomitzar* a partir del nucli nominal *àtom*:

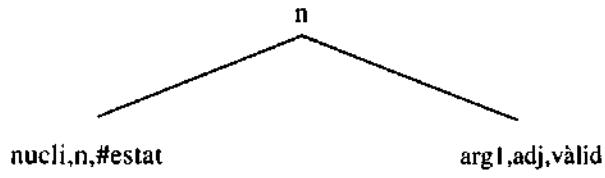
(22)



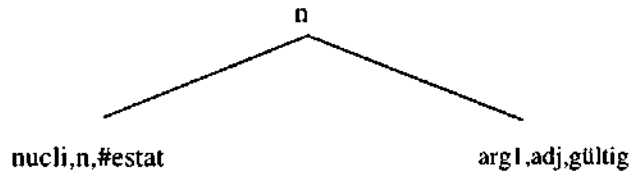

Començant per baix a la dreta, la seqüència *atom*, que pertany a la categoria nom, dóna lloc com a argument a una composició verbal amb un nucli que aporta el sufix *-itzar* i produeix *atomitzar*. Aquesta seqüència derivada intervé al seu torn com a argument en una altra derivació (deverbal) subsegüent en què, mitjançant el sufix *-ació*, produeix el resultat *atomització* amb categoria nom, gràcies també a la cooperació d'altres regles més generals associades que recomponen els marges dels segments afectats i estableixen els valors vocàlics i consonàntics de cada derivació.

Els mateixos principis de morfologia composicional permeten simplificar moltíssim les entrades dels diccionaris bilingües sempre que hi hagi, és clar, prou analogia en els processos derivacionals, com es desprèn del següent exemple sobre la nominalització de l'adjectiu *vàlid* i el seu equivalent, *gültig*, en alemany, on la indicació de significat "estat" selecciona automàticament el sufix adient:

(23) català *validesa*




alemany *Gültigkeit*




D'altra banda, hi ha els mots compostos, en qualsevol modalitat morfològica (*no_obstant, cop_de_mà, portallapis...*). En sistemes de TA per transferència sobretot entre llengües aglutinants i flexives, els compostos solen necessitar un complex tractament especial, atès que moltes solucions lèxiques a $L_a$ requereixen interpretacions sintàctiques a $L_b$ i viceversa (v. gr. alemany *Einbahnstraße* ↔ català *carrer de direcció única*).

Finalment, queda la terminologia (*aldehid, endogàmia, hipotenusa, reostat...*), en què el tractament és tot altre. La seva estricta equivalència lèxica a través de les llengües afavoreix la correspondència de patrons (3.2), molt més econòmica. Això comporta que els termes tècnics formin diccionaris específics d'utilització preferencial, almenys per dues raons: 1) perquè es tradueixen directament mot a mot sense entrar en el processament general del sistema, i 2) perquè poden compondre terminologies temàticament homogènies al marge del diccionari general. D'aquesta manera, l'eficiència dels grans sistemes de PLN augmenta molt significativament si funcionen amb dos diccionaris, un de general, amb els fenòmens gramaticals i les entrades més freqüents de la llengua, i un altre de temàtic incorporat, que recull només la terminologia especialitzada dels textos en procés.[23] Vet aquí la fesomia bàsica d'un diccionari electrònic en aquest nivell de complexitat.[24]

4.4 Molt sovint la distància tipològica entre les llengües naturals implicades en un sistema de TA és massa gran i no n'hi ha prou amb un sistema de transferència, basat, en definitiva, en un escorcoll lexicogràfic no trivial com el que acabem d'analitzar en els paràgrafs anteriors. El sistema de transferència, tot i que incorpora trets semàntics,

(23) La tipologia textual atenua una de les deficiències que caracteritza els sistemes de PLN enfront del parlant humà: la manca de transportabilitat, és a dir la incapacitat de canviar més o menys radicalment de tema sense modificar res en aparença. Això val també per la cobertura gramatical de les regles. Si, per exemple, es processa un text amb instruccions per al muntatge, manteniment i ús d'un vídeo, un sistema optimitzat negligirà fenòmens gramaticals com ara les construccions de passat, les concessives o les condicionals eventuals i irreals.

(24) El terme 'electrònic' s'aplica a vegades per designar qualsevol diccionari convencional o text classificat que es presenta en suport informàtic (programes residents, en CD-Rom, hipertextos, etc.). Es tracta, evidentment, de productes ben diferents. Altres vegades s'aplica, amb una mica més de propietat, a diccionaris amb entrades simples (arrels, derivacions, desinències...) o compostes (modismes, col·locacions...) caracteritzades per uns quants trets morfològics i sintàctics. Poden marcar (semi)automàticament extensos corpus de referència i servir per un seguit de comeses lexicològiques o gramaticals. Són, en tot cas, diccionaris electrònics de baix nivell, comparats amb els que hem vist.

té un fonament essencialment sintàctic en el sentit que radica en la subcategorització d'entrades lèxiques de diccionari, estructures arbòries i funcions oracionals com subjecte, objecte, etc. La seva eficàcia depèn del percentatge de traduccions literals o quasi-literals com les que hem vist a (9). Però no serveix quan no hi ha literalitat o, si més no, tan bon punt com la subcategorització de les entrades resulta insuficient.

Els sistemes que transcendeixen la sintaxi i creen una nova dimensió composicional s'inscriuen en el mètode d'interlingua, com hem dit. Però tot i així, encara s'hi poden identificar graus de complexitat que vénen a delimitar els camps del PLN i de la IA, segons que el processament romangui dintre d'estructures lingüístiques o les ultrapassi –diferència que abans hem descrit oposant 'interlingua' amb 'INTERLINGUA'–. En realitat, són dos trams successius d'una sola direcció, que va d'una representació semàntica fins a una representació conceptual, ambdues obertes a diverses alternatives teòriques. Examinem-ne les principals –sempre des de la perspectiva lexicogràfica, fins on sigui possible– començant per la representació semàntica a base de valències casuals (*case frames*) inspirades inicialment en l'anomenada 'gramàtica dels casos' de Ch. Fillmore.

Si comparem les següents expressions:

(24) (a) Joan obre la porta amb la clau
(b) Joan obre la porta
(c) La clau obre la porta
(d) S'obre la porta
(e) La porta és oberta per Joan
(f) La porta és oberta per la clau

admetrem sense dificultat que comparteixen molta part del significat tot i que es diferencien per l'estructura sintàctica. El subjecte, per exemple, a (a) i (b) és *Joan*, a (c) *la clau*, i a (d), (e) i (f) *la porta*. Però *Joan, la clau* i *la porta* presenten també altres funcions. I també es veu que (e) és la versió passiva de (b), com (f) l'és de (c), i que, per tant, es pot entendre que signifiquen el mateix i que provenen d'una sola representació més 'abstracta'. Nosaltres entendrem que totes aquestes expressions provenen d'un sol significat i d'una mateixa representació, lògicament semàntica.[25]

En un treball ja clàssic, Fillmore va proposar l'adopció de 'casos' semàntics profunds (ben diferents dels declinacionals: nominatiu, genitiu...) per descriure amb precisió la funció dels elements que intervenen en una acció i la seva relació mútua.[26] Així, el significat comú convingut de totes les expressions de (24) ve a establir que si Joan executa l'acció d'obrir la porta amb una clau, *Joan* rep el cas 'agent', *porta* el cas 'objecte' i *clau* el cas 'instrument'. Fillmore va definir encara més casos, com ara 'locatiu' (*al jardí*), 'benefactiu' (*per al gat*), 'direccional' (*cap a fora*) i d'altres que s'han anat incrementant i refinant amb el temps. Una representació abreujada dels exemples de (24) (a)-(f) anteriors amb valències casuals seria:

(25) El supòsit que prenem en aquesta il·lustració no és pas que les expressions de (24) hagin de tenir una mateixa interpretació –cosa que seria falsa–, sinó a l'inrevés: suposant que Joan obre efectivament la porta amb la clau, entenem que aquesta acció pot ser descrita per totes i qualsevol de les expressions de (24).

(26) «Case for case», a Bach, E. & Harms, R. (comps.) (1968) *Universals in Linguistic Theory*, Holt, Rinehart & Winston, pp. 1-88. En contrast amb la teoria transformacionalista, la contribució de Fillmore va palesar immediatament la seva aplicabilitat computacional.

(25)

(a) [OBRIR
[València casual
agent: *Joan*
objecte: *porta*
instrument: *clau*
...]
[modals
temps: present
veu: activa]]

(b) [OBRIR
[València casual
agent: *Joan*
objecte: *porta*
instrument: $\phi$
...]
[modals
temps: present
veu: activa]]

(c) [OBRIR
[València casual
agent: $\phi$
objecte: *porta*
instrument: *clau*
...]
[modals
temps: present
veu: activa]]

(d) [OBRIR
[València casual
agent: $\phi$
objecte: *porta*
instrument: $\phi$
...]
[modals
temps: present
veu: mitja]]

(e) [OBRIR
[València casual
agent: *Joan*
objecte: *porta*
instrument: $\phi$
...]
[modals
temps: present
veu: passiva]]

(f) [OBRIR
[València casual
agent: $\phi$
objecte: *porta*
instrument: *clau*
...]
[modals
temps: present
veu: passiva]]

Noteu que això permet especificar amb molta exactitud les diferències entre les distintes versions sintàctiques de (24) (a)-(f) com a mers canvis de perspectiva que l'acció pren des dels elements incursos i com una elisió d'alguns –possiblement per ignorància o negligència del parlant real–. D'aquí ve que les funcions sintàctiques superficials es permutin entre els elements i que en posició de subjecte hi pugui comparèixer l'agent de l'acció (*Joan*), l'objecte (*la porta*) o l'instrument (*la clau*). Naturalment, tot això té la seva transcendència tant teòrica com metodològica.

Per una banda, aquest procediment comporta augmentar un nivell de representació en el sistema de processament tant analític com generatiu. I encara que la representació mateixa sigui simple, el conjunt de regles requerit per connectar les funcions sintàctiques superficials amb les valències casuals és força complex. En l'exemple d'abans no és gaire difícil establir les valències d'*obrir* si en una certa estructura sintàctica apareix acompanyat de *Joan, porta* i *clau*, gràcies als trets constitutius ben diversificats de les respectives entrades lèxiques (és impensable qualsevol interpretació de tipus **La porta obre la clau amb Joan* i coses per l'estil). Habitualment, però, les estructures són alhora alambinades, ambigües i estructuralment incompletes, i es fa imprescindible harmonitzar molta informació sobre l'ordre superficial, sobre dades aparegudes en altres estructures més o menys allunyades i sobre la possible interacció dels referents implicats en el món que es descriu.

Per la banda contrària, aquest sistema forneix una plantilla molt simple i atractiva per als esquemes verbals primaris, que pot servir, en principi, per gairebé tots els verbs en totes les combinacions oracionals de superfície. L'estructura de valències d'*obrir* coincideix amb la de nombrosíssims verbs que descriuen l'acció d'un agent sobre un objecte mitjançant un instrument: *construir, destruir, pintar, escriure, arreglar, guarir*... Sovint la diferència es redueix al nombre de valències i a l'obligatorietat/opcionalitat/impossibilitat de construir-se amb tals o tals altres valències; v. gr. *preferir* requereix obligatòriament un objecte, *menjar* pot tenir-lo o no i *despertar* no pot tenir-lo. Comproveu-ho.

En conjunt, el profit extret a aquestes estratègies ha estat d'allò més remarcable. Comptant només les valències obligatòries (elidides o no, en la parla), no sols es pot bastir una classificació entre els verbs de València zero (*ploure*), u (*tossir*), dos (*voler*), tres (*regalar*) o quatre (*portar*), sinó que també es poden identificar procediments lèxico-sintàctics per passar d'una classe de verbs a la següent i viceversa.[27] Tal és el cas de la 'causativitat', una mena de recursió d'agentivitat amb el pro-verb *fer* segons una fórmula de tipus [X fa que *Y*] (on X és un agent i *Y* una construcció verbal de València *n*). Per exemple:

(26) (a) [[Joan] es $mor_1$] + [caus]: [Pere fa que [[Joan] es $mori_1$] ⇒ [Pere] $mata_2$ [Joan] + [caus]: [Carles fa que [Pere] $mati_2$ [Joan]/[Carles] *fa matar* [Joan] + [caus]: [Albert fa que [Carles faci que [Pere] $mati_2$ [Joan]/[Albert] *fa fer matar* [Joan], etc.

(b) [[El gos] $passeja_1$] + [caus]: [Pere fa que [[el gos] $passegi_1$] ⇒ [Pere] $passeja_2$ [el gos]

(c) [[Joan] $té_2$ [un llibre]] + [caus]: [Pere fa que [[Joan] $tingui_2$ [un llibre]] ⇒ [Pere] $dóna_3$ [un llibre] [a Joan]

d'on es desprèn que *matar* és el causatiu lèxic de *morir* (tal com *fer matar* és el causatiu analític de *matar*, *fer fer matar* l'és de *fer matar*, etc.), que $passejar_2$ és el causatiu lèxic de $passejar_1$ (però sense lexicalització) i que *donar* l'és de *tenir*, al marge d'altres propietats sintàctiques (com ara que *tenir* no sigui pròpiament transitiu i *donar* sí).[28]

Explotant l'enorme avantatge metodològic d'establir relacions d'aquest tipus entre esquemes de valències verbals, R. C. Schank i d'altres que treballaven en línies més o menys similars van fer veure que es podia reduir molt més encara el nombre de representacions bàsiques i que si aquestes eren prou abstractes es transcendien les estructures pròpiament gramaticals de les llengües naturals i s'arribava a una configuració veritablement INTERLINGÜÍSTICA dels significats, a la qual tendirien totes les expressions reals dels parlants.

Aquesta aproximació conceptual de les llengües ens introdueix en el domini de la IA.

(27) S'ha discutit força sobre el nombre màxim de valències obligatòries (més o menys equivalents als arguments de 4.3.1) i sembla que hi ha raons per deixar-lo en quatre: compareu [$\phi$] $plou_0$, [Joan] $tus_1$, [Joan] $vol_2$ [una moto], [Joan] $regala_3$ [fruita] [a la Maria], [Joan] $porta_4$ [llegums] [de l'hort] [a casa].

(28) Un altre dimensió transcendental és que l'examen aprofundit entre les valències dintre els esquemes verbals ha revelat característiques sèmiques –distincions entre estats, events, processos tèl·lics o no, etc.– que interactuen d'una manera molt precisa i prou complexa amb categories tan diverses com la *Aktionsart*, el temps i l'aspecte.

## 5. LEXICOGRAFIA I FORMALISMES EN IA

5.1 La idea de Schank era una versió reduccionista de les valències casuals, anomenada 'dependència conceptual', especialment dissenyada per representar esquemes prototípics d'acció. Un exemple clàssic lliga expressions com a) *Joan va donar un llibre a Maria* i b) *Maria va agafar un llibre a Joan* a través d'una sola relació conceptual de transferència batejada amb el nom de 'ATRANS' (*Abstract TRANSfer*). Formalment, i passant per alt els trets irrellevants:

| | | | |
|---|---|---|---|
| (27) | (a) | [ATRANS | (b) [ATRANS |
| | | relació: *possessió* | relació: *possessió* |
| | | actor: *Joan* | actor: *Maria* |
| | | objecte: *llibre* | objecte: *llibre* |
| | | origen: *Joan* | origen: *Joan* |
| | | recipient: *Maria* | recipient: *Maria* |

A més a més d'establir una relació de possessió, la transferència pot ser de localització, de pertinença, etc. Definida en general, la transferència es troba en la base significativa d'una bona multitud de verbs (o categories similars) en totes les llengües naturals, en el sentit que captura almenys les inferències veritatives essencials de les respectives proposicions. La dependència conceptual queda reduïda a una col·lecció d'accions primitives com ara:

(28)
- (a) ATRANS, transferència abstracta
- (b) PTRANS, canvi físic de localització
- (c) MTRANS, transferència mental d'informació
- (d) MBUILD, creació d'una nova idea o d'una conclusió a partir d'una informació
- (e) INGEST, introducció d'una substància dins el cos
- (f) PROPEL, aplicació d'una força sobre un objecte
- (g) SPEAK, producció de qualsevol mena de so

etc.

cadascuna acompanyada dels esquemes pertinents per copsar-ne subaccions igualment primitives, és a dir universals, sota el principi que hi ha efectivament moltes relacions lògiques –de tipus tot/part, hiperonímia/hiponímia, etc.– independents de qualsevol circumstància lingüística, cultural o històrica.

Examinem-ne algunes conclusions des del punt de vista de la lexicologia i de la confecció de diccionaris. En primer lloc, cal reconèixer que el procediment de la dependència conceptual permet paradoxalment recuperar valors fixos en textos oberts, cosa que només es pot pressuposar en textos tancats d'aplicació força limitada (recordeu 3.2-3). Sens dubte, és un avantatge computacionalment crucial. La trajectòria que ho

assoleix tanmateix no sols és llarga i complexa per ella mateixa, sinó que transcendeix l'àmbit del lèxic, ja que, en rigor, el diccionari consta de la col·lecció d'entrades de (29) i dels esquemes associats, i aquests valors no sols no es corresponen a res que s'assembli al lèxic morfològicament identificable de cap llengua natural, sinó que no hi són ni en són directament projectables. Ans al contrari, a la vista del que hem comentat a la secció anterior, la dependència conceptual no invalida el processament composicional morfològic i sintàctic, sinó que, almenys en teoria, l'implica. Això comporta afegir un algorisme, generalment força complex, per connectar, avant i enrera, representacions extralingüístiques amb estructures lingüístiques.

5.2 La situació, en aquest punt, es torna paradoxal. Per una banda, l'adopció d'una representació per conceptes transcendeix l'estructura purament gramatical dels textos i assumeix el coneixement dels parlants sobre l'entitat dels referents i les possibles relacions que hi ha entre ells. És una nova dimensió, inèdita a la que hem vist fins aquí, que es troba en el centre mateix de la recerca en IA. Per altra banda, això crea encara nous vincles de relació amb les estructures lingüístiques implicades atès que comporta la comprensió integral dels textos en procés. Els models oracionals no basten i cal recórrer a opcions supraoracionals –en dominis propis de la pragmàtica i la lingüística textual– que puguin interpretar inferències, pressuposicions, implicacions, etc. enmig d'estructures farcides de vaguetats, ambigüitats, elisions i fets indirectes de parla. La paradoxa sorgeix perquè, avui per avui, com més augmenta el poder teòric dels models (en fer-se més complexos i assumir més fenòmens), més en minva el pràctic, en el sentit que el seu abast d'aplicació es redueix quasi proporcionalment.[29]

Una forma habitual no tant de resoldre com de simular aquest repte consisteix a definir un àmbit, uns objectes i una xarxa de possibles relacions entre ells, seguint més o menys el primitiu model de SHRDLU, creat el 1972 per T. Winograd (amb un pàrsing d'ATN). Dintre d'un espai tridimensional determinat s'hi defineixen objectes geomètrics de diversa aparença –com cubs, esferes i piràmides– de dimensions, propietats físiques i situació relativa igualment determinada. El sistema així construït és capaç de reconèixer no sols construccions agramaticals (imprevistes en la producció de regles sintàctiques), sinó també 'impossibles', com una que pretengués col·locar una bola sobre una piràmide.

És, com deia, el domini central de recerca en IA i els seus formalismes peculiars, que comprenen sobretot les xarxes semàntiques amb multitud de modalitats: 'scripts', 'frames', 'mops', 'memettes', etc. Un domini certament apassionant obert a estratègies agosarades (com les xarxes neuronals) que supera de molt l'àmbit de la lexicografia lingüística, tot i que el travessa de cap a cap amb qualsevol dels formalismes que hem esbossat.

(29) Hi ha propostes que no passen de ser el que s'anomena 'gramàtiques de joguina', sense cap altra aplicació pràctica que servir de base d'experimentació o d'exercici pedagògic.

## 6. SIGLES D'ÚS MÉS GENERALITZAT

IA intel·ligència artificial
LC lingüística computacional
LT lingüística teòrica
PLN processament del llenguatge natural
TA traducció automàtica

RAMON CERDÀ
*Universitat de Barcelona*

## REFERÈNCIES BIBLIOGRÀFIQUES

ALSHAWI, H. (1987) *Memory and Context for Language Interpretation*. Cambridge, Camridge University Press.

BARR, A.; FEIGENBAUM, E. A. (comps.) (1986) *The Handbook of Artificial Intelligence*. Reading Mass., Addison-Wesley. Especialment el Cap. IV, Understanding Natural Language, pàgs 223-321.

BOGURAEV, B. & BRISCOE, T. (1989) *Computational Lexicography for Natural Language Processing*, London, Longman.

CHAFFIN, R. (1988) «The nature of semantic relations: a comparison of two approaches». A: EVENS, M. W. (comp.) (1988) pàgs. 289-334.

DOWTY, D. R. (1979) *Word Meaning and Montague Grammar. The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. London, Reidel.

DOWTY, D. R.; KARTTUNEN, L.; ZWICKY, A. M. (comps.) (1985) *Natural Language Parsing. Psychological, Computational, and Theoretical Perspectives*. Cambridge, Cambridge University Press.

EVENS, M. W. (comp.) (1988) *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks*. Cambridge, Cambridge University Press.

GRISHMAN, R. (1986) *Computational Linguistics. An Introduction*. Cambridge, Cambridge University Press.

MARKOWITZ, J. (1988) « An exploration into graded set membership». A: EVENS, M. W. (comp.) (1988) pàgs. 239-60.

MEL'CUK, I. (1988) «The explanatory combinatorial dictionary». A: EVENS, M. W. (comp.) (1988) pàgs. 41-74.

MEYA, M.; HUBER, W. (1986) *Lingüística computacional*. Barcelona, Teide.

NIRENBURG, S. (comp.) (1987) *Machine Translation. Theoretical and Methodological Issues*. Cambridge, Cambridge University Press.

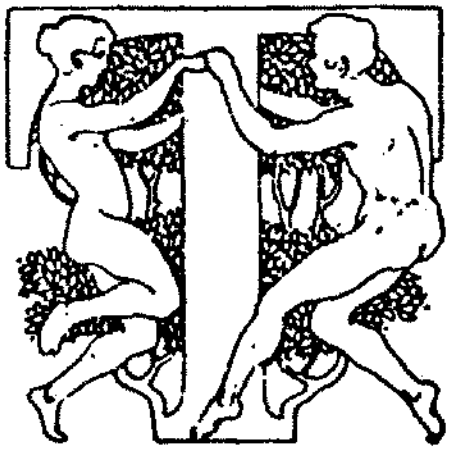
SELLS, P. (1985) *Lectures on Contemporary Syntactic Theories. An Introduction to Government and Binding Theory, Generalized Phrase Structure Grammar, and Lexical-Functional Grammar*. Stanford, Center for the Study of Language and Information. [Traducció en castellà (1989) *Teorías sintácticas actuales*. Barcelona, Teide].

SHIEBER, S. M. (1986) *An Introduction to Unification-Based Approaches to Grammar*. Stanford, Center for the Study of Language and Information. [Traducció en castellà (1989) *Introducción a los formalismos gramaticales de unificación*. Barcelona, Teide].

SLOCUM, J. (1988) *Machine Translation Systems*. Cambridge, Cambridge University Press.

SOWA, J. F. (1984) *Conceptual Structures: Information Processing in Mind and Machine*. Reading Mass., Addison-Wesley.

WINOGRAD, T. (1983) *Language as a Cognitive Process*, Volume I: Syntax. Reading Mass., Addison-Wesley.



M. PAZ BATTANER

# L'ÚS DELS DICCIONARIS ESCOLARS[1]

L'article es proposa oferir uns objectius per tal d'aconseguir l'ús satisfactori del diccionari escolar. La proposta parteix de la teoria lexicogràfica i de les obsevacions de quatre casos en una situació escolar.

Així, repassarem què és el diccionari escolar i quin és el seu ús més representatiu; com ha de ser el diccionari escolar per aconseguir-ne un ús més profitós.

## 1. QUÈ ÉS EL DICCIONARI ESCOLAR

Un diccionari escolar és un instrument d'informació monolingüe adaptat (a) als objectius de l'ensenyament d'una llengua materna, pensat per realitzar activitats i per resoldre exercicis programats pels professors; alhora que (b) esdevé un instrument de consulta autònoma, encara que aquesta funció estigui sotmesa a la primera. Aquestes dues funcions condicionen l'instrument i, en els pròlegs o introduccions, solen presentar-se en un ordre invers del que s'ha exposat aquí.

L'escola ensenya a escriure i a fer servir l'escriptura per a finalitats concretes. Més tard, l'adolescent utilitza l'escriptura per als interessos propis: escriure un diari, apuntar els vídeos que té, retenir les lletres de les cançons preferides... En efecte, l'escola ensenya a utilitzar el diccionari i, després, l'adolescent el pot utilitzar per als seus interessos.

(1) Traducció de Rosa Estopà Bagot, amb la supervisió de Montserrat Ribas Bisbal.

*Caplletra 17* (Tardor 1994), pp. 109-124

En la literatura lexicogràfica espanyola d'aquests darrers anys, s'han fet descripcions i anàlisis dels diccionaris actuals que porten l'adjectiu *escolar* o d'altres semblants.[2] Aquestes descripcions i la crítica implícita parteixen de la teoria lexicològica de J. Casares i d'uns pressupòsits pedagògics que no es concreten explícitament.[3]

H. Hernández (1989), Olarte Estampa i Garrido Moraga (1984), Martínez Marín (1991) i Costa, Viladrosa i Vilaró (1992) han aportat descripcions, han plantejat mètodes analítics i han suggerit crítiques als diccionaris que es presenten com a apropiats per a l'ensenyament.[4]

Les anàlisis citades s'ocupen de la macroestructura: del repertori lèxic o llistat que ofereixen els diccionaris analitzats, però principalment s'ocupen de la microestructura: del text de la definició, de les accepcions i de l'ordre, de les informacions contingudes, de les referències intratextuals i dels exemples.

Hartmann (1987b:122) remarca, molt encertadament, que per analitzar un diccionari és fonamental qüestionar-se les funcions que es proposa acomplir: qui el necessita, què conté, per a què s'utilitza, com es troba la informació i on i quan s'usa. Quan es tracta de diccionaris escolars s'hauria, doncs, de considerar l'ús a l'aula i l'ús espontani i individual; també s'hauria de tenir en compte el punt de vista dels destinataris com a usuaris autònoms o com a alumnes aplicats, i el punt de vista dels professors com a monitors d'activitats a la classe.

L'anàlisi dels diccionaris escolars des d'aquest punt de vista empíric encara no s'ha fet.

## 2. QUIN ÉS L'ÚS DEL DICCIONARI ESCOLAR

«Enseñar a manejar los diccionarios debería (...) constituir un objetivo destacable de la didáctica de la lengua».[5] Quan passem de l'obligatorietat potestativa a la simple realitat, constatem que la descripció de l'ús dels diccionaris en el medi escolar no està gaire avançada. Pel que fa a l'ús dels diccionaris, el projecte més pretensiós és el del professor Hartmann. Aquest treball se centra en els diccionaris, tant monolingües com bilingües, des d'una perspectiva de l'aprenentatge de la llengua estrangera.[6]

En relació amb la literatura que analitza o planteja aquest tipus de lexicografia escolar, s'acostuma a citar tant l'adquisició de les habilitats de consulta[7] com l'adaptació al currículum habitual de la llengua materna a les escoles –qüestions que tenen molts aspectes comuns–. Ara bé, el punt central d'aquests estudis és l'anàlisi del producte lexicogràfic i la crítica que se'n desprèn.

Entre tots aquells que estudien la promoció de l'ús del diccionari sobresurt Garrido Moraga (1991), el qual suggereix que els alumnes cal que tinguin un coneixement de lexicografia pràctica, cal que coneguin «la planta» del diccionari i, alhora, és necessari que entrin progressivament en la definició. Per tot això s'ofereix un programa

(2) Martínez Marín (1991: 58) recull els adjectius que apareixen en els títols dels diccionaris per a l'ús escolar.

(3) Els pròlegs dels diccionaris escolars acostumen a estar avalats per lexicògrafs prestigiosos, F. Lázaro Carreter en els diccionaris Anaya (1979 i 1991), G. Salvador Caja en l'últim diccionari de Santillana (1991), S. Gili i Gaya en els de Bibliograf, i a partir de 1976 en el de Vox. La lexicografia escolar es desenvolupa a partir de la *Ley General de Educación* de 1970.

(4) Costa, Viladrosa i Vilaró proposen una activitat formativa per a la formació del professorat perquè conegui els principals diccionaris catalans monolingües, precisament no escolars.

(5) Aquesta afirmació de H. Hernández (1989) és sostinguda per molts altres autors, Alvar Ezquerra 1982, Fontanillo 1983, Olarte Estampa i Garrido Moragas 1984.

(6) Hartmann (1989:103); i també Ibrahim i Zalessky (1989).

(7) Garrido Moraga (1991: 37-40) proposa uns punts de lexicografia i de lexicologia per a l'aplicació del diccionari a l'ensenyament de la llengua.

d'activitats que sembla que s'hagi pensat per a alumnes d'ensenyament secundari.[8]

Pel que fa al cicle inicial, Fontanillo (1983) ofereix una sèrie d'exercicis per saber utilitzar el diccionari, més concretament, per aprendre l'ús de l'ordre alfabètic –cosa que ocupa la part del lleó (pàgs. 65-79)– i explica un cas pràctic per comprendre les claus i les abreviatures d'un diccionari (pàgs. 80-86). El treball acaba amb la descripció d'una classe de primera etapa d'EGB en què s'ensenya què és un diccionari, es fan exercicis de recerca de paraules i es treballa el sufix *-ista* a partir de paraules que el professor suggereix i per a les quals no es proposa la consulta al diccionari.

L'exigència de Hartmann d'observar per tal de conèixer quin és l'ús real dels diccionaris presenta uns obstacles de difícil solució, si més no actualment. Les enquestes,[9] els protocols d'ús, l'observació exterior, la informació recollida i l'anàlisi de les guies proporcionen dades que són delicades d'interpretar per elles mateixes, ja que sempre apareixen tenyides de l'habilitat i de la competència de l'usuari, que sovint té una experiència anterior difícil de conèixer.

A fi d'avançar per aquest embrollat camí, tot seguit es presenten les observacions d'una situació de consulta lexicogràfica.

L'experiència, que és l'objecte d'observació i el punt de partida d'aquest treball, consisteix en unes proves que van passar-se a quatre escolars de setè curs d'Educació General Bàsica, d'un col·legi públic de la ciutat de Barcelona. Els escolars tenien entre dotze i tretze anys quan van ser observats. Els alumnes es van agrupar de dos en dos per realitzar les proves.[10] Se'ls va gravar en vídeo i es van recollir les respostes, fins i tot en determinats casos se'ls va demanar que les escrivissin.

Les observacions presenten unes limitacions estrictes. Només es van analitzar quatre casos –i els quatre del mateix centre encara que de diferent grup– en una situació netament d'observació, en un ambient escolar i sobre qüestions relacionades amb activitats principalment escolars.

La prova exigia des de l'inici l'ús del diccionari. Es van proporcionar diccionaris als escolars. No es va considerar, doncs, l'ús espontani ni tampoc no es va plantejar el reconeixement de problemes lèxics. Les qüestions que els escolars havien de resoldre presentaven paraules que ja havien estat aïllades o subratllades, cosa que no amaga el problema, que va denunciar Chall (1987), de la capacitat d'aïllar veus no conegudes.

La prova (vegeu apèndix) té uns objectius precisos i delimitats: la possibilitat d'extreure dades sobre les habilitats de localització de veus i d'accepcions precises, i sobre la informació obtinguda, per tal de satisfer la comprensió, facilitar l'ús i resoldre qüestions de la programació escolar.

Les veus de la prova es van limitar a 1) substantius coneguts, referenciables físicament, ***elefante, abertura, prensa;*** 2) substantius desconeguts, referenciables físicament, ***rebeco***; 3) substantius abstractes, no refenciables físicament, ***apertura, obertura, quietud;*** 4) adjectius i verbs, ***ejercer, ejercitar, listo-a;*** 5) paraules gramaticals (nexes, ***por supuesto,*** i determinatius, ***nuestro-a***; 6) formes morfològiques irregulars, ***conduje.***

(8) Per a l'ús del diccionari entre estudiants de batxillerat, vegeu Alvar Ezquerra 1982.

(9) L'enquesta que presenta el treball de H. Hernández (1989: 110-116), segons paraules d'ell, és iniciadora d'aquest enfocament: «Sólo pretendemos iniciar un nuevo camino», «somos conscientes de las limitaciones de nuestro estudio». L'enquesta està plantejada per trobar les deficiències dels diccionaris. Es pregunten «problemes» (pregunta 3a del primer qüestionari i pregunta 2a del segon qüestionari) i no pas «facilitats» o «solucions» per als usuaris. Es pregunten «deficiències (pregunta 2a del tercer qüestionari) i no pas «aportacions» per als professors.

(10) Manifesto el meu agraïment al Col·legi de Pràctiques número 2 de Barcelona i a la seva direcció, que em va facilitar la realització de la prova. La col·laboració del professor de l'àrea de llengües, Jaime Fernández Pérez, va permetre la materialització de l'experiència amb l'elecció dels alumnes i la creació d'un clima d'interès i de normalitat que era necessari. Javier Lahuerta va ajudar a gravar en vídeo i va aportar el seu interès pels diccionaris d'aprenentatge.

Els escolars observats van utilitzar dos diccionaris de característiques clarament escolars: *Diccionario Anaya de la Lengua* (1991) i *Diccionario Esencial Santillana de la Lengua Española* (1991). L'ús de dos diccionaris durant les proves tenia com a objectiu un camp d'observació més ampli i no tant una comparació entre els diccionaris.

L'anàlisi posterior del vídeo, de les notes de camp que vam anotar i de les solucions que els escolars van aportar sobre les qüestions que havien de resoldre amb l'ús del diccionari, permeten treure algunes apreciacions, que no poden considerar-se conclusions atès el caràcter d'«estudi de casos» que l'experiència presenta.

Els escolars de setè d'EGB d'un col·legi públic de la ciutat de Barcelona que van ser observats:

a) dominen les habilitats de localització de les entrades, és a dir, dominen l'ordre alfabètic dels diccionaris semasiològics.

b) no dominen el reconeixement d'accepcions (és el cas de 3 ***ejercer/ejercitar,*** de 4 ***supuesto,*** i de 6 ***listo***, aquest últim cas, però, el resolen d'una manera més satisfactòria).

c) saben que poden resoldre dubtes de morfologia flexiva (és el cas de 7 en què es demana una forma del pretèrit simple de ***conducir***, i no vacil·len quan han de buscar-la; encara que en un dels diccionaris utilitzats els resulta difícil perquè s'ha d'aconseguir la informació en un apèndix a partir del mot ***producir***).

d) desconeixen a vegades l'ús de remissions dintre dels articles (tenen dificultat quan han de trobar concretament la informació de 7, la forma del pretèrit simple de ***conducir*** a partir de la forma ***producir***).

e) descodifiquen les paraules i les accepcions que coneixen de manera força satisfactòria (és el cas d'***elefante***, paraula amb què construeixen diverses frases; això no obstant, no aconsegueixen entendre la frase «*elefante blanco*» del DUE «*cosa que cuesta mucho mantener y no produce utilidad*»).

f) descodifiquen d'una manera aproximada quan no coneixen la paraula (és el cas d'1 ***rebeco*** que ha estat observat mitjançant una codificació poc precisa «*el cazador mató al rebeco*», «*el rebeco es un animal que vive en zonas de bosque*»).

g) codifiquen aprofitant la informació del diccionari quan coneixen la paraula (són els casos d'1 ***elefante***: «*el hombre mata el elefante para vender el marfil de sus colmillos*», «*el elefante coge cacahuetes con la trompa*», i de 2 ***abertura, apertura, obertura*** encara que amb alguns dubtes i algunes limitacions.

h) codifiquen amb errors evidents o fan servir sintagmes prefixats, quan no coneixen la paraula o no els és habitual ( és el cas d'1 ***rebeco***: «*con un rebeco limpié el sillón que estaba manchado*» cosa que aconsegueixen mitjançant la identificació lexicogràfica *rebeco=gamuza;* el cas de 2 ***apertura***: «*es un local que lo inaguran*», «*la apertura de un local*» a partir de la construcció «*apertura de un local*» dels mitjans de comunicació; o el cas de 3 ***ejercer***: «ejercer el poder»).

i) aconsegueixen de reunir un vocabulari acceptable sobre una paraula que se'ls ha donat com a centre d'interès (és el cas de 5 ***quietud*** i ***prensa***, dels sinònims i antònims de 6 ***listo-a***; per demostrar-ho podem prendre com a referència l'explicació de la seva tasca: *«te salían sinónimos y antónimos...los buscábamos y daban otros»*).

j) desconeixen o tenen grans dificultats per comprendre les abreviatures (és el cas d'1 ***elefante, rebeco***, substantiu, de 3 ***ejercitar*** com a *prnl.*, pronominal, de 4 ***supuesto*** substantiu i locució adverbial, i de 8 ***nuestro-a***, adjectiu i pronom possessiu. Ara bé, són conscients que hi ha una llista al començament del diccionari que les clarifica, perquè les hi busquen quan detecten algun problema.

k) no saben fer servir les informacions de la categoria gramatical (en el cas de 8 ***nuestro-a***, arriben a identificar les abreviatures o a llegir-les en un diccionari que no les fa servir per aquest cas, però no dominen de manera operativa els conceptes d'adjectiu i de pronom, així, per exemple, comprenen millor l'expressió substantivada *«los nuestros»* que no pas l'adjectiu o el pronom ***nuestro-a***; contràriament sembla que dominin el concepte de «determinant» que no és usual als diccionaris. En el cas de 3 ***ejercitar***, verb transitiu, intransitiu, pronominal, les informacions «els sonen» a anàlisi, encara que no els són operatives).

D'aquestes observacions es dedueixen dues línies de continuïtat: com millorar els diccionaris i com aconseguir un ús més satisfactori d'aquests diccionaris. Tot seguit abordarem la segona qüestió a partir de la primera.

## 3. QUÈ HA SER EL DICCIONARI ESCOLAR

Aquesta és una pregunta a la qual responen, amb més o menys incidència, tots els treballs de lexicografia i sobretot els de lexicografia teòrica.[11] La reflexió pròpiament teòrica porta a plantejar-se unes preguntes bàsiques: quina és la informació i quin és el coneixement satisfactori dels parlants sobre les unitats lèxiques en un estat de llengua determinat, i com tot això es pot reflectir en un diccionari.

Així, què és una unitat lèxica dins d'un enfocament lingüístic també és una qüestió bàsica per a la teoria i per a la pràctica lexicogràfiques.[12]

Les veus (o unitats lèxiques o lexies) són entitats lingüístiques (és a dir, fonètiques, gramaticals i semàntiques) generalment opaques, no regulars, amb un significat prou estable com per, tot i mantenir-se, poder variar de sentit i construcció en funció dels diversos contextos.

Les veus tenen una estructura interna que s'organitza amb la intervenció de diferents components: la realització fonètica, la distinció fonològica, el comportament oracional, els components i les variacions morfològics, el significat estable i adaptable, els valors de registre, i les prescripcions pragmàtiques. En cada un d'aquests aspectes s'estableixen relacions paradigmàtiques i sintagmàtiques amb altres veus.[13] Tot això fa que

(11) Títols com *El diccionario que deseamos*, *El diccionario ideal*, *Los diccionarios escolares del español: lo que son y lo deberían ser* patentitzen aquesta orientació. Sota una altra mena de títols també trobem expressions d'exigència ideal que van unides a projectes lexicogràfics. A més, tot el que concerneix al diccionari presenta un discurs netament pedagògic que sempre és «desiderat iu».

(12) En el número especial de *Le Français dans le Monde*, A. H. Ibrahim i M. Zalessky (coords) 1989, Greimas, Martinet, Pottier i Gross contesten la pregunta «Qu'est-ce qu'un mot», pàgs. 58, 84, 120, 207, respectivament.

(13) J. Aitchinson (1987) recull els aspectes psicolingüístics del lèxic en el treball fonamental *Words in the Mind*.

l'estructura d'una veu sigui molt complexa, lingüísticament i conceptualment, i que, a més, el coneixement d'una unitat comporti el domini de moltes altres veus.

Pel que fa al processament de les unitats lèxiques d'una llengua, el fet de tenir una gran memòria és fonamental ja que els inventaris de veus arriben a contenir milers d'unitats, a més, dintre de cada unitat s'emmagatzemen altres informacions i entre totes elles s'estableixen relacions.

Tradicionalment l'objecte que ha representat aquesta memòria lèxica dels parlants d'una llengua ha estat i continua sent el diccionari, el suport més habitual del qual és el d'un llibre gruixut.

De totes les relacions que les unitats estableixen, les fòniques-ortogràfiques, pel que fa a la nostra tradició escrita, i les morfològiques són les més senzilles per característiques de regularitat; les semàntiques, les sintàctiques, les de registre i les pragmàtiques presenten molta varietat, per això es consideren les relacions purament lèxiques. Són aquestes relacions idiosincràtiques les que els diccionaris intenten proporcionar de manera més explícita i descriptiva.

La categoria lèxica de les veus sembla lexicogràficament obligada en totes les entrades i en la majoria d'accepcions, ja que indica el comportament morfològic i sintàctic de la veu. Aquesta càrrega informativa, que exigeix la seva presència en cada article del diccionari, fa que per indicar-la les editorials recorrin a les abreviatures.

Les informacions morfològiques, les fòniques i les ortogràfiques només apareixen explícitament en els casos en què es presenta una irregularitat palesa,[14] cosa que sembla, avui dia, una qüestió obligatòria en els diccionaris «escolars», que tendeixen a assemblar-se als diccionaris de dubtes.

Contràriament, no passa el mateix amb les seleccions sintàctiques que les veus admeten ja que els diccionaris encara estan molt lluny de poder informar-ne, si més no des d'un punt de vista econòmic i efectiu.[15]

El significat de les veus fluctua entre el signe referencial monovalent i el magatzem d'informacions culturals. Aquests últims anys s'ha revitalitzat la denominació de *significat conceptual* per representar aquest aspecte del significat lèxic que va des de la terminologia fins als significats culturals que hi ha emmagatzemats en una veu com per exemple *«fiesta»* (Wierzbicka 1985).

Entre els enfocaments més recents sobre la unitat lèxica com a entitat lingüística interessen també, des del punt de vista de la lexicografia escolar, aquells que consideren l'adquisició de les veus i la consciència psicolingüística (Aitchinson 1987).

Aquest darrer punt es relaciona amb el nombre d'entrades i accepcions que un diccionari escolar ha de presentar. Aquest aspecte alhora enllaça amb els textos que arriben als escolars, cosa que ja ha estat considerada en l'àmbit de l'ensenyament del vocabulari.[16] La teoria i la pràctica lexicogràfica fluctuen entre limitar les entrades, que és el cas dels diccionaris de Mèxic,[17] i estendre el repertori per no decebre el possible consultant. En general els diccionaris escolars de l'espanyol prefereixen la darrera opció.

(14) Ahumada Lara (1989: 57) recomana «como imprescindible» la informació gramatical en un diccionari escolar.

(15) El projecte lexicogràfic de Mel'cuk (1982) té molt en compte les combinacions sintàctiques de les veus i comença a tenir adaptacions al món de l'aprenentatge de les llengües maternes (Leed i Nakhimovsky 1989, Steele 1989), a més de versions bilingües (Lepinette 1989).

(16) El número d'unitats lèxiques que els nens i adolescents coneixen, així com també l'establiment del vocabulari que necessiten conèixer ha estat objecte d'atenció en l'ensenyament del vocabulari des des fa molt temps; podem agafar com a exemple l'obra de Thorndike des de 1921 (Thorndike i Lorge 1944) i el treball del MEC (1989), coordinat per M. A. Casanovas i M. Ribera per a l'espanyol d'Espanya.

(17) Lara, Ham i Hidalgo (1979).

Beck i alii (1987) caracteritzen en tres subconjunts el vocabulari que arriba als escolars. Primerament, les paraules bàsiques que –segons ells– no s'han de treballar perquè ja són conegudes; segonament, les paraules específiques que s'han d'introduir amb temes específics d'estudi propis d'assignatures concretes; i finalment, les veus d'utilitat general (***conveniente, torcido, tortuoso, incidencia,*** *etc.*) que són les que han de servir de pauta a l'hora de confeccionar un diccionari escolar. Les paraules d'utilitat general faciliten l'adquisició d'altres paraules; així, de les activitats didàctiques sobre el vocabulari i de l'ús del diccionari, el consultant n'extreu un benefici que està en relació amb les veus que coneix i amb la riquesa inicial.

Els punts que s'han debatut en la lexicografia escolar coincideixen amb els que interessen a la lexicografia sense adjectius:

- El nombre de veus repertoriades.
- Els criteris de selecció de l'inventari.
- La sincronia relativa, ja que el diccionari escolar ha de respondre als textos que puguin arribar als escolars o que s'hagi establert que els arribin.
- El tractament de la polisèmia, sobretot en els diccionaris escolars perquè es creu que facilita la recerca, quan l'entrada és l'aspecte més fàcil per a un escolar.
- L'explicació de les paraules monosèmiques però polivalents. Aquest tema va ser enfocat per Moon (1987) com una de les qüestions de l'ensenyament del vocabulari.
- L'exposició de l'estructura conceptual semàntica de manera que pugui permetre la inferència de la lectura o la contextualització adequada sense que això signifiqui sobrecarregar el diccionari ni fer-lo enciclopèdic.
- La inclusió de la informació gramatical i lingüística (règim sintàctic, anomalies morfològiques, registre, etc. )
- Els tipus d'ajudes per a la descodificació de les veus: il·lustracions.
- Els tipus d'ajudes per a la codificació i l'ús amb la intenció d'apropar diccionaris i thesaurus.
- etc.

## 4. QUIN HA DE SER L'ÚS DEL DICCIONARI ESCOLAR I COM ACONSEGUIR-LO

Sovint en els diccionaris escolars apareixen orientacions de consulta que no s'han de confondre amb l'ús que del diccionari es fa. Les guies són les propostes dels equips lexicogràfics per orientar la consulta i fomentar la familiaritat amb els diccionaris escolars.[18] Quant a l'observació tractada en l'apartat 2 (quin és l'ús del diccionari escolar), els estudiants que es van observar demostren conèixer els mecanismes de consulta.

(18) És el cas de Fontanillo 1983 i d'algunes primeres pàgines de diccionaris, Anaya, Santillana i Vox.

L'ús satisfactori d'un diccionari consisteix bàsicament en l'aprofitament de la informació que el diccionari conté, tant pel que fa a la comprensió com pel que fa a la producció de textos.[19] Igualment, per establir com ha de ser un diccionari s'ha d'insistir en què és una unitat lèxica, així com també per capacitar en la manera d'utilitzar satisfactòriament un diccionari s'ha de proveir a l'usuari d'uns coneixements sobre la unitat lèxica; aquest aspecte, però, ha perdut importància en els programes escolars.[20]

L'ordre de les habilitats que s'haurien de potenciar pot agafar la forma dels següents objectius:

**a. Dominar l'ordre alfabètic, com a habilitat de consulta**

En la nostra experiència, aquest objectiu sembla que s'ha assolit.

**b. Aïllar les unitats lèxiques, reconèixer les simples i les complexes i saber acotar la fraseologia**

Aquesta capacitat indica el grau de descontextualització del lèxic que l'escolar té en relació amb la globalitat del discurs. Cal recordar que els mètodes globals d'ensenyament de la lectoescriptura posposen la capacitat de l'anàlisi. A EEUU, Chall (1987) va trobar que «l'aïllament» del lèxic pot no aparèixer fins als deu anys i Curtis (1987) va dubtar de la realitat de les unitats lèxiques en el fenomen de la comprensió lectora, cosa que des d'una posició europea i analítica és difícil de comprendre.

Aquesta habilitat suposa acotar els problemes lèxics, sobretot pel que fa al discurs oral i als enunciats que presenten frases fetes. Garrido Moraga (1989) incideix en la resistència que ofereixen les frases fetes a l'hora d'aïllar-les i comprendre-les.

**c. Reconèixer la categoria lèxica i els paradigmes morfològics per localitzar l'article i l'accepció precisa**

Pel que fa a les nostres llengües, per exemple, ser capaç de reconèixer les formes verbals, els afixos apreciatius, les locucions invariables, etc. El reconeixement adequat de l'entrada també és una habilitat de consulta.

Aquesta capacitat només es pot observar en el cas de 4, ***supuesto***, encara que no el resolen bé. En les relacions del llatí que es practicaven en l'antic ensenyament secundari, el fet de reconèixer la categoria de la paraula que resultava «opaca» era un dels principals esculls abans de trobar-la al diccionari.

**d. Comprendre els articles lexicogràfics de veus dominades**

El discurs lexicogràfic, encara que sigui molt escolar i didàctic, requereix un apropament i un ensenyament. Les definicions[21] i les informacions gramaticals[22] dels articles de les veus que els adolescents coneixen, ensenyen les possibilitats virtuals de la informació que contenen els diccionaris, però també en mostren les deficiències, i és a través

(19) Corbin (1989:32).

(20) Battaner (en premsa).

(21) La definició, el text de la definició, és el punt més conflictiu de la lexicografia. Algunes propostes didàctiques fan treballar sobre la definició (Pleyán 1988, Coromina 1988): Garrido (1991) també ho proposa en la programació de la presentació del diccionari. En aquest article es manté exclusivament un apropament de comprensió i d'extracció d'informació de les definicions lexicogràfiques habituals. És explicable que, en l'intent de definir, els batxillers comprenguin les insatisfaccions dels diccionaris.

(22) Cowie (1983:183).

d'aquestes mancances que es poden preveure o eludir. Un escolar no ha de començar consultant veus que desconeix al diccionari; sinó que ha de buscar les paraules que coneix bé i les paraules que coneix «a mitges».

Sorprenen preguntes com «Fa falta definir, per exemple, per a escolars el mot **elefante**?».[23] El treball de Wiezbircka (1985) precisament tracta de la dificultat d'establir l'estructura semàntico-conceptual de les veus amb un referent concret, com animals, plantes, roba... L'equiparació de ***rebeco*** a ***gamuza*** als diccionaris utilitzats durant l'observació indueix a un ús desviat (qüestió que sempre es justifica atenent a la insuficiència dels diccionaris). Però la riquesa de contextualització que els escolars observats mostren quan fan servir ***elefante***, en contraposició amb la que mostren quan utilitzen ***rebeco***, dóna la raó a Wierzbicka (1985: 222-227) sobre la càrrega de conceptualització diferent que les veus mostren dintre de les cultures i de quina manera queda reflectida als diccionaris. ***Trompa, colmillos, marfil*** figuren sota ***elefante*** (en perífrasis lexicogràfiques); en canvi, no apareixen trets d'aquest tipus sota ***rebeco***.

Aquesta capacitat s'ha de posar en relació amb la inferència necessària en la contextualització de les veus: saber adaptar a un nivell particular i concret els trets d'informació semàntica general que els diccionaris proveeixen (Esshout-Mohr i Van Daalen-Kapteijnis 1987). Aquesta adaptació és la raó de l'exigència d'exemples en els diccionaris escolars (H. Hernández 1989a, 1989b).

Però quina classe d'exemples? ¿els estereotipats en els «prototipus» o els que marquen inferències que s'han aconseguit a partir de l'estructura conceptual de la veu? S'ha d'estudiar millor les possibilitats dels exemples. Aquest punt de la lexicografia demana atenció i experimentació (Drysdale 1987).

Sternberg (1987) suggereix que els trets que ajuden a una descontextualització més precisa i a una possible contextualització s'han de tenir en consideració en els diccionaris. Aquestes trets, però, són de moltes diverses indicacions: trets amb indicacions temporals, espacials, valoratives, descriptives, funcionals, causals, classificatòries i d'equivalència.

Per exemple, la definició del substantiu ***coraza*** permet entendre el verb ***chocar***; ***palanquín*** pot relacionar-se amb la ***sumisión de los porteadores*** sense que en un text s'hagi d'explicar.

Però, ¿com es poden extreure les inferències d'aquest tipus en els diccionaris? Primerament s'ha d'assajar amb paraules molt conegudes i observar quins trets de la definició les permeten. Trobem així que el text de la definició promou un coneixement actiu d'altres veus i de les seves relacions.

(23) F. Lázaro Carreter, «Prólogo a la presente edición», *Diccionario Anaya de la Lengua*, Madrid, Anaya, 1991, pàg. 6, l'alternativa no queda dirimida, però l'opinió que es manté és que aquestes veus, conegudes per als escolars, ocupen un espai ociós en un diccionari. Apareix, així, una concepció de diccionari restringida a la informació «referencial» exclusivament.

### e. Reconèixer algunes de les principals relacions lèxiques a partir també de paraules que els escolars dominen

Els casos observats permeten afirmar que els escolars coneixen les relacions sinonímiques, antonímiques i possiblement, –encara que no se'n tingui constància en aques-

ta experiència–, les oposicions inverses que en els llibres escolars apareixen amb els antònims. També tenen adquirides les relacions de derivació (etimològiques).

Unes altres relacions útils per entendre les possibilitats lèxiques d'una veu són les metafòriques, en les quals encara no s'insisteix suficientment.[24] Encara que els alumnes entenguin fàcilment alguns sentits figurats no aprofiten prou la informació lexicogràfica per entendre metàfores, per mesurar la precisió de les metàfores, per saber quin aspecte semàntic dels que figuren a les definicions d'una veu és el que permet la metaforització d'una veu en un context, per exemple.

A més, en els programes escolars falten encara uns altres conceptes sobre la paraula que són útils per captar els valors de les veus: pejoratius, melioratius, hiperònims, *pluralia* i *singularia tantum*, jerarquitzacions, quantificadors...

Així mateix són útils els conceptes que matisen els significats de processos, estats i accions: perfectiu, resultatiu, imperfectiu, causatiu, incoatiu, continuatiu, iteratiu i locatiu.

Per indicar aquests conceptes es podria utilitzar una terminologia més propera als escolars a través, per exemple, de significats de paraules confrontades que ja coneguessin i que permetessin més endavant aplicar els conceptes adquirits a les paraules noves i desconegudes.

La reflexió sobre les unitats lèxiques ha de precedir a altres tipus de reflexió lingüística a l'escola. Aquesta qüestió no es discuteix com caldria quan s'aborden els programes escolars, com ja s'ha remarcat abans.

És veritat que aquesta directriu pot canviar. Leed i Nakhimovsky (1990) i Steele (1990) arriben, per exemple, a aplicar les funcions lèxiques que va establir Mel'cuk per a l'ensenyament de llengües estrangeres i per a l'ensenyament de la llengua materna respectivament. De la mateixa manera, P. Nobili (1991) les aplica a l'ensenyament de l'italià com a llengua materna. Per bé que les funcions determinades per Mel'cuk i els seus col·laboradors representen una reflexió especialment acurada i complexa sobre la paraula que s'ha de simplificar d'alguna o altra manera.

Només amb l'adquisició dels conceptes que representen algunes de les principals funcions i relacions lèxiques es pot utilitzar profitosament un diccionari.

### f. Comprendre articles lexicogràfics de veus desconegudes o de veus conegudes insuficientment

El penúltim objectiu és passar finalment a la funció més genuïna del diccionari: omplir els buits d'informació lèxica dels parlants d'una llengua. Aconseguir-ho pot requerir encara monitorització, perquè en aquesta fase ressorgeixen totes les deficiències dels diccionaris que sovint passen desapercebudes quan es consulten veus conegudes.

A més de sospesar la informació que els articles del diccionari proporcionen en aquesta fase s'ha de practicar obertament l'adaptació de les veus consultades en un discurs, i s'han d'analitzar les inferències que permeten i les relacions que estableixen

(24) M. Morera Escarré (1993).

amb les altres informacions del discurs en el qual s'insereixen: no només relacions d'hiponímia, sinó també de coherència semàntica, coneixement del món, nivell de registre, relacions sintagmàtiques i adequació pragmàtica. Aquesta habilitat sempre s'ha de practicar a partir d'un text (comprensió o producció); mai a partir de veus aïllades.

g. **Retenir, confrontar i comparar la composició i disposició d'articles lexicogràfics de paraules que tinguin una estructura lèxica semblant**.

Una vegada hagi aconseguit l'estadi anterior, l'escolar es troba en situació d'utilitzar el diccionari de manera espontània i autònoma. L'estructuració sòlida d'un vocabulari de base i dels conceptes lingüístics per analitzar les funcions i les relacions li ha de permetre extreure informació vàlida dels dicionaris i per tant de fer-ne un ús personal.

Aquest últim objectiu té com a finalitat ajudar a estructurar el lèxic en la ment de l'estudiant de manera que conegui reflexivament el comportament de determinats verbs, d'alguns tipus de substantius i les seves accepcions, dels valors permanents o puntuals dels afixos, de les limitacions de la predicació en els adjectius, dels valors sintàctics d'alguns adverbis, per citar algunes possibilitats entre moltes altres.

La consulta en aquesta etapa ha de ser crítica, és a dir, ha d'admetre comparació i valoració en relació a unes necessitats conegudes.

Quan s'ha arribat a aquest objectiu, l'estudiant pot apreciar la informació que li subministra el diccionari per tal que, posteriorment, l'impulsi a la consulta habitual.

En síntesi, quan l'escolar aconsegueix tots aquests objectius és que ha hagut de passar del reconeixement intuïtiu i en bloc del que significa un text, un anunciat o un sintagma, al reconeixement raonat i analitzat que atorga qualitat i profunditat[25] al lèxic d'un parlant i el capacita per continuar aprenent vocabulari i discernint entre paraules. L'objectiu final serà, doncs, precisar el coneixement reflexiu de la llengua a partir de la paraula, cosa que exigirà a l'estudiant la utilització del diccionari durant la vida adulta.

En general, no s'acostumen a programar controls ni exàmens amb diccionari en l'ensenyament de la llengua materna perquè se sobreentén que ja es fa servir i per tant es desconfia de la quantitat d'informació que el diccionari subministra a l'estudiant. En efecte, encara que l'ús dels diccionaris aparegui en els programes escolars, no s'inclou en el currículum real. Aquest ús s'ha d'introduir decididament en el currículum escolar, tot i que no arribarà a ser real si no se l'acompanya d'un programa de reflexió acurada i precisa sobre el lèxic, cosa que justificaria la validesa i les deficiències del diccionari.

Un tema que encara no s'ha plantejat aquí, i que ara és potser el moment idoni per fer-ho, és la conveniència dels anomenats diccionaris escolars.

Només des d'una perspectiva d'objectius graduals es pot enfocar tant la conveniència d'un dicionari adaptat als cicles segon i tercer de l'Ensenyament Primari, com la graduació dels diccionaris en general.

(25) En altres treballs (Battaner 1990, Battaner (en premsa)) s'ha citat la proposta de Dale 1973 sobre quatre graus de coneixement d'una unitat lèxica: no s'ha sentit, no s'ha sentit però es coneix, es coneix aproximadament, es domina la veu fins al punt de poder-hi «jugar».

Així, sembla adequat pensar en un primer diccionari pròpiament escolar que introduexi l'adolescent en el món del lèxic i de les convencions lexicogràfiques per ser utilitzat entre els vuit i els dotze anys; un segon diccionari pensat per al ciutadà mitjà, suficientment didàctic, al qual es podria accedir en el nivell d'ensenyament secundari, encara que també fos útil al llarg de la vida; el tercer nivell escaparia de la institució escolar; seria, doncs, un dicionari per al ciutadà adult amb una professió que requerís unes informacions més precises i més específiques.

De 18.000 a 20.000 entrades el primer diccionari;[26] al voltant d'unes 40.000 entrades el segon; el nombre d'entrades del tercer és un assumpte de plantejaments que no corresponen a aquest article. En els diccionaris d'ús escolar, no tant en el primer dels proposats, pel que s'ha exposat, però sí en el segon, el repertori d'entrades ha de sobrepassar decididament els coneixements de vocabulari dels usuaris potencials.

El diccionari sembla oferir un malbaratament d'informació, encara que més tard la informació concreta sol·licitada decebi. És un fenomen semblant a la fertilitat de la naturalesa que s'ofereix ubèrrima malgrat que les condicions de perduració de vida siguin difícils per a molts éssers vius.

Intentar salvar aquesta paradoxa podria ser l'objectiu didàctic dels diccionaris i de tots els que invitin a usar-los.

## APÈNDIX

**Observació sobre usos del diccionari escolar**

**Objectiu**: recollir els comentaris que sorgeixen quan es fan servir dos diccionaris escolars i descriure les habilitats (estratègies, mancances, capacitat d'utilitzar la informació) practicades en activitats escolars per tenir protocols d'ús.

**Subjectes observats**: quatre alumnes de setè d'EGB que tenen el castellà com a llengua familiar, agrupats en dues parelles.

**Lloc i centre**: Barcelona, *Col·legi de pràctiques nº 2.*

**Duració de la prova**: una hora i mitja.

**Dates**: segon trimestre del curs 1992-1993, dimecres 10 i 17 de març de 1993.

## PROVES

1. Buscad en el diccionario ELEFANTE y REBECO. A partir de la información encontrada en el diccionario y arriesgándoos algo ¿cuántas frases podéis hacer que tengan sentido?:
   «Los elefantes...
   «Los rebecos...
   «...elefantes/elefante...»
   «...rebeco/rebecos...»

(26) El *Diccionario Básico del Español de México*, El Colegio de México, México D. F., 1986, ha optat només per 7.000 entrades i va destinat a l'ensenyament de primera i segona etapa.

¿Qué habéis aprendido sobre la palabra REBECO?
o ¿qué habéis recordado/aprendido sobre la palabra ELEFANTE? Escribid frases en que se vea.
En el artículo ELEFANTE de la fotocopia que se os ha suministrado (fotocopia del *Diccionario de uso del español* de María Moliner, Madrid, Gredos, p. 1064), ¿qué encontráis: más o menos información? ¿información más interesante o menos interesante?
Marcad lo que no entendáis.

2. ¿Sabéis la diferencia de uso y de significado entre ABERTURA, APERTURA, OBERTURA ?
Buscad estas palabras en los diccionarios y a partir de lo que leáis en ellos haced una frase bien completa con cada una de estas palabras en las que se pueda ver la diferencia que hay entre: ABERTURA, APERTURA y OBERTURA.

3. Intentad explicar la diferencia que hay entre los verbos EJERCER/EJERCITAR.
Sólo después, buscadlas en los diccionarios; valorad si os han servido de ayuda.

4. *En el SUPUESTO de que nos lo deje, intentaríamos copiarlo, Por SUPUESTO que iremos a buscarlo cuando llegue!*
¿cómo quedan diferenciados en los diccionarios estos dos usos de SUPUESTO? Escribid todas las diferencias que el diccionario explica.

5. Los diccionarios pueden servir para relacionar unas palabras con otras, buscad todas las palabras que puedan tener alguna relación con QUIETUD y PRENSA, y explicad cuál es esa relación. (familia, sinónimos, antónimos, palabras con el mismo sufijo).
Quizás os sea útil leer el artículo de QUIETO.

6. El adjetivo LISTO-A tiene varias acepciones, mirad cuántas le dan los diccionarios, haced una frase con cada una y luego probad a sustituirlo en estas frases por los sinónimos y antónimos que ofrecen los mismos diccionarios. Valoradlo.

7. ¿Sabéis buscar cómo es la primera persona del singular del Pretérito Simple de Indicativo de CONDUCIR en el diccionario? Veamos.

8. Las palabras gramaticales son difíciles de explicar en un diccionario porque generalmente se conocen muy globalmente y, sin embargo, los diccionarios dan informaciones precisas. Valorad cuál es la información que mejor entendéis en los diccionarios consultados sobre NUESTRO-A.

M. Paz Battaner
*Universitat Pompeu Fabra*

## REFERÈNCIES BIBLIOGRÀFIQUES

AHUMADA LARA, I. (1989) «Acerca de la información gramatical en los diccionarios escolares». A: F.Garrido Carabias y J.Comesaña Rincón, *Actas del VII congreso Nacional de Lingüística Aplicada*. Sevilla, AESLA- Universidad de Sevilla, 1989, pàgs. 53-59.

AITCHISNSON, J. (1987) *Words in the Mind, an Introduction to the Mental Lexicon*. Londres, Basil Blackwell.

ALVAR EZQUERRA, M. (1982) «Función del diccionario en la enseñanza de la lengua». A: *Revista de Bachillerato*, monogràfic, (9), pàgs. 49-53.

BATTANER, M. P. (1990) «Más allá de la punta de la lengua». A: M. P. Battaner y G. Pujals (coords.) *El vocabulari i l'escrit*, VII Jornades de Redacció i Escola, Barcelona, Universitat de Barcelona, pàgs. 115-133.

— (en premsa) «La investigación en la enseñanza del español como lengua materna. Vocabulario y léxico». A: *Actas del Congreso de la Lengua Española*. Sevilla, Octubre, 1992, Pabellón de España-Instituto Cervantes.

BATTENBURG, J. D. (1991) *English monolingual learners' dictionaries. A user-oriented study*. Tübingen, Max Niemeyer.

BECK, I. L.; MCKEOWN, M. G.; OMANSON, R. C. (1987) «The Effects and Uses of Diverse Vocabulary Instructional Techniques». A: Mckeown, M. G.; Curtis, M. E. *The Nature of Vocabulary Acquisition*. Hillsdale, N.J., Lawrence Erlbaum Associates, 1987, pàgs. 147-163.

CHALL, J.S. (1987) «Two vocabulary for Reading: Recognition and Meaning». A: Mckeown, M. G.; Curtis, M. E. *The Nature of Vocabulary Acquisition*. Hillsdale, N.J., Lawrence Erlbaum Associates, 1987, pàgs. 19-35.

CORBIN, P. (1989) «Lire les dictionnaires». A: Ibrahim, A. H.; Zalessky, M. *Lexiques*, número especial de *Le Français dans le Monde*, 1989, pàgs. 31-41.

COROMINA, E. (1988) *Explicar mots i fets*. Barcelona, Teide, 1a ed.

COSTA, J.; VILADROSA, M.; VILARÓ, F. (1992) «L'anàlisi dels diccionaris de definicions en un curs per a assessors lingüístics» (1). A: *Com ensenyar català a adults*, 1992, (27) pàgs. 15-20.

COWIE, A. P. (1987) «Syntax, the dictionary and the learner's communicative needs». A: A.Cowie, *The dictionary and the Language Learner*. Tübingen, Max Niemeyer Verlag, 1987, pàgs. 183-192.

DALE, E. (1973) *The Words We Know: A National Inventory*. San Palo, Ca., Field Educational Publications, Inc.

DRYSDALE, P. D. (1987) «The role of examples in a learner's communicative needs». A: Cowie, A. *The Dictionary and the Language Learner*. Tübingen, Max Niemeyer Verlag, 1987, pàgs. 213-223.

ELSHOUT-MOHR, M. i DAALEN-KAPTEIJNS, M. M. van (1987) «Cognitive processes in

Learning Word Meanings». A: Mckeown, M. G. ; Curtis, M. E. *The Nature of Vocabulary Acquisition*. Hillsdale, N. J., Lawrence Erlbaum Associates, 1987, pàgs. 53-71.

FONTANILLO MERINO, E. (1983) *Cómo utilizar los diccionarios*. Salamanca, Anaya.

GARRIDO MORAGA, A. M. (1991) «Lexicografía y enseñanza». A: H. López Morales, *La enseñanza del español como lengua materna*, Actas del II Seminario Internacional sobre Aportes de la lingüística a la enseñanza del español como lengua materna. San Juan, Ed. de la Universidad de Puerto Rico, pàgs. 29-47.

HARTMANN, R. R. K. (1987a) «Four perspectives on dictionary use: a critical review of research methods». A: Cowie, A. *The dictionary and the Language Learner*. Tübingen, Max Niemeyer Verlag, 1987, pàgs. 11-28.

— (1987b) «Dictionaries of English: The User's Perspective». A: Bailey, R. W. *Dictionaries of English, Cambridge*, C.P.U., 1987, pàgs. 121-135.

— (1989) «Sociology of the dictionary user: Hypotheses and empirical studies». A: Hausmann, F. J. i alii, *Worterbucher/Dictionaries/Dictionnaires. An International encyclopedia of lexicography*. 3 vols, Berlin, W.de Gruyter.

— (1992) «Lexicography, with particular reference to English learners' dictionaries». A: *Language Teaching*, (25/3), pàgs. 151-159.

HERNÁNDEZ, H. (1989a) *Los diccionarios de orientación escolar*. Tübingen, Max Niemeyer Verlag.

— (1989) «Semántica y lexicografía: el carácter productivo de los diccionarios». A: *Tabanque*. 5, pàgs. 45-52.

IBRAHIM, A.H.; ZALESSKY, M. (coords) (1989), *Lexiques*, número especial de *Le Français dans le Monde*.

— (1989) «Enquête: L'Usage des dictionnaires». A: *Lexiques*, número especial de *Le Français dans le Monde*, pàgs. 24-30.

LARA, L. F.; HAM, R. H.; HIDALGO, M. I. (1979) «Base estadística del *Diccionario del español de México*». A. *Investigaciones lingüísticas en lexicografía*. México, El Colegio de México, pàgs. 5-39.

LEED, R. L.; NAKHIMOVSKY, A. (1990) «Lexical functions and Language Learning». A. Steele, J. *Meaning-Text Theory. Linguistics, lexicography and implications*. Otawa, Univ. of Otawa Press, pàgs. 365-375.

LÈPINETTE, B. (1989) «Vers un dictionnaire explicatif et combinatoire bilingue». A: *Cahiers de lexicologie*. 54.

MARTÍNEZ MARÍN, J. (1991) «Los diccionarios escolares del español: lo que son y lo que deberían ser». A: Ahumada, I (coord.), *Lecciones del I Seminario de Lexicografía Hispánica*. Jaén, Facultad de Humanidades, pàgs. 53-70.

M.E.C. (1989) *Vocabulario básico en la E.G.B.*. Dos vols., dirigit per J. J. de Bustos, coordinado per Casanova, M. A.; Rivera, M. Madrid, M.E.C.-Espasa Calpe.

MEL'CUK, I. A. (1982) «Lexical functions in lexicographic description». A: *Proceedings*

*of the Eigth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, Department of Slavic Languages and Literatures, University of California, pàgs. 427-444.

MENÉNDEZ PIDAL, R. (1945) «El diccionario que deseamos», pròleg del *Diccionario Vox* dirigit per S. Gili i Gaya, reeditat com «El diccionario ideal». A: *Estudios de Lingüística*. Madrid, Espasa-Calpe, 1965.

MOON, R. (1987) «Monosemous words and the dictionary». A: Cowie, A. *The dictionary and the Language Learner*. Tübingen, Max Niemeyer Verlag, 1987, pàgs. 173-182.

MORERA ESCARRÉ, M. (1993) *Metáfora y competencia léxica (Aplicaciones didácticas en la enseñanza secundaria)*. Tesi doctoral, Universitat Autònoma de Barcelona.

NOBILI, P. (1991) «Definizioni nei dizionari e strutturazione didattica del lessico». A: *Lend*. 2, pàgs. 50-59.

OLARTE STAMPA, L.; GARRIDO MORAGA, A. M. (1984) «Diccionario y enseñanza. Aproximación a los diccionarios más usados en los niveles educativos». A: *Español Actual*. 41, pàgs. 21-29.

PLEYÁN, C. (1988) *Lengua española, 1º de B.U.P.*. Barcelona, Teide

REY-DEBOVE, J. (1989) «Dictionnaires d'apprentissage: que dire aux enfants». A: Ibrahim, A. H. i Zalessky, M. *Lexiques*, número especial de *Le Français dans le Monde*. pàgs. 18-24.

STEELE, J. (1990) «Meaning-Text Theory and First-Language Learning: Some Implications». A: Steel, J. *Meaning-Text Theory. Linguistics, lexicography and implications*. Otawa, Univ. of Otawa Press, pàgs. 376-390.

STEELE, J.; MEYER, I (1990) «Lexical functions in an Explanatory Combinatorial Dictionary: Kinds, Descriptions and English Examples». A: Steel, J. *Meaning-Text Theory. Lingüistics, lexicography and implications*. Otawa, Univ. of Otawa Press, pàgs. 41-61.

STERNBERG, R. J. (1987) «Most vocabulary is Learned from Context». A: Mckeown, M. G.; M. E. Curtis, M. E. *The Nature of Vocabulary Acquisition*. Hillsdale, N. J., Lawrence Erlbaum Associates, 1987, pàgs. 89-105.

THORNDIKE, E. L.; LORGE, I (1944) *The Teacher's Word Book of 30.000 words*. Nueva York, Teachers College - Columbia University.

WIERZBICKA, A. (1985), *Lexicography and conceptual analysis*. Ann Arbor, Karoma Publishers, Inc.

EMILI PASCUAL

# LA PROFESSIÓ DE LEXICÒGRAF

## INTRODUCCIÓ

És innegable que a hores d'ara la lexicografia ocupa un lloc important en el món de les ciències del llenguatge i que aquest fet ha comportat que es pugui parlar d'una tasca científica feta per uns professionals amb la preparació adequada: els lexicògrafs.

En aquest article intentaré de fer unes breus reflexions sobre aquest tema, amb una especial atenció a les característiques de la feina o la professió de lexicògraf.

Després d'una presentació del canvi que s'ha produït, d'uns trenta anys ençà, en la consideració de la lexicografia i en el món dels diccionaris, esbossaré la comesa del lexicògraf com a científic i tècnic en la confecció dels diccionaris. La lexicografia, certament, ha estat considerada una feina artesanal, deslligada de l'evolució de les ciències del llenguatge, però ara té una gran importància entre aquestes ciències, fins a l'extrem que algun estudiós ha pogut afirmar que el món dels diccionaris està sotmès, des de fa trenta anys, a una revolució radical que encara no ha acabat.

El lexicògraf continua sent la persona que «fa diccionaris», però ara és un professional que ha de conèixer, a un nivell suficient, les aportacions científiques de la lingüística general i de les diverses branques especialitzades, especialment els treballs sobre el lèxic, els quals aquests darrers anys s'han multiplicat abundosament (estudis sobre el concepte de mot, la lexicalització, aportacions del camp de la semiòtica, etc); ha de tenir en compte també els avenços que han tingut lloc en el camp de discipli-

*Caplletra 17* (Tardor 1994), pp. 125-136

nes com la dialectologia i la història de la llengua (en l'àmbit català, per exemple, un lexicògraf no pot ignorar la valuosa i voluminosa obra de Joan Coromines); ha de conèixer i valorar les teoritzacions dels experts tocant a la definició (tema cabdal de la lexicografia i que ben recentment ha estat estudiat i estructurat), al tractament de la macroestructura i de la microestructura dels diccionaris, als treballs sobre polisèmia, sinonímia i antonímia; ultra això, el lexicògraf avui ha de conèixer bé les possibilitats diverses de presentació del lèxic segons el tipus de diccionari que ha de fer, general o especialitzat, gramatical, ortogràfic, amb exemples o no, amb gràfics i/o il·lustracions, històric o actual, monolingüe o plurilingüe, normatiu o descriptiu, escolar o per al gran públic, etc, cosa que pressuposa, doncs, que ha de tenir molt clar quin nivell d'usuaris vol satisfer i quins registres de llengua ha de fer servir; és a dir, cal que conegui suficientment les aportacions recents de la sociolingüística.

Finalment, la lexicografia rep a hores d'ara un suport inestimable de la informàtica, de manera que es pot afirmar sense traïr la veritat que una gran part de la revolució en el món dels diccionaris ha vingut provocada pel progrés dels ordinadors aplicats al tractament de textos i de materials lexicogràfics (tenir en una base de dades lexicogràfiques milers d'informacions, perfectament classificades i fàcilment consultables, no és una utopia, sinó una realitat).

Dues darreres consideracions encara abans de finir aquestes ratlles introductòries:

a) La lexicografia és ara un treball d'equip, en el qual participen professionals de diverses especialitzacions lingüístiques (lingüistes, dialectòlegs, semiòtics, terminòlegs, etc), els quals esdevenen col·laboradors imprescindibles del lexicògraf.

b) Malgrat el que s'ha dit abans, aquesta revolució lexicogràfica no representa una solució de continuïtat dràstica entre els lexicògrafs actuals i els antics; en l'àmbit de la lexicografia, hi ha un pes important de la tradició, i una bona dosi de «factor humà», que uneix el lexicògraf actual amb els seus predecessors en la història dels diccionaris.

## I. DE LA LEXICOGRAFIA D'AFECCIONATS A LA PROFESSIONAL

Atès que la lexicografia és una disciplina recent, en la seva consideració científica, parlar de lexicògraf és aplicar una qualificació professional moderna que distingeix els autors dels diccionaris actuals dels seus predecessors. No fa pas massa anys que els autors de diccionaris eren persones interessades en la llengua que sentien la pruïja de fer obres lexicogràfiques de consulta, sovint com una activitat que els omplia les estones de lleure. Així, igual podia ser autor d'un diccionari un químic que un literat o un comerciant. Sens dubte que la major part d'aquests autors de diccionaris eren persones amb un cert nivell de cultura i d'erudició, coneixedors de la llengua si més no en el grau suficient per no cometre equivocacions greus; disposaven, a més, d'altres diccionaris o reculls lexicogràfics anteriors en què basar-se. És ben cert que la majoria

de diccionaris han partit d'obres fonamentals que, en un moment determinat, marcaren una fita important en la lexicografia, i que els autors de diccionaris posteriors, en gran part, no han fet altra cosa que «copiar-los», sense aportar altres materials lexicogràfics fiables o renovadors. Pel que fa al català, per exemple, qui pot negar que Fabra i l'Alcover-Moll han estat la font d'on ha begut tota la lexicografia que s'ha fet fins ara? No debades, algun estudiós modern de temes lexicogràfics ha parlat de la munió de plagiaris en la història dels diccionaris.

Per tot això, la confecció de diccionaris i vocabularis ha estat considerada fins fa ben poc una obra artesanal, una art o, a tot estirar, una tècnica, feina de diletants i afeccionats, però no pas de científics. Cosa que ha tingut com a conseqüència una certa menysvaloració de la feina lexicogràfica. La lexicografia era una «artesania», no una tasca científica.

Ja és prou sorprenent que els autors de diccionaris, fins al moment que s'han fixat els mètodes de treball i els continguts científics de la lexicografia moderna, mai no han teoritzat sobre la seva tasca ni han exposat els principis i la metodologia dels seus treballs. Mentre que les diverses ciències evolucionaven i explicitaven les respectives metodologies i les tècniques que empraven, la confecció de diccionaris es mantenia com una feina obscura i artesanal, fruit d'uns autors que produïen obres abundoses, sense, però, explicar en cap moment, ni en els prefacis de les obres, els mètodes de treball, les fonts de procedència dels materials, les tècniques de tractament d'aquests i els principis teòrics en què es basaven. Més d'un d'aquests artesans s'hauria sorprès si se li hagués preguntat de quina manera i amb quin utillatge teòric i pràctic feia el seu diccionari.

Afortunadament, el progrés de la lingüística en conjunt, en totes les seves branques, ha comportat l'establiment d'una lexicologia científica i d'una ciència de la confecció dels diccionaris (la lexicografia). Entre els estudiosos es plantegen modernament qüestions teòriques de terminologia en relació amb les diverses disciplines que tracten del lèxic, tot intentant de delimitar amb claredat els camps pertinents de la lexicologia, la morfologia i la semàntica, de la lexicologia i la pragmàtica, de la lexicologia i la lexicografia i, finalment, de la lexicografia i la terminologia. No cal entrar ara en disquisicions teòriques sobre aquests temes ni sobre la consideració de ciència separada o no, subordinada o no, necessària o no, de la lexicologia i la lexicografia. Per al nostre propòsit és suficient que es pugui afirmar que hi ha una especialització de la lingüística aplicada que és la lexicografia, estretament vinculada al seu torn amb una teorització sobre el lèxic que és la lexicologia.

El que és important és que l'estudi, la crítica i la confecció de diccionaris ja no es pot considerar una activitat artesanal, en mans de persones afeccionades o diletants, sinó que és una tasca que es porta a terme amb principis de la ciència lingüística i amb una metodologia i unes eines adequades, que cada vegada s'afinen més i resten a les mans d'uns professionals experts que són els lexicògrafs.

La lexicografia ha passat de ser «l'art de fer diccionaris» a ser «la ciència i la tècnica dels diccionaris». La lexicografia repensa i estructura els mètodes de treball, analitza els components que configuren els diccionaris; estableix les condicions científiques i tècniques de l'obra lexicogràfica i estudia fins i tot el producte final que ha d'oferir al gran públic, consultor potencial dels diccionaris, com també aporta dades sobre la funció social d'aquest tipus d'obra des de perspectives sociolingüístiques.

En aquest canvi de perspectiva han influït factors molt diversos: després de la Segona Guerra Mundial, les ciències lingüístiques han estat sotmeses a una evolució ràpida i profunda. Augmenten, d'altra banda, la intercomunicació lingüística i l'intercanvi de documentació, amb la necessitat de traduccions i d'aprenentatge de llengües estrangeres. Els experts converteixen el diccionari en objecte d'estudi, i l'usuari, més preparat i més crític, exigeix del diccionari respostes més científiques i més fiables. No solament es «fan» diccionaris, sinó que «s'estudia» el diccionari (les seves característiques, les condicions que ha de complir, els principis que han de sustentar-ne l'elaboració i l'ús). Més encara, els diccionaris tenen ja com a autors especialistes del camp de la lingüística. S'estableixen també les tipologies dels diversos diccionaris, mentre que, alhora, augmenten els estudis crítics sobre els treballs lexicogràfics precedents i se sistematitzen les dades pertinents per a una història dels diccionaris i del diccionari.

## II. ESTABLIMENT DE LA LEXICOGRAFIA CIENTÍFICA

En la fixació de la lexicografia moderna hi ha un moment important i ineludible: la celebració del Congrés Internacional de lexicografia celebrat a Exeter, Gran Bretanya, l'any 1983. Aquest Congrés recull les aportacions i el fruit dels estudis duts a terme a partir sobre tot dels anys 60 i 70. Diversos estudiosos del lèxic s'hi van plantejar la tasca de la lexicografia i la necessitat d'establir les bases científiques per a un treball competent i avaluable dels lexicògrafs. Arran d'aquest Congrés es va fundar EURALEX (European Association for Lexicography). Es pot dir que en l'establiment de la lexicografia moderna i en la fixació de les tasques dels lexicògrafs hi ha un abans i un després del Congrés d'Exeter. Cosa que no vol dir, evidentment, que la labor de fer diccionaris comenci en aquest moment. Tots els estudiosos del lèxic afirmen que la tasca lexicogràfica és molt antiga. La majoria dels autors reconeixen que la lexicografia, com a activitat pràctica, és més antiga que la lingüística, però que és modernament que la lexicografia s'estructura científicament i desenvolupa una teoria pròpia i una metodologia adequada, en estreta relació, com hem dit, amb els progressos de la lingüística general i de la lingüística aplicada.

Situada la lexicografia en el rang que li correspon entre les disciplines lingüístiques, es pot definir segons els termes en què coincideixen la majoria dels estudiosos: «La lexicografia és la disciplina de la lingüística aplicada que s'encarrega dels problemes

teòrics i pràctics que planteja l'elaboració de diccionaris»; o també: «La lexicografia és la part de la lingüística aplicada que s'ocupa de la temàtica i els principis de l'elaboració de diccionaris».

Aquestes definicions pressuposen, però, que el lexicògraf no solament ha de fer diccionaris, sinó que la seva competència l'ha de portar a fer la crítica dels diccionaris ja existents i a establir els principis teòrics que possibiliten la tasca lexicogràfica. D'altra banda, és palès que un lexicògraf pot ser un estudiós de la lexicografia i/o un redactor de diccionaris (o d'un diccionari), atès que, al seu torn, la lexicografia pot ser l'estudi dels diccionaris i/o la confecció dels diccionaris.

Els teòrics del Congrés d'Exeter establiren també les relacions que la lexicografia havia de mantenir amb les altres disciplines de la lingüística i intentaren de definir tots els components de què hauria de constar la disciplina lexicogràfica, i, doncs, quins són els àmbits de coneixements científics que han de ser exigits al lexicògraf per a desenvolupar competentment la seva professió.

Així, es van proposar les connexions de la lexicografia amb la lingüística general i amb la lingüística computacional, i també el suport que el treball lexicogràfic rep de les tecnologies de la informació. D'altra banda, es va postular l'existència d'una metalexicografia de la qual formarien part la història de la lexicografia, la teoria general de la lexicografia, la investigació sobre l'ús del diccionari i la crítica dels diccionaris. Ultra això, es van estudiar les activitats bàsiques i necessàries per a la tasca del lexicògraf: projecte del diccionari; establiment de la base del diccionari; activitats directament relacionades amb l'escriptura dels textos del diccionari.

Notem de passada que en l'àmbit de la llengua catalana d'un temps ençà comencen a sorgir els treballs acadèmics i de divulgació que tenen com a tema la crítica dels diccionaris existents o simplement la presentació classificada dels reculls lexicogràfics, cosa que facilita la feina dels estudiosos que dedicaran els seus esforços a una història dels diccionaris catalans o a presentar al gran públic la procedència dels materials lexicogràfics i la dependència d'uns diccionaris respecte d'uns altres.

L'establiment de la lexicografia entre les ciències lingüístiques planteja òbviament la necessitat també d'establir àrees d'estudi especialitzat a les Facultats de filologia. Dit altrament, cal una preparació adient dels professionals de la lexicografia, els lexicògrafs, i és així que en diversos àmbits acadèmics es treballa seriosament en aquest aspecte. Igualment, hi ha un moviment paral·lel conduït pels lexicògrafs i per les editorials dedicades als diccionaris per a educar el gran públic tocant a la importància i la necessitat d'aquestes obres i, doncs, que impulsa la formació dels usuaris dels diccionaris. Perquè és evident que part del públic té una gran ignorància i un gran desconeixement del diccionari; molts usuaris de la llengua no saben manejar el diccionari ni coneixen les funcions que acompleix. El lexicògraf resta sovint perplex davant l'actitud de tantes persones que no aprofiten el recurs tan elemental de consultar el diccionari abans de penjar rètols a les botigues i altres llocs públics amb faltes d'ortografia elementals.

## III. LA FEINA DEL LEXICÒGRAF

En aquest tercer apartat de l'article intentaré de resumir, a través de consideracions breus, les característiques que al meu entendre defineixen la feina del lexicògraf.

a) La lexicografia pertany a l'àrea de la lingüística aplicada i això ens fa veure, d'antuvi, que el lexicògraf no és pròpiament un teòric de la llengua. El seu treball no pertany a la teorització lingüística, sinó a la praxis de la llengua; és un professional pràctic, més que teòric, un tècnic, quan treballa en la confecció de diccionaris. No pot dedicar el seu temps a meres disquisicions lingüístiques ni treballa un món de teories on es barregen consideracions especulatives de caire semàntic, semiòtic i lingüístic. El lexicògraf ha de fer una obra concreta que es resol, en darrera instància, en un producte útil, concret, amb unes característiques funcionals i pràctiques: un llibre, el diccionari.

Per tant, malgrat que el lexicògraf ha de conèixer el món teòric de la lexicologia, la semàntica i la semiòtica, la seva feina s'ha de desenvolupar forçosament al nivell de temàtiques molt concretes i ha de donar solucions també concretes, sense que li sigui lícit de deixar problemes sense resoldre o de donar com a excusa, en un moment determinat, que no té la competència suficient.

La tasca lexicogràfica comporta, a més, un seguit de condicionaments pragmàtics que els teòrics de les ciències lingüístiques desconeixen la majoria de vegades: exigències editorials, temps d'edició de l'obra, compromisos contractuals, limitacions d'espai, rendibilitat de l'obra, resposta del públic, etc; un diccionari és un producte que s'ha de vendre, i resta, per tant, sotmès a les lleis de la compra-venda i a les condicions de la producció i del mercat.

El lexicògraf, doncs, ha de tenir en compte els progressos de la semàntica, per exemple, en el camp de l'estudi dels mots i dels significats, però s'ha de deseeixir dels problemes teòrics que aquesta disciplina planteja si vol desenvolupar eficaçment la seva feina. Altrament restaria bloqat per les disquisicions teòriques i no podria avançar en la presentació dels materials lexicogràfics als usuaris de la llengua.

b) Com a conseqüència, el lexicògraf és conscient que la seva feina té sempre uns components d'imperfecció, unes limitacions gairebé congènites, que provenen del fet que li és molt difícil d'aplicar en la pràctica les conclusions teòriques de les disciplines sobre el lèxic en la confecció dels diccionaris. Un diccionari serà sempre una obra útil, però incompleta; necessària, però imperfecta. Els estudis recents sobre la definició, per exemple, han fet paleses les deficiències estructurals d'aquest element en molts diccionaris; però, tot i amb això, el lexicògraf ha de definir, i ho ha de fer segons unes pautes que puguin connectar amb el consultor de l'obra lexicogràfica, malgrat que teòricament algun sistema definitori sigui considerat insuficient o millorable. El lexicògraf ha d'optar per la solució que li sembla millor i no pot plantejar constantment problemes teòrics quan té davant seu la tasca, per exemple, d'enllestir centenars o milers

de definicions. Té, però, al seu favor el fet que treballa dins del sistema de la llengua, atès que definir no és una operació metalingüística fonamentalment; el lexicògraf, de fet, no fa formalitzacions lingüístiques, sinó que tracta la llengua emprant la mateixa llengua.

D'altra banda, l'aplicació estricta de determinades teoritzacions sobre el tractament dels lexemes i sobre les definicions conduiria a un tipus de diccionari apte solament per a persones molt enteses i per a científics de la llengua, però que esdevindria inservible per als usuaris normals i habituals dels diccionaris. En aquest punt també cal remarcar que un diccionari és un tipus d'obra que pressuposa en el consultor un determinat nivell de competència en l'ús de la llengua i un desenvolupament cultural suficient, cosa que permet de superar moltes deficiències i de fer útil la consulta. Certament, determinats diccionaris (com els pensats per als infants, els escolars, etc.) tenen una especificitat que els diferencia dels diccionaris generals de la llengua. Cosa que també afecta els diccionaris bilingües, els quals comporten unes exigències de fiabilitat, d'exactitud en les equivalències dels significats i d'exemplificacions d'ús de les formes que fins ara no sempre s'han tingut en compte en aquesta mena d'obres. En aquest camp, la lexicografia actual té també molt a dir i una gran feina a fer.

c) Tothom està d'acord que la feina del lexicògraf consisteix bàsicament en tres operacions:

- selecció dels materials, és a dir, recollida de les dades que constituiran el corpus del diccionari. En la majoria dels diccionaris aquesta selecció és, de fet, l'establiment de les entrades de l'obra (el que s'anomena la «macroestructura» del diccionari).
- descripció dels materials seleccionats, segons les característiques del diccionari; aquesta descripció comporta habitualment l'adscripció a categories gramaticals dels lexemes, les definicions, les etimologies, els exemples i altres explicacions complementàries (elements que configuren l'anomenada «microestructura» del diccionari).
- finalment, el lexicògraf ha de presentar l'obra, li ha de donar forma a través d'un procés tècnic que té continuïtat en totes les etapes de l'edició fins arribar a situar en mans dels usuaris un producte acabat i definit: el diccionari.

Podem afegir que el lexicògraf ha de tenir sempre ben present que treballa per a uns usuaris concrets i que la seva feina està al servei dels consultors potencials del diccionari. Aquests usuaris són molt diversos, i els diccionaris s'especialitzen cada vegada més per servir un determinat tipus de consultors.

d) La teoria del lèxic i dels significats planteja avui un seguit de problemes que el lexicògraf ha de conèixer i ponderar, ja que afecten la seva feina. D'uns anys ençà, la teoria sobre el concepte de «mot», per exemple, ha fet veure la necessitat de repensar el tractament de les formes lèxiques en els diccionaris i ha incidit positivament

en el tractament de la macroestructura. El lexicògraf sap que hi ha problemàtiques molt complexes plantejades al voltant de temes com els lexemes compostos unisignificants («banc de dades», «peu de rei», per exemple), les formes lèxiques coincidents en la grafia, però de significats allunyats («cabra» -insecte, bolet, crustaci), la fraseologia o el tractament de la sinonímia. Per què, i és un altre exemple, «blat de moro» és tractat generalment dins de l'entrada «blat», si el fet és que el blat de moro no és blat? El lexicògraf, per tant, és conscient que ja no es tracta solament de les dificultats de definir, sinó que ha de tenir en compte molts altres problemes que li són plantejats teòricament.

A més, la Lexicografia actual intenta d'aprofundir en temes que han passat pràcticament desapercebuts en els diccionaris fins ara, com les relacions de funcionament morfosintàctic de les formes: usos verbals, preposicionals i adverbials; col·locació dels mots; combinatòria d'aquests, etc. Així, per exemple, el lexicògraf ha d'intentar informar els consultors del diccionari sobre la «col·locació»: en català és possible que la forma «avui» funcioni en contextos com «Qui ha vingut avui?» i «Avui, qui ha vingut?», en frase interrogativa; però no és possible «Qui avui ha vingut?». Igualment, la combinatòria: «nosa» i «punyetes» es combinen amb «fer», no pas amb «donar»; i, doncs, no és possible, en català, «Vés a donar punyetes» o «Em dónes nosa». Aquests aspectes de la llengua, tradicionalment menystinguts o no considerats en els diccionaris, la moderna lexicografia exigeix que es tinguin en compte, especialment en un món en què els intercanvis culturals, els viatges i la multiplicitat d'informació fan possible un bescanvi lingüístic que era impensable només fa cinquanta anys. Els diccionaris de llengua, d'una llengua determinada, són avui consultats pels estrangers, els quals necessiten sovint molt més les informacions d'ús de les formes que les definicions.

Diversos projectes de diccionaris moderns apliquen ja tècniques i mètodes que permeten de comprendre el funcionament estructural de la llengua de què es tracti, insistint en els aspectes de la combinatòria i de la col·locació dels elements en la cadena morfosintàctica. Els diccionaris tradicionals suplien aquestes mancances, de manera sovint insuficient, a través dels exemples.

e) Ja he dit abans que el lexicògraf no pot deixar de prendre decisions en la seva feina, de la mateixa manera que no pot deixar de definir; ha de decidir en la selecció de materials, en el seu tractament, en les definicions, malgrat que sovint es troba amb problemes no resolts i amb qüestions difícils. Els recursos de què disposa no sempre són suficients (altres diccionaris, recerques dels estudiosos, etc). I és que, de fet, són pocs els investigadors de la llengua que hagin treballat en totes les àrees i puguin aportar dades de primera mà. S'imposaria, aleshores, una recerca de camp que la majoria de vegades no és possible. Un treball de camp és en la pràctica molt difícil, i només és possible en grups lingüístics minoritaris.

f) El lexicògraf esdevé, sense pretendre-ho, només per les conseqüències de la seva feina, un dels principals factors d'estandardització de l'idioma i de fixació de les formes «cultes» de la llengua. De fet, la majoria de les llengües tenen el diccionari normatiu, en el qual per una convenció entre les autoritats acadèmiques i els usuaris de la llengua resten enregistrades les formes que configuren els usos «correctes» i normalitzats de l'idioma. Però deixant de banda aquest tipus de diccionari, l'usuari de la llengua cerca en les obres dels lexicògrafs la seguretat en l'ús de les formes, la confirmació de la seva competència lingüística en el camp del lèxic, la recta escriptura de les formes de la parla, etc.

De fet, es genera en la tasca lexicogràfica un doble corrent entre els usuaris de la llengua i els autors de diccionaris: el lexicògraf no pot deixar de recollir els usos de la comunitat parlant, i aquesta troba confirmada la seva competència en les pàgines del diccionari. El lexicògraf no pot inventar la llengua; en tot cas, explicita, fixa, deixa constància del que han fet o del que fan els parlants de la llengua. Fins i tot en la formació dels tecnicismes, ha de comptar amb la metodologia que la Terminologia ha establert i amb les regles de derivació i de formació de compostos que pertanyen al sistema de la llengua.

D'altra banda, per a molts usuaris de la llengua el diccionari ha tingut i té encara una consideració gairebé sacra: el que es troba en el diccionari és la veritat de la llengua; el que no es troba en el diccionari no és bo o no és correcte.

g) El lexicògraf té avui en la informàtica unes variades possibilitats d'ajut en la seva feina. De tal manera que ara ja no és concebible una feina lexicogràfica feta amb fitxes de paper, tisores, goma d'enganxar i arxivadors. Tot allò que era llarg i lent en les tasques de confecció d'un diccionari, com copiar, classificar, seleccionar o triar, alfabetitzar, etc, és agilitat fiablement pels ordinadors. Avui es pot afirmar que la majoria, per no dir tots, els projectes de diccionaris es fan, en menor o major grau, amb assistència d'ordinadors.

Els estudis teòrics sobre el processament de dades lexicogràfiques i sobre treballs lèxics progressen notablement, mentre que a la pràctica la confecció dels diccionaris i dels vocabularis s'informatitza cada vegada més.

El treball amb ordinadors permet un control ràpid dels materials lexicogràfics, l'alfabetització automàtica, la supressió o la inclusió de dades, la consulta de la feina ja feta, el control de les remissions i de la freqüència d'ús de determinades formes, etc.

La informatització, d'altra banda, abraça tot el procés de la feina del lexicògraf, atès que el resultat final és un llibre, producte que també és tractat informàticament (picatge, correcció, maquetació, etc).

No cal, evidentment, que el lexicògraf sigui un informàtic. Però ha de saber el mínim que li permeti de parlar tècnicament amb els informàtics, dels quals rebrà el suport i l'ajut que calgui, i ha de saber moure's amb la soltesa suficient entre els laberints de les màquines i dels programes.

El progrés informàtic aplicat a la lexicografia ha permès ja l'existència de diccionaris en suport informàtic, que poden ser consultats pels usuaris a través de disquets. Es treballa seriosament en la creació de diccionaris generals de la llengua constantment actualitzats que seran consultats pels usuaris a través de serveis de minitel i de teletext.

Cal, però, tenir en compte que la llengua és un objecte de treball que no consisteix en un cúmul de dades exactes perfectament formalitzables i informatitzables. I que, doncs, la tasca del lexicògraf és imprescindible; ell aporta aquella dosi de «factor humà» que converteix la labor de fer diccionaris en una tasca alhora apassionant i feixuga, perquè la llengua no és una realitat fixa i immutable, sinó el mitjà de comunicació d'una comunitat viva en constant evolució.

## CONCLUSIÓ

Crec que a les pàgines anteriors han quedat resumides les característiques que configuren avui la professió de lexicògraf. L'establiment de la lexicografia com una branca de la ciència lingüística ha arrossegat la figura del lexicògraf com un tècnic i un científic de la confecció de diccionaris, ultra configurar-lo també com la persona entesa que pot criticar els diccionaris i establir la metodologia i les eines pròpies dels treballs lexicogràfics.

La lexicografia té a hores d'ara una bibliografia abundant i actual, i ací i allà sorgeixen els estudis i les obres que apliquen ja la metodologia i els principis d'aquesta branca de la lingüística aplicada. Els fruits de tot això es recolliran, ho esperem, en obres lexicogràfiques modernes i renovadores, verament útils als usuaris de la llengua, les quals, malgrat tot, entroncaran amb una llarga tradició en la història dels diccionaris. Ho esperem i ho desitgem.

EMILI PASCUAL FERRANDO
*Enciclopèdia Catalana*

## REFERÈNCIES BIBLIOGRÀFIQUES

HARTMANN, R.R.K., ed. (1983) *Lexicography: Principles and Practice*. London, Orlando..., Academic Press, Inc.

HERNÁNDEZ, Humberto (1989) «Los diccionarios de orientación escolar». A: *Lexicographica*. Series Major 28. Tübingen, Niemeyer.

ILSON, Robert, ed. (1987) *A spectrum of Lexicography*. Amsterdam/Philadelphia, John Benjamins Publ. Co.

LANDAU, Sidney I. (1989) *Dictionaries. The Art and Craft of lexicography*. Cambridge University Press.

QUEMADA, Bernard (1990) «La nouvelle lexicographie». A: «La lingüística aplicada», cicle de conferències. Barcelona, Universitat de Barcelona/Fundació Caixa de Pensions.

SEBEOK, Thomas A., ed. (1971) *Approaches to semiotics*. París, Mouton.

# Part II:

# LEXICOGRAFIA CATALANA

M. TERESA CABRÉ

# LA LEXICOGRAFIA CATALANA ACTUAL

s un fet que la lexicografia catalana disposa d'un bon nombre d'obres representatives que la situen en un nivell científic molt alt. Aquesta és l'opinió que sostenen una bona part de lingüistes i filòlegs que han estudiat aquesta temàtica. Així, Germà Colon afirma el 1976, amb motiu de la celebració del IV Col·loqui Internacional de Llengua i Literatura Catalanes, que la lexicografia catalana ha aconseguit un nivell molt alt de perfecció, fent referència a cinc obres lexicogràfiques catalanes fonamentals: el *Diccionari Català-Valencià-Balear* d'Alcover i Moll, el *Diccionari Aguiló*, el *Diccionari Balari*, el *Glossarium mediae latinitatis Cataloniae* i el *Diccionari Etimològic i Complementari de la Llengua Catalana* de Coromines, aleshores encara en fase d'elaboració:

> *Aquestes realitzacions i d'altres de més modestes ens permeten de considerar la lexicografia catalana com a una disciplina que ha assolit un grau elevat de perfecció. Si resultats tan tangibles han estat aconseguits fins ara, com mancaríem d'esperança?*

I Jordi Carbonell comparteix amb Colon la satisfacció pel nivell assolit per la lexicografia catalana. Així ho fa públic en la «Introducció» de la primera edició de la *Gran Enciclopèdia Catalana* (1979):

> *El català, llengua moderna de cultura, té una bibliografia lexicogràfica i gramatical qualitativament semblant a la de les llengües que han tingut un desenvolupament que podríem anomenar normal.*

*Caplletra 17* (Tardor 1994), pp. 139-156

Efectivament, sembla del tot cert afirmar que la lexicografia catalana és molt rica si considerem la qualitat de les seves obres més representatives i l'elevat nombre de diccionaris de què disposa, sobretot tenint present el pes demogràfic de la població de la zona catalanoparlant. Però una anàlisi més profunda del tema que posi en correlació la quantitat de diccionaris editats amb la tipologia d'aquests diccionaris, revela aspectes que convé tractar més críticament. Aquests punts crítics afecten, d'una banda, les mancances lexicogràfiques que es donen; i, d'altra banda, les repeticions de tipus de diccionaris, que semblen de tot punt innecessàries.

El propòsit d'aquest article és fer una panoràmica crítica de la lexicografia catalana actual, mostrant d'una banda els aspectes positius de la seva evolució, i, d'altra banda, els punts febles que presenta, molts dels quals són explicables per l'evolució política i social de la llengua i de la societat del domini català en el seu conjunt.[1]

## 1. LA LEXICOGRAFIA CATALANA CONTEMPORÀNIA: ASPECTES QUANTITATIUS

El nombre de diccionaris catalans publicats entre 1940 i 1990 és de 281, una xifra notable si tenim en compte el pes demogràfic de la societat catalanoparlant i les vicissituds polítiques de 1940 a 1977.

Aquesta producció es distribueix irregularment al llarg del període 1940-1990:

| | xifra absoluta | % |
|---|---|---|
| 1940-1950 | 1 | 0,3% |
| 1951-1960 | 4 | 1,4% |
| 1961-1970 | 24 | 8,5% |
| | | .....10% |
| 1971-1980 | 69 | 24,5% |
| 1981-1990 | 183 | 65,5% |
| | | .....90% |

Aquest quadre mostra com la producció lexicogràfica ha seguit una corba progressiva incessant: la dècada de 1940 a 1950 representa el 0,3% de la producció; i la de 1980 a 1990, el 65,5%. En els darrers vint anys (1970-1990) es concentra el 90% de la producció total, enfront del 10% d'un període de 40 anys (1940-1970).

Dins del període més productiu (1970-1990) l'edició de diccionaris es distribueix regularment i progressivament:

| | |
|---|---|
| 1970-1980 | 24,5% |
| 1980-1990 | 65,5% |

(1) El corpus de dades que constitueix la base d'aquest article és el catàleg de diccionaris de Cabré i Lorente (1991), ampliat fins al 1992. Aquest catàleg pretén ser un inventari exhaustiu de diccionaris catalans publicats a tot l'àmbit catalanoparlant. S'ha ampliat amb dos apèndixs no publicats: 1989-1990 i 1990-1992.

Algunes de les dades que aquí es presenten han estat difoses en una comunicació que vaig presentar al XX Congrés Internacional de Lingüística i Filologia Romànica celebrat a Zuric l'abril de 1992. Un fragment d'aquesta comunicació ha aparegut en les Actes d'aquest congrés.

I la mateixa progressió i regularitat s'observa en analitzar any per any les dues dècades:

Període 1970-1980:

| | xifra absoluta | % |
|---|---|---|
| 1971 | 1 | 1,4% |
| 1972 | 6 | 8,7% |
| 1973 | 6 | 8,7% |
| 1974 | 7 | 10% |
| 1975 | 9 | 13% |
| 1976 | 4 | 5,8% |
| 1977 | 11 | 16% |
| 1978 | 6 | 8,7% |
| 1979 | 11 | 16% |
| 1980 | 8 | 11,7% |
| TOTAL | 69 | |

Període 1980-1990:

| | xifra absoluta | % |
|---|---|---|
| 1981 | 15 | 8,2% |
| 1982 | 18 | 9,8% |
| 1983 | 15 | 8,2% |
| 1984 | 20 | 11% |
| 1985 | 27 | 14,5% |
| 1986 | 21 | 11,5% |
| 1987 | 14 | 7,5% |
| 1988 | 24 | 13% |
| 1989 | 9 | 4,9% |
| 1990 | 21 | 11,5% |
| TOTAL | 183 | |

Una fragmentació d'aquests dos períodes en segments de cinc anys mostra que hi ha un punt d'inflexió en el qual la producció de diccionaris augmenta espectacularment:

| | xifra absoluta | % |
|---|---|---|
| 1940-1945 | 0 | 0% |
| 1946-1950 | 1 | 0,4% |
| 1951-1955 | 1 | 0,4% |
| 1956-1960 | 3 | 1% |
| 1961-1965 | 6 | 2,1% |
| 1966-1970 | 18 | 6,4% |
| 1971-1975 | 29 | 10,3% |
| 1976-1980 | 40 | 14,2% |
| 1981-1985 | 94 | 33,5% |
| 1986-1990 | 89 | 31,6% |
| TOTAL | 281 | |

En efecte, dels 24 diccionaris de la dècada de 1960 a 1970, 23 es publiquen entre 1965 i 1970; aquests 23 diccionaris representen el 96% de la producció lexicogràfica d'aquesta dècada. Sembla ser, doncs, que el punt d'inflexió de la producció de diccionaris és pròpiament el 1965, ja que és a partir d'aquest any que es produeix una publicació en massa de diccionaris. Així, podem establir, d'acord amb aquesta nova dada, tres grans etapes lexicogràfiques:

1) 1940-1964, el desert lexicogràfic
2) 1965-1980, el desvetllament de la lexicografia
3) 1980-1990, la productivitat lexicogràfica

Aquests tres segments de temps ofereixen la següent panoràmica quantitativa:

| | xifra absoluta | % |
|---|---|---|
| 1940-1964 | 6 | 2,1% |
| 1965-1980 | 92 | 32,7% |
| 1980-1990 | 183 | 65,2% |

Les dades que acabem de mostrar sobre l'evolució de la producció lexicogràfica permeten suggerir inicialment algunes conclusions parcials sobre el tema:

1. La importància quantitativa de la producció de diccionaris és considerable si es té en compte la situació viscuda per la llengua i les condicions polítiques del context durant una bona part del període ressenyat.
2. Hom remarca una concentració gairebé exclusiva de la producció de diccionaris en el període de 1965-1990. L'any 1965, quinze anys abans del restabliment de les autonomies i de la legalització de l'ús de la llengua catalana, és el punt

d'inflexió en la producció lexicogràfica; fet que indica que la revitalització de la producció de diccionaris no s'inicia amb la recuperació institucional, sinó que té una base fonamentalment social.

3. Un cop iniciat el procés de revitalització lexicogràfica, la recuperació de les institucions és el factor que fa augmentar considerablement, i regularment, la producció de diccionaris. Els diccionaris publicats en el període 1980-1990 representen el 66% del total.

## 2. LA LEXICOGRAFIA CATALANA CONTEMPORÀNIA: ASPECTES TIPOLÒGICS

Una primera anàlisi dels diversos tipus de diccionaris que apareixen al llarg del període 1940-1990 mostra que aquests pertanyen a una àmplia gamma tipològica; però una anàlisi més detallada del tema revela una irregularitat manifesta tant en les mancances inexplicables de classes de diccionaris que s'observen al llarg d'aquest període de temps, com en les repeticions innecessàries de determinats tipus.

Amb aquesta finalitat hem establert 18 tipus lexicogràfics, tots ells corresponents a publicacions existents.[2] Cada tipus representa una classe de diccionaris definida a partir d'una sola dimensió, segons quin sigui l'aspecte més rellevant que fa que un diccionari pertanyi a una determinada classe:[3]

| | |
|---|---|
| G | diccionaris generals |
| E | diccionaris enciclopèdics |
| V | diccionaris visuals |
| D | diccionaris dialectals |
| ET | diccionaris etimològics |
| T | diccionaris terminològics |
| C/C | diccionaris bilingües català/castellà |
| B | diccionaris correctius |
| BIL | diccionaris bilingües (excepció de C/C) |
| ESC | diccionaris escolars |
| HOM | diccionaris d'homònims |
| SIN/ANT | diccionaris de sinònims i/o antònims |
| LOC | diccionaris de locucions |
| INV | diccionaris inversos |
| IDEOL | diccionaris ideològics |
| ARG | diccionaris d'argot |
| VOCBAS | vocabularis bàsics |
| ORT | diccionaris ortogràfics i de pronunciació |
| ESP | casos especials |

(2) L'absència en la classificació de determinats tipus de diccionaris és un indicador de la seva inexistència en la lexicografia catalana fins al 1990.

(3) Una classificació multidimensional dels tipus de diccionaris que tingués en compte la diversitat d'informacions que inclou cada diccionari, més que no pas el tipus prioritari de diccionari que constitueix, sembla que no conduiria a resultats substancialment diferents dels que permet obtenir una classificació unidimensional.

Una anàlisi del grau de correlació entre la riquesa quantitativa dels diccionaris editats i la diversitat que presenta el panorama lexicogràfic d'aquest període ens permet analitzar si la riquesa quantitativa es correspon amb una gran diversitat de diccionaris. Per abordar aquesta qüestió ens valem de tres observacions:

a) com es distribueixen els cinc tipus lexicogràfics més freqüents
b) en quin període de temps apareix el primer exemplar de cada tipus de diccionari
c) amb quina freqüència apareix cada tipus lexicogràfic

a) L'evolució dels diversos tipus de diccionaris en períodes de 10 anys dóna els resultats següents:

| | 1940-50 | 1951-60 | 1961-70 | 1971-80 | 1981-90 | total |
|---|---|---|---|---|---|---|
| G | | | 2 | 5 | 9 | 16 |
| E | | | 1 | | 1 | 2 |
| V | | | 1 | 2 | 8 | 11 |
| D | | | 1 | | | 1 |
| ET | | | | 1 | | 1 |
| T | 1 | 2 | 5 | 29 | 109 | 146 |
| C/C | | 1 | 9 | 14 | 17 | 41 |
| B | | | 2 | 1 | 5 | 8 |
| BIL | | | 1 | 4 | 15 | 20 |
| ESC | | | | 2 | 5 | 7 |
| HOM | | | | | 1 | 1 |
| SIN/ANT | | | 2 | 4 | 1 | 7 |
| LOC | | 1 | | 2 | 1 | 4 |
| INV | | | | | 3 | 3 |
| IDEOL | | | | 1 | 1 | 2 |
| ARG | | | | 1 | | 1 |
| VOCBAS | | | | | 1 | 1 |
| ORT | | | | 1 | 1 | 2 |
| ESP | | | | 2 | 4 | 6 |
| TOTAL | 1 | 4 | 24 | 69 | 183 | 281 |

on la mitjana (X) és igual a 14,7, i la desviació estàndard (n) és igual a 32,3. Aquesta desviació tan alta revela una dispersió dels tipus de diccionaris molt superior a la mateixa mitjana.

Les dades que aquest quadre ofereix mostren que, del total dels 18 tipus de diccionaris avaluats, només cinc tipus superen la mitjana, amb aquest ordre decreixent:

1) els diccionaris terminològics
2) els diccionaris bilingües català-castellà
3) els diccionaris bilingües d'altres llengües
4) els diccionaris generals
5) els diccionaris visuals

Sobre aquests resultats és important de remarcar que sembla lògic que la terminologia i els diccionaris bilingües d'altres llengües formin part de la llista dels tipus de diccionaris més freqüents, ja que presumiblement cada diccionari o bé ofereix una temàtica específica (les terminologies), o bé inclou llengües diverses (els diccionaris bilingües). Sorprèn, en canvi, la gran quantitat de diccionaris bilingües català-castellà que mostra el quadre, quantitat que contrasta amb la pobresa de la gran majoria dels altres tipus. I aquests resultats semblen suggerir, si analitzem les dades independentment del context en què s'han produït, una certa manca de racionalitat en la producció lexicogràfica.

La proporció que s'estableix entre els diccionaris bilingües català-castellà i els diccionaris bilingües d'altres llengües, per posar un cas, és realment desequilibrada. D'un total de 61 diccionaris bilingües en general, 41 són català-castellà (67%) i 20, d'altres llengües (33%):

| | | |
|---|---|---|
| diccionaris bilingües català-castellà | 41 | 67% |
| diccionaris bilingües d'altres llengües | 20 | 33% |

L'observació del ritme d'aparició dels diversos tipus lexicogràfics al llarg d'aquest període explicita que els cinc tipus de diccionaris apareixen progressivament seguint una corba ascendent, si bé en alguns tipus la corba és més pronunciada i abrupta que en d'altres:

| | 1940-1964 | 1965-1970 | 1971-1990 |
|---|---|---|---|
| T | 8 | 29 | 109 |
| C/C | 10 | 14 | 17 |
| BIL | 1 | 4 | 15 |
| G | 2 | 5 | 9 |
| V | 1 | 2 | 8 |

És interessant de destacar el grau d'accentuació de la corba progressiva que presenten les terminologies (8-29-109) i els diccionaris bilingües d'altres llengües (1-4-15) en contrast amb la suavitat de la corba (encara que també progessiva en tots els casos) que correspon als bilingües català-castellà. Aquest fet separa els casos de tipus que, malgrat que es presenten al llarg de les diverses dècades, en un període determinat

augmenten per damunt del que s'hauria d'esperar i produeixen un nombre de diccionaris molt important respecte al nombre de diccionaris del mateix tipus que han aparegut en el període anterior. Sembla lògic atribuir aquest comportament diferenciat a l'aparició d'algun factor sociopolític específic, que fa esclatar la producció d'un determinat tipus de diccionari.

b) L'anàlisi de l'aparició del primer diccionari de cada tipus (marcat al quadre següent amb un *) pot ser útil per determinar en quin segment de temps es produeix una diversificació lexicogràfica més important. En efecte, les dades mostren que entre 1960 i 1980 apareixen pràcticament tots els tipus de diccionaris de què la llengua catalana disposa. Aquesta diversificació s'atura gairebé radicalment en la dècada de 1980-1990, que només presenta dos nous tipus respecte als de les dècades anteriors: els diccionaris inversos, facilitats per la difusió de la informàtica, i els d'homònims.

| 1940-50 | 1951-60 | 1961-70 | 1971-80 | 1981-90 |
|---|---|---|---|---|
| T* | T | T | G | G |
| | | C/C | T | E |
| | | G* | C/C | V |
| | | C/C* | B | T |
| | | E* | BIL | C/C |
| | | V* | SIN/ANT | B |
| | | D* | V* | BIL |
| | | B* | ET* | ESC |
| | | BIL* | ESC* | SIN/ANT |
| | | SIN/ANT* | LOC* | LOC |
| | | | ARG* | VOCBAS |
| | | | ORT* | ORT |
| | | | IDEOL* | IDEOL |
| | | | | HOM* |
| | | | | INV* |

Una interpretació d'aquestes dades ens fa pensar que la capacitat de concebre i publicar nous diccionaris per part d'autors i d'editorials sembla ser bastant constreta, i no pas perquè la producció lexicogràfica catalana estigui totalment coberta. La justificació sembla que rau en les regles del mercat: es produeixen només aquells tipus de diccionaris de venda segura, encara que repeteixin (sovint sense gaires divergències) alguns tipus ja existents. El conservadorisme comercial s'imposa en aquest cas al risc econòmic que podria suposar treure al mercat un tipus de diccionari nou. I aquesta posició conservadora s'imposa també a la voluntat de dotar la llengua catalana d'un

inventari lexicogràfic similar al d'altres llengües de cultura, tot i que proporcional a la població parlant.

c) Observarem finalment amb quina freqüència es donen els diferents tipus de diccionaris al llarg dels diversos segments de temps per establir el grau de constància de cada un dels tipus lexicogràfics, i posteriorment creuarem aquestes dades amb l'índex de rellevància de la repetició:

| 1966-70 | 1971-75 | 1976-80 | 1981-85 | 1986-90 | totals |
|---|---|---|---|---|---|
| | G | G | G | G | 4 |
| E | | | E | | 2 |
| V | V | V | V | | 5 |
| | | | | ET | 1 |
| | | D | | | 1 |
| T | T | T | T | T | 5 |
| C/C | C/C | C/C | C/C | C/C | 5 |
| B | B | | B | B | 4 |
| BIL | BIL | | BIL | BIL | 4 |
| | | ESC | ESC | ESC | 3 |
| | | | | HOM | 1 |
| SIN | SIN | | | SIN | 3 |
| | LOC | LOC | LOC | | 3 |
| | | | INV | INV | 2 |
| | | ARG | | | 1 |
| | | | VOCBAS | | 1 |
| | ORT | | | ORT | 2 |
| | | IDEOL | | IDEOL | 2 |

D'acord amb aquestes dades, podem establir tres sèries de tipus lexicogràfics segons el grau de constància d'aparició:

1) tipus constants

repetició 5:

terminologia
bilingües català-castellà
visuals

repetició 4:

generals
bilingües d'altres llengües
correctius

repetició 3:

escolars
sinònims i antònims
locucions

2) Tipus inconstants

repetició 1:

dialectals
etimològics
homònims
argot
enciclopèdics

repetició 2:

ideològics
inversos

El màxim grau d'inconstància d'un tipus lexicogràfic és un indicador de la necessitat de reforçar aquest tipus lexicogràfic amb noves publicacions de característiques similars, si no actua de neutralitzador l'índex de rellevància. El grau màxim de constància suposa una repetició innecessària, si no és neutralitzat per l'índex de rellevància.

L'índex de rellevància de la repetició és un indicador que permet caracteritzar els diversos tipus de diccionaris segons la redundància de la repetició. Un tipus de diccionari és redundant si la repetició, en lloc de conduir a un diccionari de nova informació, condueix simplement a un diccionari del mateix tipus amb un augment o una disminució de la informació. Així, dues terminologies, que pertanyen al mateix tipus lexicogràfic, seran redundants si tracten de la mateixa temàtica, encara que hagin estat concebudes per a usuaris diferents i una contingui només una selecció de les dades de l'altra. En canvi, no seran redundants dues terminologies de temàtica diferent o dos diccionaris bilingües de llengua diferent.

L'índex de rellevància permet distingir, dintre dels diccionaris repetits, els que ho són de forma rellevant i, per tant, introdueixen un nou diccionari, dels que ho són de manera irrellevant, i solen ser innecessaris en qualitat de tipus lexicogràfic.[4]

(4) Per a aquest índex de rellevància, no hem tingut en compte en aquest article la diferenciació a partir dels usuaris a què està destinat un diccionari. Des d'aquesta perspectiva algunes de les repeticions que considerem redundants podrien no ser-ho. Tot i així, aquest matís sembla no afectar la relació rellevància-irrellevància dels bilingües català-castellà.

1) Diccionaris de repetició rellevant:
terminologia
bilingües d'altres llengües

2) Diccionaris de repetició irrellevant:
bilingües català-castellà
visuals

Sembla clar, com ja hem dit, que tant la terminologia com els diccionaris bilingües d'altres llengües són presumiblement tipus no redundants perquè se suposa que cada exemplar introdueix o bé una temàtica nova o bé una llengua diferent. No sembla ser aquest, però, el cas dels diccionaris bilingües català-castellà, dels visuals, dels correctius o dels generals. Un repàs dels diccionaris visuals, per exemple, revela que tots ells s'adrecen a infants amb l'única finalitat de facilitar l'aprenentatge de la lectura, i que no es publica en aquest període cap diccionari visual al mercat, ni general ni especialitzat, destinat a adults, ni tan sols orientat als escolars d'ensenyament primari més avançat o d'ensenyament mitjà. La repetició innecessària d'aquest tipus lexicogràfic (vista la realitat i les necessitats) és, doncs, palesa.

Pel que fa als diccionaris generals, es pot comprendre clarament que, vista la situació restrictiva del diccionari normatiu i la manca d'elaboració d'altres diccionaris generals de referència per part de les institucions acadèmiques, el lèxic català comú havia de comptar amb altres diccionaris generals que n'ampliessin i en modernitzessin l'inventari general. D'aquí que, no pas en un període immediat, però sí cap a la dècada dels 60, comencen a publicar-se alguns diccionaris generals (pocs, però de manera ininterrompuda), que, en un cert sentit i sense discutir-ne l'autoritat normativa, són alternatius al *Diccionari General de la Llengua Catalana*. La falta d'actualització del diccionari normatiu condueix les editorials a publicar diccionaris que responen més adequadament a les noves necessitats dels usuaris.

D'entre els diccionaris bilingües d'altres llengües modernes (no considerem el cas del llatí i del grec) només tenen repetició els corresponents a 3 llengües, tots els altres són exemplars únics per a la llengua en qüestió:

| | |
|---|---|
| anglès | 5 |
| francès | 4 |
| alemany | 2 |
| hongarès | 1 |
| italià | 1 |
| japonès | 1 |
| rus | 1 |

En contrast amb els diccionaris generals i els bilingües d'altres llengües, no sembla tenir justificació científica (encara que sociolingüísticament és explicable) la repetició, clarament innecessària, de diccionaris correctius (també anomenats diccionaris de barbarismes), que no introdueixen, ni en els continguts ni en els destinataris, diferències importants respecte als que ja hi havia publicats. Els diccionaris correctius suposen una de les repeticions més freqüents en la lexicografia catalana d'un determinat període (1940-1970) i estan tots ells destinats a «corregir» la interferència de la llengua castellana en la catalana. En la darrera dècada (1980-90) aquest tipus de diccionari va desapareixent progressivament.

Una breu anàlisi de l'apèndix dels diccionaris publicats el 1990 (que no s'havien recollit en el moment que vam fer aquest estudi) indica que en la producció de 1990 s'hi han d'afegir 8 terminologies, dada que no fa canviar els resultats comentats anteriorment.

Pel que fa als diccionaris editats el 1991 i el 1992, els resultats que tenim són els següents:

| | 1991 | 1992 |
|---|---|---|
| G | 1 | – |
| E | – | – |
| V | – | 1 |
| D | 1 | – |
| ET | – | – |
| T | 39 | 25 |
| C/C | – | – |
| B | – | – |
| BIL | 2 | 1 |
| ESC | – | – |
| HOM | – | – |
| SIN/ANT | – | 1 |
| LOC | – | – |
| INV | – | – |
| IDEOL | – | – |
| ARG | 2 | – |
| VOCBAS | 1 | – |
| ORT | – | – |
| ESP | 1 | 1 |
| TOTAL | 47 | 29 |

L'única diferència que sembla alterar la corba manifestada fins l'any 1990 és la desaparició dels diccionaris centrats en la relació català-castellà, sovint amb propòsits prescriptius.

## 3. SOBRE ELS DICCIONARIS BILINGÜES CATALÀ-CASTELLÀ

Els diccionaris bilingües català-castellà constitueixen el segon grup de diccionaris més freqüent en el panorama lexicogràfic de la llengua catalana.

| | 1940-1990 | 1965-1990 |
|---|---|---|
| T | 51,2% | 54,1% |
| C/C | 14,6% | 15,2% |
| B | 2,8% | 3% |
| BIL | 7,1% | 7,4% |

Aquest tipus de diccionari representa el 14,6% de la producció de 1940 a 1990 (d'un total de 281 diccionaris) i el 15,2% de la producció de 1965 a 1990.

| 1940-50 | 1951-60 | 1961-70 | 1971-80 | 1981-90 | total |
|---|---|---|---|---|---|
| 0 | 1 | 9 | 14 | 17 | 41 |

14,6% (1940-1990)
15,2% (1965-1990)

Des del punt de vista de la constància, és un tipus molt freqüent al llarg del període 1965-1990. Des del punt de vista de la rellevància de la seva repetició, és clarament redundant.

L'evolució que segueix aquest tipus de diccionari, en contrast amb els altres tipus, és la següent:

| | 1940-50 | 1951-60 | 1961-70 | 1971-80 | 1981-90 | total |
|---|---|---|---|---|---|---|
| **C/C** | | **1** | **9** | **14** | **17** | **41** |
| G | | | 2 | 5 | 9 | 16 |
| E | | | 1 | | 1 | 2 |
| V | | | 1 | 2 | 8 | 11 |
| D | | | 1 | | | 1 |
| ET | | | | 1 | | 1 |
| T | 1 | 2 | 5 | 29 | 109 | 146 |
| B | | | 2 | 1 | 5 | 8 |
| BIL | | | 1 | 4 | 15 | 20 |
| ESC | | | | 2 | 5 | 7 |
| HOM | | | | | 1 | 1 |
| SIN/ANT | | | 2 | 4 | 1 | 7 |
| LOC | | 1 | | 2 | 1 | 4 |
| INV | | | | | 3 | 3 |
| IDEOL | | | | 1 | 1 | 2 |
| ARG | | | | 1 | | 1 |
| VOCBAS | | | | | 1 | 1 |
| ORT | | | | 1 | 1 | 2 |
| ESP | | | | 2 | 4 | |
| TOTAL | 1 | 4 | 24 | 69 | 183 | 281 |

Durant la dècada de 1960-1970, i concretament l'any 1966, comença l'edició de diccionaris bilingües català-castellà. A partir d'aquesta data, la producció de diccionaris d'aquest tipus segueix una corba progressivament ascendent al llarg d'aquesta dècada:

| | |
|---|---|
| 1961 | 1 |
| 1962 | 0 |
| 1963 | 0 |
| 1964 | 0 |
| 1965 | 1 |
| 1966 | 2 |
| 1967 | 0 |
| 1968 | 2 |
| 1969 | 2 |
| 1970 | 1 |
| Total | 9 |

I la corba progressiva continua també en el període posterior:

| | |
|---|---|
| 1966-70 | 7 |
| 1971-80 | 14 |
| 1981-90 | 17 |
| Total | 39 |

La quantitat absoluta de diccionaris bilingües català-castellà (38) en el període 1966-1990 és de per si ben significativa, sobretot si es posa en correlació amb els altres tipus freqüents de constància irrellevant:

| | | |
|---|---|---|
| **C/C** | **38** | **14,1%** |
| G | 14 | 5,2% |
| V | 11 | 4% |
| D | 1 | 0,4% |
| ET | 1 | 0,4% |
| SIN | 7 | 2,6% |
| IDEOL | 2 | 0,7% |
| INV | 2 | 0,7% |
| ORT | 2 | 0,7% |

Però també és significativa en relació amb els diccionaris freqüents de constància rellevant:

| | | |
|---|---|---|
| **C/C** | **38** | **14,1%** |
| T | 142 | 52,6% |
| BIL | 20 | 7,4% |

Ara bé, la importància d'aquest tipus de diccionari en la lexicografia catalana és encara més remarcable si als 38 bilingües català-castellà s'hi afegeixen tots aquells diccionaris generals i especialitzats bilingües català-castellà, com és el cas dels correctius i de les terminologies bilingües (1965-1990). Així, la producció de diccionaris que presenten la relació català-castellà representa gairebé la meitat de tota la producció lexicogràfica catalana.

| | | |
|---|---|---|
| **C/C** | **38** | **14,1%** |
| B | 8 | 3% |
| T | 71 | 26,3%[5] |
| Total | 117 | 43,4% |

(5) Incloem en aquest grup les terminologies català-castellà. D'un total de 142 terminologies, 71 són català-castellà i representen el 50% de terminologia.

## 4. ALGUNS COMENTARIS SOBRE ELS RESULTATS

De l'anàlisi qualitativa de la producció lexicogràfica catalana es poden extreure inicialment algunes conclusions que semblen bastant definitives:

1. La correlació entre les dades purament quantitatives sobre els diccionaris publicats i els tipus de diccionaris de què la llengua catalana disposa revela que la quantitat no es correspon necessàriament amb la diversitat; o en altres mots, que la riquesa quantitativa que la primera anàlisi mostra no es correspon necessàriament amb una gran diversitat lexicogràfica, sinó que més aviat demostra que en el panorama de la lexicografia catalana actual es produeixen una sèrie de repeticions racionalment innecessàries, encara que històricament explicables.
2. Aquesta correlació també mostra una tendència general ascendent en gairebé tots els tipus de diccionaris, si bé aquesta evolució és de caràcter més abrupte en uns casos que no pas en d'altres. Aquesta circumstància està relacionada amb algun esdeveniment polític que condiciona les característiques del mercat dels diccionaris. En canvi, la progressió suau d'altres tipus es relaciona o bé amb la progressiva recuperació natural de la llengua dins del marc sociopolític més estabilitzat, o bé amb determinats «tics» socials respecte a la llengua.
4. L'anàlisi de l'aparició del primer diccionari de cada tipus revela com en el període intermedi (1965-1980) la llengua catalana disposa de gairebé els mateixos tipus de diccionaris que en l'actualitat. Aquesta constatació és una mostra de l'interès relatiu de les editorials de diccionaris per completar el panorama lexicogràfic, probablement a causa de la incertesa de mercat. S'observa una tendència més acusada a «repetir» tipus de diccionaris ja existents que tenen un mercat segur que no pas a «diversificar» l'inventari de tipus, dotant la llengua catalana de la diversitat d'eines lexicogràfiques que tota llengua moderna de cultura posseeix.[6]
5. Una anàlisi dels diccionaris bilingües revela que la proporció de diccionaris bilingües català-castellà, que constitueixen un tipus repetitiu irrellevant, és molt alta (14%) en relació amb la dels bilingües d'altres llengües (7,4%), que són un tipus repetitiu rellevant i, per tant, la seva proporció hauria de ser molt més alta que la dels primers.
6. La relació català-castellà constitueix un dels principals eixos de la concepció i edició de diccionaris de llengua catalana de 1965 ençà, encara que sembla endevinar-se una certa tendència a la reducció de publicacions d'aquestes característiques. Aquest paràmetre específic de la lexicografia catalana es fa explícit de manera espectacular si, al nombre de diccionaris pròpiament bilingües català-castellà, li sumem altres tipus de diccionaris que inclouen el castellà al costat del català, com les terminologies bilingües català-castellà i els diccionaris cor-

(6) Aquesta apreciació sembla que ha començat a canviar en els tres darrers anys. Aquest canvi, però, serà objecte d'un estudi posterior.

rectius. El nombre total d'aquests diccionaris representa pràcticament la meitat de tota la producció lexicogràfica catalana contemporània.

És evident que qualsevol llengua moderna de cultura disposa d'una sèrie d'eines lingüístiques que en codifiquen els recursos expressius. La llengua catalana, com molt enraonadament afirma Carbonell a la «Introducció» de la *Gran Enciclopèdia Catalana*, també posseeix evidentment aquestes eines en forma de diccionaris i gramàtiques, fet que la situa al nivell d'altres llengües de cultura.

Però si bé no hi ha dubte que la llengua catalana disposa de diccionaris que codifiquen el seu lèxic i de gramàtiques, més o menys completes, que recullen les regularitats del seu sistema gramatical (la fonologia, la morfologia, la sintaxi) i la seva representació gràfica convencional (l'ortografia), no per això es pot afirmar mecànicament que aquest panorama de recursos és totalment satisfactori, si no es compta amb una revisió crítica no només de l'inventari d'aquests recursos sinó també de la seva qualitat i adequació. I aquesta revisió crítica encara no s'ha fet.

Una anàlisi d'aquesta mena permetria detectar, en primer lloc, si el conjunt lexicogràfic català cobreix o no les necessitats dels usuaris; en segon lloc, permetria saber si els recursos de què es disposa són, no només de qualitat, sinó també adequats als diversos usuaris de la llengua.

Amb una informació periòdica correcta sobre el panorama de la lexicografia catalana en correlació amb el d'altres llengües modernes de cultura, el món editorial català (impulsat pel món acadèmic i institucional) disposaria de dades suficients per planificar la seva producció lexicogràfica, situant la llengua catalana al mateix nivell real d'altres llengües amb què se la pot comparar. I amb una anàlisi crítica que disposés de models d'avaluació lexicogràfica (actualment inexistents), se situaria la lexicografia catalana en el lloc real que li correspon, a partir del qual ha de planificar la seva evolució.

M. Teresa Cabré Castellví
*Universitat de Barcelona*

## REFERÈNCIES BIBLIOGRÀFIQUES

Cabré, M. T.; Lorente, M. (1991) *Els diccionaris catalans. De 1940 a 1988*. Barcelona, Publicacions de la Universitat de Barcelona.

Cabré, M. T. ; Bastardas, J. (1993) «Normativa lèxica del català: l'actualització del diccionari normatiu». A: *Actes del XX Congrés Internacional de Lingüística i Filologia Romàniques*. Zuric.

CABRÉ, M. T. (1993) «Sobre la lexicografia catalana contemporània». A: *Actes del XX Congrés Internacional de Lingüística i Filologia Romàniques*. Zuric.

CARBONELL, J. (1979) «Introducció» a la *Gran Enciclopèdia Catalana*. Barcelona, Enciclopèdia Catalana.

COLON, G. (1977) «La lexicografia catalana: realitzacions i esperances». A: Colon, G. (1977) *Actes del IV Col·loqui Internacional de Llengua i Literatura Catalanes*. Barcelona, Publicacions de l'Abadia de Montserrat.

LORENTE, M. (1991) «Funció lexicogràfica i contacte de llengües: Els diccionaris de barbarismes». A: *Actes del IX Col·loqui Internacional de Llengua i Literatura Catalanes (Alacant-Elx)*. Barcelona, Publicacions de l'Abadia de Montserrat (en premsa).

QUEMADA, B. (1990) «La nouvelle lexicographie». A: Cabré, M. T. (coord.) (1990) *La Lingüística aplicada avui*. Barcelona, Publicacions de la Universitat de Barcelona-Fundació Caixa de Pensions.

JORDI COLOMINA I CASTANYER

# UN TAST DELS VOCABULARIS BILINGÜES VALENCIANS ACTUALS

ovint les valoracions dels repertoris lexicogràfics es fan a partir d'uns quants mots triats «a l'atzar». Per evitar aquesta mena de subjectivitat, he fet un tast una mica més substanciós –setanta mots– a deu vocabularis bilingües valencià-castellà i castellà-valencià –publicats entre 1967 i 1993–. Per raons d'eficàcia he cregut més convenient treballar amb els vocabularis castellà-valencià, car el fet de partir del mot castellà permet comparar més fàcilment les diverses solucions proposades pels lexicògrafs. He procurat que els setanta mots triats fossen mots conflictius des del punt de mira valencià i que permetessen observar: *a*) el grau de particularisme/unitarisme, *b*) el grau de fidelitat a la normativa lèxica –i de vegades fonètico-gràfica–, i *c*) la riquesa lèxica i la variació sinonímica.

He agrupat els mots en els vuit camps semàntics següents: *a*) el món rural i la meteorologia: 1 *arado*, 2 *colmena*, 3 *estrella*, 4 *lloviznar* i 5 *chaparrón*; *b*) gastronomia: 6 *bacalao*, 7 *bocadillo*, 8 *cucurucho (de papel)*, 9 *palomitas de maíz*, 10 *solomillo*, 11 *requesón* i 12 *tragón, glotón*; *c*) fruita i verdura: 13 *patata*, 14 *pimiento (fruto)*, 15 *fresa (fruta)*, 16 *níspero (fruto)* i 17 *zanahoria*; *d*) la casa: 18 *chimenea*, 19 *hollín*, 20 *servilleta*, 21 *lavaplatos, lavavajillas*, 22 *bandeja*, 23 *herramienta*, 24 *basura*, 25 *basurero (oficio)* i 26 *bombilla*; *e*) fitonímia: 27 *tomillo*, 28 *romero*, 29 *junco*, 30 *juncia*, 31 *hiedra*, 32 *fresno*, 33 *jazmín* i 34 *raíz*; *f*) zoonímia: 35 *cola*, 36 *aguijón*, 37 *cangrejo*, 38 *lagarto*, 39 *saltamontes*, 40 *mariposa*, 41 *murciélago*, 42 *pavo*, 43 *jilguero*, 44 *golondrina*, 45 *abubilla* i 46 *jauría*; *g*) el vestit i el cos humà: 47 *abanico*, 48 *deshilar*, 49 *deshilachar*, 50 *calzoncillos*, 51 *latir*, 52 *estómago*, 53 *cogote, pescuezo*, 54 *cosquillas*, 55 *agujetas*, 56 *bálsamo*, 57 *muñeca (del brazo)* i 58 *párpado*;

*Caplletra 17* (Tardor 1994), pp. 157-178

vària: 59 *pequeño*, 60 *pequeñito*, 61 *larguirucho*, 62 *defender*, 63 *resbalar*, 64 *salpicar*, 65 *empujar*, 66 *empujón*, 67 *hinchar*, 68 *en cuclillas*, 69 *juguete* i 70 *llamarada*.

Les obres estudiades són les següents: *Vocabulari* de Francesc Ferrer Pastor (1967, 1972 i 1977),[1] *Diccionari Galmar* (1977), *Vocabulari Bàsic* del GAV (1980), *Vocabulari* del GAV (1983), *Vocabulari Elemental* d'Emili Miedes (1983), *Vocabulari Fullana-Miedes* (1985), *Diccionari Gregal* de Vicent Pascual (1987), *Vocabulari Fonamental* d'Enric Valor (1988), *Vocabulari Brúixola* (1990) i *Diccionari* de la Real Acadèmia de Cultura Valenciana (1992-1993).[2] Perquè es veja d'on beu algun vocabulari –p.e. F. Pastor, Galmar, Pascual o Valor– els he confrontats quasi sistemàticament amb dos vocabularis català-castellà –Albertí (1961 i 1975) i DEC 1985– i més esporàdicament amb Moll (1977). També faig al·lusió un parell de vegades a F. Pastor-J. Giner (1956), F. Ferrer Pastor (1960 i 1985) i al vocabulari «diferencial» de Recio (1985). Entre claudàtors indique mots que malgrat no trobar-se en el castellà-valencià, sí que apareixen en el valencià-castellà.

1. cast. **arado**.[3] Albertí 1961, Galmar 1977, DEC 1985 i Pascual 1987: *arada, aladre, llaura*; F. Pastor 1967: *arada, aladre, llaura, forcat*; GAV 1983 i Acadèmia 1993: *aladre*; Miedes 1983: *forcat*; Fullana-Miedes 1985: *aladre, charuga, cheruga*; Valor 1988: *aladre*, reg. *arada*; Brúixola 1990: *arada*.

2. cast. **colmena**. Albertí 1961 i Galmar 1977: *arna, buc, rusc, abell, arnera, tou, casera*; F. Pastor 1967: *arna, buc, rusc*; DEC 1985 i Pascual 1987: *rusc, buc, arna*; Moll 1977: *rusc, buc d'abelles, casa d'abelles, casera, arna, tou, baso*; GAV 1980, Miedes 1983 i Fullana-Miedes 1985: *colmena*; GAV 1983: *colmena, abellar;* ; Valor 1988: *rusc, buc d'abelles*; Brúixola 1990: *abeller, rusc, buc, arna, tova*<sic>;[4] Acadèmia 1993: *colmena, gayola, baso*.

3. cast. **estrella**. Albertí 1961, F. Pastor 1967, Galmar 1977, Moll 1977 i DEC 1985: *estel, estrella*; GAV 1980, Miedes 1983 i Acadèmia 1993: *estrela*; GAV 1983: *estrel, estrela, estela, estel*; Fullana-Miedes 1985: *estrela, estel*; Pascual 1987: *estel, estrella, estrela*; Valor 1988: *estrela / estel*; Brúixola 1990: *estrela, estrella, estel*.

4. cast. **lloviznar**. Albertí 1961: *plovisquejar, ploviscar, plovinejar, roinar, roinejar, brusquinejar, borrimejar, repixar*; F. Pastor 1967: *plovisquejar, ploviscar*; Galmar 1977: els vuit sinònims d'Albertí (1961) més *brusquejar, gotellimar, gotellinar*;[5] GAV 1980: *purnejar*; GAV 1983: *plovinejar*; Miedes 1983: *plovisnar, purnejar*; Fullana-Miedes 1985: *plovinejar, gotejar, pluinejar, pluinar*; DEC 1985: *plovisquejar, ploviscar, plovinejar, roinejar, repixar*; Pascual 1987: *plovinejar, ploviscar, plovisquejar, repixar*; Valor 1988: *plovitejar, plovisquejar, llagrimejar* 'casi imperceptiblemente'; Brúixola 1990: *plovisquejar, ploviscar, gotejar*; Acadèmia 1993: *purnejar, plovitejar, borrimar, gotejar, gotinyar*.

(1) Quasi sempre coincideixen les tres edicions: llavors cite només F. Pastor 1967; la de 1972 és de vegades una mica més reduïda que la de 1977, la definitiva i la que sense més variacions –només la data d'impressió– s'ha publicat fins a l'hora d'ara.

(2) Vull agrair a l'amic i col·lega Emili Casanova els seus suggeriments, sempre ben valuosos; i a ell i a Lluís Alpera, Joan J. Ponsoda, Biel Sansano i Josep Tormo el seu suport bibliogràfic.

(3) GAV 1980 no enregistra *arado*.

(4) *Tova* és errada per *tou* (Andorra, Alt Pallars, Organyà, *DCVB*).

(5) Tots onze sinònims són al *DFa*, que no sempre presenta la mateixa riquesa sinonímica. Em fa l'efecte que Galmar reprodueix ací alguna edició posterior –i doncs ampliada– d'Albertí (1961).

5. cast. **chaparrón.**[6] Albertí 1961 i Galmar 1977: *ruixat, ruixada, ram, ramassat, ramassada, grop, aiguarrada, glopada*; F. Pastor 1967: *ruixada, ramassada, aiguarrada*; GAV 1983: *aiguafort, aiguaço, espetec*; Miedes 1983: *arruixo, ruixada*; Fullana-Miedes 1985: *chapallo, ramassada, arruixo*; (chaparrón de agua <sic>) *aiguaço*; DEC 1985: *xàfec, ruixat*; Pascual 1987: *arruixada (o ruixada), ruixat, xàfec*; Valor 1988: *arruixada, ruixat, ramassada*; Brúixola 1990: *xàfec, ruixat, ramassada*; Acadèmia 1993: *arruixada, arruixo, aiguarrada, aiguaço, sacso d'aigua, alaboc, ramassada, glopada, garrotada, diluviada, chapallo* [*araboga* 'aguanieve' <sic>].

6. cast. **bacalao.** Albertí 1961, F. Pastor 1967, Galmar 1977, Moll 1977 i DEC 1985: *bacallà*; GAV 1980 i 1983: *abadejo*; Miedes 1983: *abaejo*; Fullana-Miedes 1985: *abaejo, bacallar, abadeig*; Pascual 1987: *bacallà, abadejo*; Valor 1988: *bacallà* [*abadeig*]; Brúixola 1990: *abadejo, bacallà*; Acadèmia 1993: *bacallar, abadejo.*

7. cast. **bocadillo.**[7] Tots els vocabularis donen *entrepà*, i a més: Albertí 1961: *sandvitx, badall*; F. Pastor 1967: *sandvitx*; Galmar 1977: *badall, emparedat, sandvitx*; GAV 1983: *mossegadeta*; DEC 1985: *badall*; Pascual 1987: *badall, emparedat*; Acadèmia 1993: *cantell.*[8]

8. cast. **cucurucho (de papel).**[9] Albertí 1961, F. Pastor 1967 i Galmar 1977: *cucurull, cucurutxo, paperina*; GAV 1983: *paperada, paperina, cucurruig*; Miedes 1983: *cucuruig, mesureta*; Fullana-Miedes 1985: *paperina, paperada, mesura*; DEC 1985: *cucurull, cucurutxo*; Pascual 1987: *cucurull, cucurutxo, paperina*; Valor 1988: *cucurull, paperina*; Brúixola 1990: *cucurutxo, mesura, paperina*; Acadèmia 1993: *mesura, paperada, papera.*

9. cast. **palomitas de maíz.**[10] Albertí 1961: *bufa, rosa*; Miedes 1983: *borles, roses*; DEC 1985: *roseta, crispeta*; Pascual 1987: *rosa, roseta*; Brúixola 1990: *borla, crispeta*; Acadèmia 1993: *rosa, esclafito, clocha.*

10. cast. **solomillo.**[11] Albertí 1961 i Galmar 1977: *llom* (en carnisseria); F. Pastor 1967: *llom, entrecot;* GAV 1983: *rellomello, sanguisolet* <sic>; Miedes 1983: *llomello*; Fullana-Miedes 1985: *rellomello*; DEC 1985: *filet, rellom*; Pascual 1987: *filet*; Valor 1988: *llom* / (de cerdo) *llomell* o *llomello;* Brúixola 1990: *llom;* Acadèmia 1993: *rellomello, llistell* <sic>.[12]

11. cast. **requesón.**[13] Albertí 1961: *mató, recuit, brossat*; F. Pastor 1967: *mató, recuit*; Albertí 1975 i Galmar 1977: *mató, recuit, brossat, brull*; GAV 1983: *mato,*[14] *formaget*; Miedes 1983: *brull, brullo*; Fullana-Miedes 1985: *formaget, brull*; DEC 1985: *mató, brossat, recuit / quallada, quall*; Pascual 1987: *mató, brull, brossat, recuit / quallada*; Valor 1988: *brull, mató*; Brúixola 1990: *brull, recuit, mató*; Acadèmia 1993: *brullo, formaget.*

12. cast. **tragón, glotón.**[15] Albertí 1961 i Galmar 1977: *golafre, gormand, menjador, menjaire, barrut;* F. Pastor 1967-77: *golafre, menjador, barrut*; F. Pastor 1972: *golafre, menjador;* GAV 1983: *golut, golafre, glotó;* Fullana-Miedes 1985: *golafre, golut, galduf, farto, buc*; DEC 1985: *golafre, menjador, menjaire, fart,*

(6) GAV 1980 no enregistra *chaparrón.*

(7) Fullana-Miedes 1985 i Valor 1988 no enregistren *bocadillo.*

(8) *Cantell* 'cantero, trozo' «cantell de pa» (B. Metge, *DCVB*).

(9) GAV 1980 no enregistra *cucurucho.*

(10) No enregistren *palomitas de maíz* ni F. Pastor 1967, Galmar 1977, Moll 1977, GAV 1980 ni 1983, Fullana-Miedes 1985 ni Valor 1988.

(11) GAV 1980 no enregistra *solomillo.*

(12) *Llistell* 'filete, listel, listón (arqueologia) / filete (de carn), solomillo'; si la primera accepció és correcta, la segona és pura invenció, per ampliació abusiva de significat, car *llistell* en català –i els seus equivalents francès, castellà i italià– només està documentat com a terme d'arquitectura (*llistell* 'adorn de columna', OPou, *DCVB*).

(13) GAV 1980 no enregistra *requesón.*

(14) *Mató* no apareix en Fullana (1921).

(15) GAV 1980 i Miedes 1983 no enregistren *tragón, glotón.*

*fartaner*; Pascual 1987: *glotó, golut, golafre, fartaner*; Valor 1988: *glotó, golafre, fartó;* Brúixola 1990: *menjador, destraler; golafre, golut, fartó, fart;* Acadèmia 1993: *farto, fartana, fartera, engaldidor; galduf, gloto, galdirot, golut, golos, galafre, galima, gulo, buco.*

13. cast. **patata**. Albertí 1961, Moll 1977, Galmar 1977 i DEC 1985: *patata, trumfa;* F. Pastor 1967: *patata, trumfa, creïlla;* GAV 1980 i Acadèmia 1993: *creïlla;* GAV 1983: *creïlla, pataca;* Miedes 1983: *quereguilla, creïlla* [*quereïlla*]; Fullana-Miedes 1985: *quereguilla, creïlla, criadilla*;[16] Pascual 1987: *creïlla, patata, trumfa;* Valor 1988: *patata, creïlla;* Brúixola 1990: *creïlla, patata, pataca, trumfa.*

14. cast. **pimiento (fruto)**.[17] Albertí 1961: *pebrotera, pebroter, bitxo, pebrina / pebrot*; F. Pastor 1967: *pebrera, pebrotera, pebroter, bitxo, pebrina;* Galmar 1977: *pebrotera, pebroter, bitxo, pebrina;* Moll 1977 i Pascual 1987: *pebre, pebrot;* GAV 1983: *pebrotera, pebrot, pimento, pebrereta;* Miedes 1983: *pimiento* <sic>, *peminto* [*pimento*]; Fullana-Miedes 1985: *peminto, pimento, pebrot*; DEC 1985: *pebrot;* Valor 1988: *pebrot*, reg. *pebrera*, reg. *pimentó;* Brúixola 1990: *pebre, pebrera, pebrot*; Acadèmia 1993: *pimento, pebrot / pebrera.*[18]

15. cast. **fresa (fruta)**. Albertí 1961, F. Pastor 1967, Galmar 1977, Valor 1988, DEC 1985 i Brúixola 1990: *maduixa, fraula*; GAV 1980: *fraura* [*fraula*]; GAV 1983: *fraula, fresa;* Miedes 1983: *fraura, fresa*; Fullana-Miedes 1985: *fresa, fraula*; Pascual 1987: *maduixa* [*fraula*]; Acadèmia 1993: *fraula, fraura.*

16. cast. **níspero (fruto)**.[19] Albertí 1961 i Pascual 1987: *nespra, nespla;* Moll 1977: *nespla, nespra /* (fruto dels níspero del Japón) *nispro*;[20] GAV 1983: *nyesple* <sic>, *nyespro, nispro*; Fullana-Miedes 1985: *nyespla, nisparo*; DEC 1985: *nespla, nespra*; Valor 1988: *nespla / nespra*, reg. *nispro;* Brúixola 1990: *nespra, nyespla*; Acadèmia 1993: *nyispro.*

17. cast. **zanahoria**.[21] Albertí 1961, F. Pastor 1967 i Galmar 1977: *pastanaga, safanòria*; GAV 1983: *safanoria, pastanaga*; Miedes 1983 i Acadèmia 1993: *safanòria*; Fullana-Miedes 1985: *pastanaga, safanoria*; Pascual 1987 i Valor 1988: *safanòria, pastanaga*; DEC 1985: *pastanaga*; Brúixola 1990: *carlota, safanòria, pastanaga.*

18. cast. **chimenea**. Albertí 1961: *xemeneia*; F. Pastor 1967: *xemeneia, ximenera;* Albertí 1975: *xemeneia,* (conducte) */ fumeral*; Moll 1977, Valor 1988[22] i Brúixola 1990: *fumeral, xemeneia*; Galmar 1977: *xemeneia, fumera* [*fumeral*]; GAV 1980: *fumeral;* GAV 1983: *fumeral, ximenera;* Miedes 1983 i Fullana-Miedes 1985: *xumenera, ximenera, fumeral*; DEC 1985: *xemeneia, fumeral*; Pascual 1987: *ximenera, xemeneia, fumeral;* Acadèmia 1993: *fumeral, chimenera.*

19. cast. **hollín**.[23] Albertí 1961, F. Pastor 1967-77, Galmar 1977 i DEC 1985: *sutge, sutja, estalzí, estalzim*; F. Pastor 1972: *sutge, sutja, estalzí;* Moll 1977: *sutge, sutja, estalzí, estalzim, follí, teranyina*; GAV 1983: *estalcim, solli* <sic> [*suja, sollim*]; Miedes 1983: *suja*; Fullana-Miedes 1985: *estalcim, suja, sollim, mascarada*; Pascual 1987: *sutja* (o *sutge*), *follí, estalzí* (o *estalzim*); Valor 1988: *sutja*, reg. *estalzim*, reg. *follí*; Brúixola 1990: *sutja, sutge*; Acadèmia 1993: *suja, folli, sollim, estalzim.*

(16) En Fullana (1921) només trobem *criadilla* 'patata' i *pataca* 'batata'.

(17) GAV 1980 no enregistra *pimiento (fruto).*

(18) Acadèmia 1992 enregistra *pebrera* només com a 'guindilla / pimienta' i no enregistra *pebrot.*

(19) Albertí 1975, Galmar 1977 i GAV 1980 no enregistren *níspero (fruto)*; F. Pastor 1967 tampoc, però en l'invers dóna *nespla, nespra*; Miedes 1983 tampoc, però en l'invers enregistra *nyespla.*

(20) Ja el *DFa* enregistra *nispro* 'nom vulg. dels fruits del *nesprer*, vulg. *nisprer*.

(21) GAV 1980 no enregistra *zanahoria.*

(22) Valor dóna a més en l'invers *ximenereta* (en armes antigues) 'fogón'.

(23) GAV 1980 no enregistra *hollín.*

20. cast. **servilleta**. Albertí 1961, F. Pastor 1967, Galmar 1977, Valor 1988, DEC 1985 i Brúixola 1990: *tovalló*; Moll 1977 i Pascual 1987: *tovalló, torcaboques;* GAV 1980: *torcaboques;* GAV 1983 i Miedes 1983: *torcaboca, servilleta*; Fullana-Miedes 1985: *servilleta, torcaboca*; Acadèmia 1993: *torcaboca, tovalloleta.*

21. cast. **lavaplatos, lavavajillas.**[24] Albertí 1961, F. Pastor 1967-77 i Galmar 1977: *rentaplats, llavaplats;* F. Pastor 1972: *llavaplats;* Moll 1977: *escuraplats, rentaplats*; GAV 1980 i Miedes 1983: *escuradora*; DEC 1985: *rentavaixella, màquina de rentar plats*; Pascual 1987: *rentavaixelles, màquina de rentar plats* [*rentaplats*].

22. cast. **bandeja.**[25] Albertí 1961, F. Pastor 1967, Galmar 1977, DEC 1985 i Brúixola 1990: *safata*; GAV 1980 i 1983: *font;* Miedes 1983: *font / batea* <sic>; Recio 1985: *fritera*; Pascual 1987: *safata, plata;* Acadèmia 1993: *fritera, fruitera.*

23. cast. **herramienta**. Albertí 1961 i Galmar 1977: *eina / einam, ferramenta*; F. Pastor 1967: *ferramenta, eina;* GAV 1980: *ferramenta;* GAV 1983: *ahina* 'herramenta <sic>', *ferrament, ferramenta;* Miedes 1983: *ferramenta, aïna*; Fullana-Miedes 1985: *ferrament, ahïna, aïna*; (cast. *herramientas*) *ferramenta, ferrage*; DEC 1985: *eina / utillatge, einam, eines*; Pascual 1987: *eina, ferrament / ferramenta*; Valor 1988: *ferrament, eina / ferramenta*; Brúixola 1990: *eina / ferramenta*; Acadèmia 1993: *ferramenta, ferrament / ferramenta* [*aïna* 'utensilio (de cuina o ferramenta)'].

24. cast. **basura.**[26] Albertí 1961, Galmar 1977 i DEC 1985: *escombraries*; F. Pastor 1967: *fem, porqueria;* Moll 1977: *fems, escombraries*; GAV 1980 i 1983: *fem*; Miedes 1983: *fem, desperdicis*; Pascual 1987: *fem, escombraries;* Valor 1988: *escombraries, brossa, fem;* Brúixola 1990: *fem(s), escombraries*; Acadèmia 1993: *fem, brossa.*

25. cast. **basurero (oficio).**[27] DEC 1985: *escombriaire, escombrariaire.* Tots els altres vocabularis enregistren *femater*, i de més a més: Albertí 1961, Pascual 1987 i Brúixola 1990: *escombriaire*; Moll 1977: *escombriaire, merdacaner*; Galmar 1977: *escombraire*; GAV 1983: *femeter*; Fullana-Miedes 1985: *merder.*

26. cast. **bombilla.**[28] Albertí 1961 i F. Pastor 1967: *bombeta (elèctrica);* Moll 1977: *bombeta, pera*; Galmar 1977, DEC 1985, Pascual 1987, Valor 1988 i Brúixola 1990: *bombeta;* GAV 1980 i 1983 i Miedes 1983: *perilla, pera;* Acadèmia 1993: *pera.*

27. cast. **tomillo.**[29] Albertí 1961, F. Pastor 1967 i Galmar 1977: *farigola, frígola, timó, timonet;* Moll 1977: *timó, farigola, frígola, tem*; GAV 1983: *tomello, timonet;* Miedes 1983: *tomello, timo*; Fullana-Miedes 1985: *timo, timonet, tomello, farigola*; DEC 1985: *farigola, timó, tem, frígola*; Pascual 1987: *timó, farigola, tem, frígola*; Valor 1988: *timó, frígola, farigola;* Brúixola 1990: *timó, farigola;* Acadèmia 1993: *timo, tomello, timonet, frigola.*

28. cast. **romero (planta).**[30] Albertí 1961, F. Pastor 1967, Moll 1977 i Galmar 1977: *romaní*; GAV 1983 i Miedes 1983: *romer*; Fullana-Miedes 1985: *romer, rosmari* <sic>; DEC 1985 i Valor 1988: *romaní, romer*; Pascual 1987 i Brúixola 1990: *romer, romaní*; Acadèmia 1993: *romer, romani.*

(24) Albertí 1975, GAV 1983, Fullana-Miedes 1985, Valor 1988, Brúixola 1990 ni Acadèmia 1993 no enregistren *lavaplatos, lavavajillas.*

(25) Fullana-Miedes 1985 i Valor 1988 no enregistren *bandeja.*

(26) Fullana-Miedes 1985 no enregistra *basura.*

(27) Valor 1988 no enregistra *basurero.*

(28) Fullana-Miedes 1985 no enregistra *bombilla.*

(29) GAV 1980 no enregistra *tomillo.*

(30) GAV 1980 no enregistra *romero.*

29. cast. **junco**.[31] Albertí 1961 i Galmar 1977: *jonc, jonc d'estores, jonquera;* F. Pastor 1967, Albertí 1975, Moll 1977 i DEC 1985: *jonc, jonquera*; GAV 1980 i 1983, Miedes 1983, Fullana-Miedes 1985 i Acadèmia: *junc;* Pascual 1987: *jonc*.

30. cast. **juncia**[32] Albertí 1961, F. Pastor 1967, Galmar 1977 i Pascual 1987: *jonça, castanyola*; Moll 1977: *jonça, junça*; GAV 1983, Miedes, Fullana-Miedes 1985 i Acadèmia: *junça*; DEC 1985: *jónçara*,[33] *serrana*; Valor 1988: *jonça*, [*junça* 'juncia / castañuela'].

31. cast. **hiedra**.[34] Albertí 1961, F. Pastor 1967 i Galmar 1977: *heura, heurera;* GAV 1983 i Fullana-Miedes 1985: *hedra, eura;* Miedes 1983: *hèura, eura*; DEC 1985, Pascual 1987, Valor 1988 i Brúixola 1990: *heura*; Acadèmia 1993: *hedra*.

32. cast. **fresno**.[35] GAV 1983 i Acadèmia 1993: *fleix*; Miedes 1983: *freix*; Fullana-Miedes 1985: *freix, fleix*. La resta de vocabularis: *freixe*.

33. cast. **jazmín**. Albertí 1961, F. Pastor 1967 i Galmar 1977: *llessamí, gessamí;* Albertí 1975 i DEC 1985: *gessamí, llessamí*; Moll 1977: *gesmí, gesminer, gessamí, gessaminer, gesmil, gesmiler, llessamí, llessaminer;* GAV 1980 i Miedes 1983: *gesmil; gesmiler;* GAV 1983: *englatina* <sic>, *gesmil, gessami*; Fullana-Miedes 1985: *gesmil, gesmiler, gessami*; F. Pastor 1985: *gesmil, gessamí, llessamí; gesmiler, gessaminer;* Pascual 1987: *gesmil, gesmiler, gessamí, llessamí;* Valor 1988: (del árabe-persa «jasmín») *gesmí*, reg. *gessamí*, dial. *gesmir, gesmil;* Brúixola 1990: *gesmil, gessamí, llessamí*; Acadèmia 1993: *gesmil*.

34. cast. **raíz**. GAV 1983: *rail, arrail*; Tots els altres vocabularis enregistren *arrel*, i a més: Albertí 1961, F. Pastor 1967, Galmar 1977, DEC 1985, Pascual 1987 i Brúixola 1990: *rel*; Moll 1977: *rel, raïl*; Miedes 1983: *rel, rabassa*; Fullana-Miedes 1985: *arrail*; Valor 1988: reg. *raïl*; Acadèmia 1993: *arraïl, raïl*.

35. cast. **cola (de animal)**. Albertí 1961, F. Pastor 1967, DEC 1985, Pascual 1987 i Valor 1988: *cua*. Albertí 1975 i Galmar 1977: *cua, coa*; Moll 1977, GAV 1980 i Brúixola 1990: *coa, cua*; GAV 1983: *cua, rabo*[36] [*coa*]; Miedes 1983: *coa, cua, rabo;* Fullana-Miedes 1985: *coa, cüa, rabo*; Recio 1985 i Acadèmia 1993: *coa*.

36. cast. **aguijón (de insecto)**.[37] Tots els vocabularis enregistren *fibló*, i a més: Albertí 1961 i Galmar 1977: *agulló, fible*; F. Pastor 1967: *agulló* [*fible*]; Moll 1977: *agulló, fiblet*; GAV 1983: *agujo* <sic>, *fiço*; Fullana-Miedes 1985: *agujo, agullo*; DEC 1985: *fible, agulló*; Pascual 1987: *agulló*; Valor 1988: *fiçó;* Brúixola 1990: *agulló, esperó*; Acadèmia 1993: *agullo, fiço*.

37. cast. **cangrejo**.[38] Albertí 1961, Moll 1977, Galmar 1977, DEC 1985, Pascual 1987 i Valor 1988: *cranc;* F. Pastor 1967 i Brúixola: *cranc, carranc;* GAV 1983 i Fullana-Miedes 1985: *carranc, cranc;* Miedes 1983, Recio 1985 i Acadèmia 1993: *carranc*.

38. cast. **lagarto**.[39] Albertí 1961 i F. Pastor 1967:[40] *llangardaix, lluert;* Moll 1977 i Galmar 1977: *llangardaix, lluert, fardatxo*; GAV 1983: *fardaig;* Miedes 1983: *fardaig, fardacha* 'lagarta'; DEC 1985: *llangardaix*; Pascual 1987, Valor 1988 i Brúixola 1990: *fardatxo, llangardaix;* Acadèmia 1993: *fardacho, sarvacho*.

(31) Brúixola 1990 no enregistra *junco*; Valor 1988 tampoc, però en l'invers dóna *jonc*, i com a dial(ecta l). *junc*.

(32) GAV 1980 i Brúixola no enregistren *juncia*.

(33) *Jónçara* –*DAg*, Masclans (1954), DEC 1982– és grafia contrària a la fonètica històrica catalana –on (vg. Colomina [en premsa]) les *aa* postòniques internes esdevenen *ee*–: la grafia correcta –la que empren *DCVB* i *DCor*– és *jóncera*.

(34) GAV 1980 no enregistra *hiedra*.

(35) GAV 1980 no enregistra *fresno*.

(36) *Rabo* ja apareix en Fullana (1921).

(37) GAV 1980 i Miedes 1983 no enregistren *aguijón*.

(38) GAV 1980 no enregistra *cangrejo*.

(39) GAV 1980 i Fullana-Miedes 1985 no enregistren *lagarto*.

(40) F. Pastor-Giner 1956 i F. Pastor 1985 enregistren també *fardatxo* i *sarvatxo*, i F. Pastor 1960 *fardatxo*.

39. cast. **saltamontes.**[41] Albertí 1961 i Galmar 1977: *llagost, llagosta, saltarel·la, saltamartí*; F. Pastor 1967 i DEC 1985: *llagosta, saltamartí*; Moll 1977: *llagost, saltamartí*; GAV 1983: *saltamarti, llangosta*; Fullana-Miedes 1985: *salto* <sic>; Pascual 1987: *llagostí, llagosta, saltamartí;* Valor 1988: *llagostí, saltamartí;* Brúixola 1990: *llangostí, llangosta, llagosta;* Acadèmia 1993: *saltamarti.*

40. cast. **mariposa (insecto).** Albertí 1961, Galmar 1977 i DEC 1985: *papallona, papalló, voliaina, voliana*; F. Pastor 1967: *papallona, papalló;* Moll 1977: *papalló, papallona, voliana, voliaina, paloma, palometa, pampallola, papiola, babaiana*; GAV 1980: *palometa;* GAV 1983: *papallona, papallo, paloma;* Miedes 1983 i Acadèmia 1993: *paloma, palometa*; Fullana-Miedes 1985: *paloma*; Pascual 1987: *paloma, papallona;* Valor 1988: *papallona*, reg. *paloma;* Brúixola 1990: *palometa, papallona.*

41. cast. **murciélago.**[42] Albertí 1961, Galmar 1977 i DEC 1985: *rata-pinyada, rat-penat, rata-penada, muricec, muriac, voliac*; F. Pastor 1967-77: *rat penat, rata penada, muricec, muriac, voliac;* F. Pastor 1972 i Valor 1988: *rat penat, rata penada;* GAV 1983: *mucilac* <sic>,[43] *rat penat*; Fullana-Miedes 1985: *rat penat, rateta penada, mucilac* <sic>; Pascual 1987: *rata-penada, rat-penat, rata-pinyada, muricec;* Brúixola 1990: *rata-penada, rat-penat*; Acadèmia 1993: *rata penada, rat penat.*

42. cast. **pavo (común).**[44] Tots els vocabularis enregistren *titot*, i a més: Albertí 1961, F. Pastor 1967-77 i Galmar 1977: *gall dindi, indiot, pioc;* F. Pastor 1972: *gall, dindi* <sic>; GAV 1983: *totestiu* <sic>,[45] *tito, pavo*; Miedes 1983: *tita, polit, pavo*; Fullana-Miedes 1985: *pavo, tito*; (pavo <sic> carbonero) *totestiu*; DEC 1985 i Pascual 1987: *gall dindi, indiot;* Valor 1988: *gall dindi*, reg. *titot*, reg. *indià*, reg. *indiot;* Brúixola 1990: *polit, indiot*; Acadèmia 1993: *tito, polit.*

43. cast. **jilguero.**[46] Albertí 1961, F. Pastor 1967-77 i Galmar 1977: *cadernera, cardina, carderola, cadernell;* F. Pastor 1972, Pascual 1987 i Valor 1988: *cadernera, cardina;* GAV 1983 i Acadèmia 1993: *cagarnera, cadarnera;* Miedes 1983: *cagarnera*; Fullana-Miedes 1985: *cadarnera, cagarnera*; DEC 1985: *cadernera, cadernell, cardina, carderola*; Brúixola 1990: *cadernera.*

44. cast. **golondrina (ave).** Albertí 1961: *oreneta, oronella, oronell, orenola*; F. Pastor 1967: *oreneta, oronella*; Moll 1977: *oronella, oreneta, orenola, orandella, vinjolita*; Galmar 1977: *oronella, oreneta, oronell, orenola;* GAV 1980: *oroneta;* GAV 1983: *oroneta* [*oronella, oronell* 'golondrino']; Miedes 1983: *ornella* <sic>, *oroneta, oronell* 'golondrino'; Fullana-Miedes 1985: *ornella* <sic>, *oroneta, fumarell*; DEC 1985: *oreneta, oronell, oronella, orenola*; Pascual 1987: *oroneta, oreneta, oronella;* Valor 1988: *oroneta, oronella;* Brúixola 1990: *oronella, oroneta;* Acadèmia 1993: *oroneta, oronella, orandella.*

45. cast. **abubilla.**[47] Albertí 1961, F. Pastor 1967, Galmar 1977, DEC 1985, i Pascual 1987: *puput, but-but*; Moll 1977, Valor 1988 i Brúixola 1990: *puput*; GAV 1983: *palput, putput*; Fullana-Miedes 1985: *putput, palput*; Acadèmia 1993: *palput.*

(41) GAV 1980 i Miedes 1983 no enregistren *saltamontes.*

(42) GAV 1980 i Miedes 1983 no enregistren *murciélago.*

(43) L'errada ja ve de Fullana (1921): *mucilac* 'mucílago, murciélago'.

(44) GAV 1980 no enregistra *pavo.*

(45) L'errada ve de Fullana (1921) *totestiu* 'pavo [per *paro*] carbonero, monje, ave'.

(46) GAV 1980 no enregistra *jilguero.*

(47) GAV 1980 i Miedes 1983 no enregistren *abubilla.*

46. cast. **jauría.**[48] Albertí 1961, F. Pastor 1967 i Galmar 1977: *canilla, gossada*; Moll 1977: *canilla, gossada, gossam*; DEC 1985 i Pascual 1987: *gossada, canilla;* Valor 1988: *canilla, jama*; Acadèmia 1993: *currica.*[49]

47. cast. **abanico.**[50] Albertí 1961, F. Pastor 1967 i Galmar 1977: *ventall, vano*;[51] Moll 1977, DEC 1985, Pascual 1987 i Brúixola 1990: *ventall*; GAV 1980 i 1983 i Recio 1985: *palmito;* Miedes 1983 i Acadèmia: *palmito, ventall*; Fullana-Miedes 1985: *ventall, aventador, paypay,*[52] *palmito.*

48. cast. **deshilar.**

49. cast. **deshilachar.**[53] Albertí 1961, F. Pastor 1967[54] i Galmar 1977: *esfilegar, esfilegassar, desfilar; esfilagarsar;* Moll 1977: *desfilar; esfilar, desfilagassar, esfilarsegar, esblenar, sargaiar*; GAV 1983: *esfilagarrar* <sic> [*esfilagarsar*]; Miedes 1983: *esfilagarsar; desfilagarchar*; Fullana-Miedes 1985: *desfilar*; *d/esfilagarsar-se*; DEC 1985: *desfilar, esfilegar, esfilegassar; esfilagarsar*; Pascual 1987: *desfilar*; no enregistra *deshilachar* [*esfilagarsar* 'deshilachar']; Brúixola 1990: *desfilar; esfilagarsar;* Acadèmia 1993: *desfilar; desfilarjar.*[55]

50. cast. **calzoncillos.**[56] Albertí 1961, Galmar 1977 i DEC 1985: *calçotets, calçons blancs;* F. Pastor 1967, Miedes 1983, Pascual 1987 i Brúixola 1990: *calçotets;* GAV 1983: *calço, calçotets, saraguells;* Fullana-Miedes 1985: *calço, eslip*; Recio 1985 i Acadèmia 1993: *calço(ns).*

51. cast. **latir (el corazón / el perro).** Albertí 1961, F. Pastor 1967 i Galmar 1977: *bategar, batre, polsar / clapir, glapir*; Moll 1977: *bategar, debategar, esbategar, batre / glapir, glatir*; GAV 1980: *glatir;* GAV 1983: *depir* <sic>, *glapir* [*clapir*]; Miedes 1983: *clapir, glapir, glatir*; Fullana-Miedes 1985: *clapir, bategar, glapir, batre*; DEC 1985: *bategar, batre / clapir, glapir*; Pascual 1987: *bategar, batre* [*glapir* el gos 'ladrar']; Valor 1988: *bategar / clapir*; Brúixola 1990: *bategar, batre;* Acadèmia 1993: *glatir / clapir.*

52. cast. **estómago.** F. Pastor-J. Giner 1956 i Valor 1987: *estómec*;[57] Albertí 1961, Galmar 1977, DEC 1985 i Pascual 1987: *estómac, païdor, pedreny, ventrell;* F. Pastor 1967: *estómac, païdor, ventrell;* GAV 1980: *estomac, ventre, pancha;* GAV 1983: *ventrell, estomac;* Miedes 1983: *estomac, ventrell*; Fullana-Miedes 1985: *estomac*; Recio 1985 i Acadèmia 1993: *estomec*; Brúixola 1990: *estómac.*

53. cast. **cogote; pescuezo.** Albertí 1961 i Galmar 1977: *clatell, clotell; bescoll, batcoll;* F. Pastor 1967: *clatell, clotell, bascoll*;[58] *bescoll, batcoll*; Moll 1977: *clatell, clotell, bescoll; clatell, clotell, betcoll, bescoll*; GAV 1980 i Valor 1988: *bascoll;* GAV 1983: *clatell* [*clotell, bascoll*]; Miedes 1983: *clatell, bascòll*; Fullana-Miedes 1985: *bascoll, clatell, clotell*; DEC 1985: *clatell, bescoll, nuca*; *clatell, bescoll, batcoll*; Pascual 1987: *bescoll, clatell, nuca;* Brúixola 1990: *bescoll, clatell;* Acadèmia 1993: *tos; bascoll.*

54. cast. **cosquillas.**[59] Albertí 1961, F. Pastor 1967, Galmar 1977 i DEC 1985: *pessigolles*; Moll 1977: *pessigolles, cossigolles, cossiguetes, còssigues, joques*; GAV 1983: *cosconelles, cosgolles* <sic>, *cosquerelles, coscarrelles, sigolletes, pessigolles,*

(48) GAV 1980 i 1983, Miedes 1983, Fullana-Miedes 1985 i Brúixola 1990 no enregistren *jauría*.

(49) Acadèmia 1992 proposa llegir *currica*, però és a tort, car la pronúncia correcta és *cárrica* (Escrig, 1851; València, Casinos, *DCVB*).

(50) Valor 1988 no enregistra *abanico* però en l'invers dóna *ventallar-se* i *ventall*.

(51) *Vano* –manlleu no valencià del cast. *abano*, forma anterior del modern *abanico*– no figura al *DFa* ni al DEC 1982, ni tampoc en F. Pastor-Giner (1956), ni en F. Pastor (1960, 1985).

(52) *Paypay* no figura a Fullana (1921). És un tipus de ventall de cartó i d'estil japonès; a Alcoi és viva l'expressió *fer-se aire amb un paipai*.

(53) GAV 1980 i Valor 1990 no enregistren ni *deshilar* ni *deshilachar*.

(54) F. Pastor 1967 canvia amb encert la grafia a tort amb *a* de Fabra i Albertí *esfilagarsar* pel més correcte *esfilegarsar*. Diguem de passada que *esfilegarsar* és una simple variant ultracorrecta –amb *rs* per *ss*, com en *nevarsar* per *nevassar*– de *esfilegassar*.

(55) *Desfilarjar* és grafia ultracorrecta de *desfilarxar*, explicable per la procedència apitxada dels redactors del vocabulari de l'Acadèmia.

(56) GAV 1980 no enregistra *calzoncillos*. Valor 1988 tampoc no enregistra *calzoncillos* però en l'invers dóna *calçó* 'm. indum. Vd polaina / pl. calzones / – de cuiro (per a caça i camp) zahones o zafones'.

(57) Vegeu documentació antiga i moderna de *estómec* en Colomina (en premsa).

(58) Vegeu en Colomina (1993: 32-33) per què *bascoll* és una mala grafia de *bescoll*.

(59) GAV 1980 no enregistra *cosquillas*.

*cuscanelles* [*cosigolles*]; Miedes 1983: *cosconelles, pessigolles*; Fullana-Miedes 1985: *cosconelles, pessigolles, cosquerelles*; Pascual 1987 i Brúixola 1990: *cosquerelles, pessigolles*; Valor 1988: *pessigolles, cuscanelles, cosquerelles*; Acadèmia 1993: *cosquerelles, picoretes, sorelles* [*coscanelles, cosquelles, cosquinogues, cassonies, cassonigues*].

55. cast. **agujetas (en los músculos doloridos)**.[60] Albertí 1961 i Galmar 1977: *punxades que hom nota als músculs adolorits a causa d'un exercici violent*; Moll 1977 i Valor 1988: *esbraonament*;[61] Miedes 1983: *punchades*; DEC 1985: *adoloriment, cruiximent, fiblades*; Pascual 1987: *adoloriment, cruiximent*; Brúixola 1990: *dolor muscular causat per l'exercici físic*.

56. cast. **bálsamo**.[62] F. Pastor-J. Giner 1956 i Valor 1988: *bàlsem*;[63] F. Pastor 1967, Galmar 1977, DEC 1985, Pascual 1987 i Brúixola 1990: *bàlsam;* Miedes 1983: *balsam, balsem*; GAV 1983, Fullana-Miedes 1985 i Acadèmia 1993: *balsem*.

57. cast. **muñeca (del brazo)**. Albertí 1961, Galmar 1977: *puny, monyó, canell*; F. Pastor 1967: *monyó, canell;* Moll 1977: *canell, monyeca, braó*; GAV 1980: *gobanilla;* GAV 1983: *munyiqueta, gobanilla, munyeca, monyica;* Miedes 1983: *gobanilla, monyica;* Fullana-Miedes 1985: *munyeca, gobanilla, monyica/munyica*; DEC 1985: *puny, canell, monyó*; Pascual 1987: *monyica, puny, canell;* Valor 1988: *canell*, reg. *monyica;* Brúixola 1990: *canell, puny, monyó*; Acadèmia 1993: *gobanella, monyica*.

58. cast. **párpado**.[64] Albertí 1961, F. Pastor 1967-77 i Galmar 1977: *parpella, palpebra, pipella*;[65] F. Pastor 1972: *parpella, palpebra*; Moll 1977, DEC 1985 i Pascual 1987: *parpella, pipella, palpebra;* GAV 1983: *parpalla, parpall*; Miedes 1983 i Fullana-Miedes 1985: *parpall*; Valor 1988 i Brúixola 1990: *parpella*; Acadèmia 1993: *parpall, parpalla*.

59. cast. **pequeño**.[66] Albertí 1961: *petit*; F. Pastor 1967, Albertí 1975, Galmar 1977 i DEC 1985: *petit, menut*; Moll 1977: *petit, menut, xic*; GAV 1980: *chicotét, menút, chic;* GAV 1983: *petorri*; Miedes 1983: *menut, chicotet;* Pascual 1987: *menut, petit;* Valor 1988: *petit, menut*, fam. *xicotet;* Brúixola 1990: *xicotet, menut, petit*; Acadèmia 1993: *chicotet, menut, petorri*.

60. cast. **pequeñito**.[67] Moll 1977: (cast. *pequeñuelo*) *petitó, menudó, xicotet, xicarró, esquiterell, fiteu, fitó, graneu, mató, tetó*; GAV 1980: *chicotíu, chicotét, menudíu;* GAV 1983: *chicotet, chicotiu*; Miedes 1983: *chicotiu;* Fullana-Miedes 1985: *chicotet, petorri*; DEC 1985: *petitet, petitó, menudet*; Acadèmia 1993: *chicotiu, chiconiu, chicorriniu*.

61. cast. **larguirucho**.[68] Albertí 1961, Galmar 1977, DEC 1985, Pascual 1987 i Brúixola 1990: *llargarut*; F. Pastor 1967, GAV 1983, Valor 1988 i Acadèmia 1993: *llarguerut*;[69] Fullana-Miedes 1985: *llarguerut, desllanguit*.

62. cast. **defender**. Albertí 1961, F. Pastor 1967, Moll 1977, Galmar 1977, DEC 1985 i Brúixola 1990: *defensar, defendre*; GAV 1980 i 1983 i Acadèmia: *defendre;*

(60) F. Pastor 1967, GAV 1980 i 1983, Fullana-Miedes 1985 i Acadèmia 1993 no enregistren *agujetas*.

(61) Al *DFa* trobem *esbraonament* 'estat del qui està esbraonat' i *esbraonar* 'fer perdre el braó, el vigor, fatigar en alt grau'; DCVB *esbraonar* 'adolorir dels <sic> braons i extremitats del cos per excés de treball o per un exercici massa violent', refl. 'cansar-se de les extremitats ... cast. *extenuarse*' (mall., men.).

(62) GAV 1980 no enregistra *bálsamo*.

(63) Vegeu documentació antiga i moderna de *bàlsem* en Colomina (en premsa).

(64) GAV 1980 no enregistra *párpado*.

(65) *Pipella* és deformació balear de *parpella* (*DCVB*).

(66) Fullana-Miedes 1985 no enregistra *pequeño*.

(67) Albertí 1961, F. Pastor 1967, Galmar 1977, Pascual 1987, Valor 1988 i Brúixola 1990 no enregistren *pequeñito*.

(68) GAV 1980 i Miedes no enregistren *larguirucho*.

(69) Vegeu per què *llarguerut* és la grafia correcta en Colomina (1992: 135-6).

Miedes 1983: *defendre, amparar;* Fullana-Miedes 1985 i Pascual 1987: *defendre, defensar*; Valor 1988: *defensar.*

63. cast. **resbalar**. Albertí 1961: *relliscar, lliscar, llenegar, resquitllar;* F. Pastor 1967: *relliscar, lliscar;* Moll 1977: *llenegar, lliscar, relliscar, esvarar, rossolar, solsir, resquitllar, esllisar, patinar*; Galmar 1977: *lliscar, relliscar, rossolar, llenegar, resquitllar*; GAV 1980: *esvarar;* GAV 1983: *esbarar, esvarallar, esvarar;* Miedes 1983: *esvarar, esguillar, esvarallar*; Fullana-Miedes 1985: *esvarallar, esvarar, esbarar*; DEC 1985: *relliscar, lliscar, llenegar, rossolar*; Pascual 1987: *esvarar* (o *esvarar-se), relliscar, llenegar;* Valor 1988: *relliscar, esvarar-se;* Brúixola 1990: *esvarar, esquitllar-se, relliscar;* Acadèmia 1993: *esvarar, esgollar, esguillar-se, eslliçar.*

64. cast. **salpicar**. Albertí 1961 i Galmar 1977: *esquitxar, esquitxotejar, escatxicar, brufar;* F. Pastor 1967: *esquitxar;* Moll 1977: *esquitar, esguitar, esquitxar, escarritxar, brufar, esbrufar, xarbotar*; GAV 1980 i Acadèmia 1993: *esguitar;* GAV 1983: *salpicar, esgarrifar, esguitar, esborronar*;[70] Miedes 1983: *esguitar, salpicar*; Fullana-Miedes 1985: *salpicar, arruixar, asperjar, esborronar, esgarrifar, esguitar*; DEC 1985: *esquitxar, escatxicar*; Pascual 1987: *esguitar, esquitxar, esquitar*; Valor 1988: *esquitar o esquitxar;* Brúixola 1990: *esguitar, esquitxar, estufar.*

65. cast. **empujar**; 66. cast. **empujón**. Albertí 1961 i Galmar 1977: *empènyer; empenta, espenta;* F. Pastor 1967: *empényer; empenta, espenta*; Moll 1977: *empènyer, empentejar, sempentejar, espitjar*; *empenta, sempenta, forfollada*; GAV 1980: *espentar; espenta;* GAV 1983: *espejar* <sic>, *espentar, espentejar, empenyer; espenta, empenyiment;* Miedes 1983: *espentar, empenyer; espenta*; Fullana-Miedes 1985: *empenyer, espenyer, espentejar; espenta, empenyiment, espentejada*; DEC 1985: (mover) *empènyer* / (atropellar) *empentar*; *empenta*; Pascual 1987: *empényer* (o *empènyer*) / (dar empujones) *empentar, empentejar, espentar, espentejar*; *empenta, espenta*; Valor 1988: *empényer, espentar, espitjar*, no enregistra *empujón* [*empentar, espentejar, espenta* 'empujón']; Brúixola 1990: *espentar, empentar, espényer, empényer; espenta, empenta;* Acadèmia 1993: *espentar, espentejar; espenta, espento.*

67. cast. **hinchar**. Albertí 1961, F. Pastor 1967, Moll 1977, Galmar 1977 i DEC 1985: *inflar*; GAV 1980 i Recio 1985: *unflar*; GAV 1983: *reinflar, unflar, inflar, abotinflar-se;* Miedes 1983: *unflar, bufar;* Fullana-Miedes 1985: *reinflar, unflar, abotinflar-se*; Pascual 1987 i Valor 1988: *inflar, unflar;* Brúixola 1990: *inflar, unflar, embotir;* Acadèmia 1993: *unflar, botinflar.*

68. cast. **en cuclillas (ponerse)**.[71] Albertí 1961, Galmar 1977 i DEC 1985: *a la gatzoneta*; Pascual 1987: *emponnat* (o *aponat), a la gatzoneta;* Valor 1988: *a la gatzoneta, emponnat;* Acadèmia 1993: *ajocar* [*aponar-se, emponar-se*].

69. cast. **juguete**. Albertí 1961, F. Pastor 1967 i DEC 1985: *joguina;* Albertí 1975 i Galmar 1977: *joguina, jugueta;* Moll 1977: *jugueta, joguina, juguet*; GAV 1980 i 1983, Miedes 1983 i Acadèmia 1993: *joguet*; Fullana-Miedes 1985: *joguet, trasto*

(70) *Esborronar* 'salpicar' (Fullana 1921) no apareix a Escrig, ni Llombart, ni Martí i Gadea; *esgarrifar* 'salpicar' (Fullana 1921) ja es troba en Escrig.

(71) F. Pastor 1967 no enregistra *cuclillas (en)* però l'any 1960 arreplegava *aponar-se, aponat*; GAV 1980 i Miedes 1983 no l'enregistren; GAV 1983 i Fullana-Miedes 1985 no l'enregistren, però en l'invers donen *emponar-se*; Brúixola 1990 tampoc l'enregistra però a l'invers dóna *a la gatzoneta, emponnar-se i emponnat.*

<sic>, *joquina* <sic>; Pascual 1987 i Brúixola 1990: *joguet, joguina*; Valor 1988: *joguina, joguet.*

70. cast. **llamarada**.[72] Albertí 1961 i Galmar 1977: *flamarada, foguerada, xera, brimarada;* F. Pastor 1967: *flamerada,*[73] *foguerada*; GAV 1980 i Miedes 1983: *flamerada;* GAV 1983: *flamera*; Fullana-Miedes 1985: *flamerada, flama*; DEC 1985, Pascual 1987 i Brúixola 1990: *flamarada*; Acadèmia 1993: *flamerada / llampada.*

## Particularisme *vs.* unitarisme en els vocabularis valencians

Per veure en quina mesura els vocabularis bilingües valencians actuals són, pel que fa a la tria lèxica, unitaristes –«catalanitzants»– o particularistes –«secessionistes»–, hem triat quinze entrades lèxiques en les quals és possible de trobar quatre solucions: *a*) únicament el lexema considerat característic de Catalunya –que en una part dels casos (p.e. els números 6, 35, 37, 59, 62, 64 i 67) és també valencià–, *b*) el lexema característic de Catalunya i després el del País Valencià; *c*) el lexema característic valencià i després el català; i *d*) únicament el lexema considerat valencià. Els resultats apareixen en el quadre 1.

(72) Valor 1988 no enregistra *llamarada*.

(73) Vegeu sobre la grafia correcta *flamerada* Colomina (1992: 136-141).

## Quadre 1

| CATALÀ | CATALÀ-VALENCIÀ | VALENCIÀ-CATALÀ | VALENCIÀ |
|---|---|---|---|
| 6. *bacallà*<br>F. Pastor, Galmar, Valor | *bacallà-abadejo*<br>Pascual, Acadèmia | *abadejo-bacallà*<br>Fullana-M., Brúixola | *abadejo*<br>GAV 80 i 83, Miedes |
| 13. *patata*<br>Galmar | *patata-creïlla*<br>F. Pastor, Valor | *creïlla-patata*<br>Pascual, Brúixola | *creïlla*<br>GAV 80 i 83, Acadèmia, Miedes, Fullana-M. |
| 27. *farigola* | *farigola-timó*<br>F. Pastor, Galmar | *timó-farigola*<br>Fullana-M., Pascual, Valor, Brúixola | *timo*<br>GAV 83, Miedes, Acadèmia |
| 28. *romaní*<br>P. Pastor, Galmar | *romaní-romer*<br>Valor | *romer-romaní*<br>Pascual, Brúixola, Acadèmia | *romer*<br>GAV 83, Miedes |
| 35. *cua*<br>F. Pastor, GAV 83, Pascual, Valor | *cua-coa*<br>Galmar | *coa-cua*<br>GAV 80, Brúixola, Miedes, Fullana-M. | *coa*<br>Acadèmia |
| 37. *cranc*<br>Galmar, Pascual, Valor | *cranc-carranc*<br>F. Pastor, Brúixola | *carranc-cranc*<br>GAV 83, Fullana-M. | *carranc*<br>Miedes, Acadèmia |
| 38. *llangardaix*<br>F. Pastor | *llangardaix-fardatxo*<br>Galmar | *fardatxo-llangardaix*<br>Pascual, Valor, Brúixola | *fardatxo*<br>GAV 83, Miedes, Acadèmia |
| 40. *papallona*<br>F. Pastor, Galmar | *papallona-palom(et)a*<br>GAV 83, Valor | *palom(et)a-papallona*<br>Pascual, Brúixola | *palom(et)a*<br>GAV 80, Miedes, Fullana-M. Acadèmia |
| 54. *pessigolles*<br>F. Pastor, Galmar | *pessigolles-cosquerelles*<br>Valor | *cosquerelles-pessigolles*<br>GAV 83, Miedes, Fullana-M., Pascual, Brúixola | *cosquerelles*<br>Acadèmia |
| 59. *petit*<br>F. Pastor, Galmar, Pascual | *petit-xicotet*<br>Valor | *xicotet-petit*<br>Brúixola | *xicotet*<br>GAV 80, Miedes, Acadèmia |
| 62. *defensar*<br>Valor | *defensar-defendre*<br>F. Pastor, Galmar, Brúixola | *defendre-defensar*<br>Fullana-M., Pascual | *defendre*<br>GAV 80 i 83, Miedes, Acadèmia |
| 63. *relliscar*<br>F. Pastor, Galmar | *relliscar-esvarar*<br>Valor | *esvarar-relliscar*<br>Pascual, Brúixola | *esvarar*<br>GAV 80 i 83, Miedes, Fullana-M., Acadèmia |
| 64. *esquitxar*<br>F. Pastor, Galmar | *esquitxar-esguitar* | *esguitar-esquitxar*<br>Pascual, Valor, Brúixola | *esguitar*<br>GAV 80 i 83, Miedes, Fullana-M., Acadèmia |
| 67. *inflar*<br>F. Pastor, Galmar | *inflar-unflar*<br>Pascual, Valor, Brúixola | *unflar-inflar*<br>GAV 83 | *inflar*<br>GAV 80, Miedes, Fullana-M., Acadèmia |
| 68. *a la gatzoneta*<br>Galmar | *gatzoneta-emponat*<br>Valor | *emponat-gatzoneta*<br>Pascual | *emponat* |

Les dades del quadre 1 distribuïdes segons els vocabularis apareixen en el quadre 2.

Quadre 2

| | CATALÀ | CATALÀ-VALENCIÀ | VALENCIÀ-CATALÀ | VALENCIÀ |
|---|---|---|---|---|
| F. Pastor 1967 | 10 | 4 (*creïlla, timó, carranc, defendre*) | 0 | 0 |
| Galmar 1977 | 11 | 3 (*timó, coa, fardatxo*) | 0 | 0 |
| GAV 1980 | 0 | 0 | 1 (*cua*) | 8 |
| GAV 1983 | 1 (*cua*) | 1 (*papallona*) | 3 (*cranc, pessigolles, inflar*) | 8 |
| Miedes 1983 | 0 | 0 | 2 (*cua, pessigolles*) | 12 |
| Fullana-Miedes 1985 | 0 | 0 | 6 | 5 |
| Pascual 1987 | 3 (*cua, cranc, petit*) | 2 (*bacallà, inflar*) | 10 | 0 |
| Valor 1988 | 4 (*bacallà, cua, cranc, defensar*) | 8 | 3 | 0 |
| Brúixola 1990 | 0 | 3 (*cranc, defensar, inflar*) | 11 | 0 |
| Acadèmia 1993 | 0 | 1 (*bacallà*) | 1 (*romanó*) | 12 |

Les dades del quadre 2 ens permeten distribuir els vocabularis valencians en quatre grups:

*a*) *catalanitzants*: El més radical és Galmar, que en 11 dels 15 mots es decanta per la sola forma catalana i només en tres casos admet en segon lloc la forma valenciana (*timó, coa, fardatxo*). Molt semblant és el criteri de F. Ferrer Pastor que en 10 dels 15 casos només dóna la forma catalana i solament accepta en segon lloc la solució valenciana en quatre casos (*creïlla, timó, carranc, defendre*). Molt lluny de Galmar i F. Pastor trobem Valor (amb quatre casos d'única solució catalana: *bacallà, cua, cranc, defensar*), Pascual (amb tres: *cua, cranc, petit*) i GAV 1983 (amb un: *cua*);

*b*) *equilibrats però amb preferència per la forma catalana*: el qui es destaca en aquesta solució és de bon tros Valor, amb vuit casos (sobre els quinze possibles), seguit ja a bona cosa de distància per F. Pastor (quatre casos), Galmar i Brúixola (tres), Pascual (dos) i GAV 1983 i Acadèmia (un: *papallona; bacallà*);

*c*) *equilibrats però amb preferència per la forma valenciana*: Brúixola (amb 11 casos) i Pascual (amb 10) són els campions d'aquesta solució, seguits a mitjana distància per Fullana-Miedes (sis casos) i ja molt més lluny per GAV 1983 i Valor (amb tres), Miedes (amb dos) i GAV 1980 i Acadèmia (amb un); i

*d*) *secessionistes*: no solament cerquen diferenciar-se de Catalunya en el lèxic, sinó que també s'aparten de la tradició gràfica –no accentuen, redueixen els dígrafs tradicionals *tg, tj, tl* a *g, j, l*, rebutgen les grafies *x* inicial i postconsonàntica i *tx*–. Formen un grup compacte encapçalat per Miedes i Acadèmia (amb 12 casos de només la forma valenciana), seguit pels GAVs (vuit casos) i una mica més lluny per Miedes-Fullana (cinc casos).

## Quadre 3

| | F.P. | GAL. | GAVb | GAV | Mie. | F-M | Pas. | Val. | Brúi. | Aca. | total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3. estrela | | | + | + | + | + | + | + | + | + | 8 |
| 6. abadejo | | | + | + | + | + | + | + | + | + | 8 |
| 10. rellomello | | | | + | | + | | | | + | 3 |
| 13. creïlla | + | | + | + | + | + | + | + | + | + | 9 |
| 14. pimentó | | | | + | + | + | | + | | + | 5 |
| 19. follí | | | | | | | + | + | | + | 3 |
| 27. tomello | | | | + | + | + | | | | + | 4 |
| 28. romer | | | | + | + | + | + | + | + | + | 7 |
| 29. junc | | | + | + | + | + | | + | | + | 6 |
| 31. hedra | | | | + | | + | | | | + | 3 |
| 33. gesmil | | | + | + | + | + | + | + | + | + | 8 |
| 34. (ar)raïl | | | | + | | + | | + | | + | 4 |
| 36. fiçó | | | | + | | | | + | | + | 3 |
| 40. palom(et)a | | | + | + | + | + | + | + | + | + | 8 |
| 44. oroneta | | | + | + | + | + | + | + | + | + | 8 |
| 47. palmito | | | + | + | + | + | | | | + | 5 |
| 52. estómec | | | | | | | | + | | + | 2 |
| 54. cosquerelles | | | | + | | + | + | + | + | + | 6 |
| 56. bàlsem | | | | + | + | + | | + | | + | 5 |
| 58. parpall(a) | | | | + | + | + | | | | + | 4 |
| 59. xicotet | | | + | + | + | + | + | + | | + | 7 |
| 60. llarguerut | + | | | + | | + | | + | | + | 5 |
| 63. esvarar(-se) | | | + | + | + | + | + | + | + | + | 8 |
| 64. esguitar | | | + | + | + | + | + | | + | + | 7 |
| 65. espentar | | | + | + | + | | + | + | + | + | 7 |
| 67. unflar | | | + | + | + | + | + | + | + | + | 8 |
| 68. empon(n)ar-se | | | | + | | + | + | + | + | + | 6 |
| 69. joguet | | | + | + | + | + | + | + | + | + | 8 |
| 70. flamerada | + | | + | | + | + | | | | + | 5 |
| total | 3 | 0 | 15 | 25 | 20 | 25 | 22 | 15 | 15 | 29 | |

Lògicament, els vocabularis «catalanitzants» són els que obliden més formes valencianes: F. Pastor (sis omissions: no enregistra [n.e.] *estrela*, n.e. *abadejo*; n.e. *junc*; n.e. *palmito*, n.e. *coa*; n.e. *unflar*; i igualment Galmar 1977 (6 omissions: n.e. *estrela*, n.e. *abadejo*, n.e. *creïlla*, n.e. *junc*; n.e. *palmito*; n.e. *unflar*. La resta de vocabularis ometen dos o tres mots valencians, però mentre els «equilibrats» obliden els més característicament valencians (*junc* [Pascual], *palmito* [Pascual, Valor i Brúixola], *coa* [Pascual, Valor], *aladre* [Brúixola], els «secessionistes» exclouen els més característicament catalans (*bacallà* [GAV 1980 i 1983, Miedes], *ventall* [GAV 1980 i 1983], *inflar*[74] [GAV 1980, Miedes, Acadèmia], *cua*[75] [Acadèmia].

## Fidelitat a la normativa lèxica i fonètico-gràfica

Sobre una selecció de 29 casos possibles, el quadre 3 ens permetrà veure comparativament les innovacions valencianes al lèxic normatiu (*DFa*); marquem amb el signe + els vocabularis on es troba el mot assenyalat.

Els mots valencians no enregistrats al *DFa* admesos per més vocabularis valencians són per orde decreixent els següents. En 9 vocabularis: *creïlla*; en 8 vocabularis: *estrela, abadejo, gesmil, palom(et)a, oroneta, esvarar(-se), unflar, joguet*; en 7 vocabularis: *romer, xicotet, esguitar, espentar;* en 6 vocabularis: *junc, cosquerelles, empon(n)ar-se*; en 5 vocabularis: *pimentó, palmito;* en 4 vocabularis: *tomello, (ar)raïl, parpall(a)*; en 3 vocabularis: *rellomello, follí, hedra, fiçó.* Pel que fa a les variants fonètico-gràfiques *bàlsem, llarguerut* i *flamerada* apareixen en cinc vocabularis, mentre que *estómec* només en dos. D'una manera paral·lela –i previsible– al que hem vist en l'apartat anterior, els vocabularis «catalanitzants» introdueixen poques innovacions al lèxic normatiu –fabrià–, mentre que els «secessionistes» es caracteritzen per la plètora innovadora. Sobre 29 possibles introduccions, els vocabularis valencians es comporten de la manera següent. En primer lloc destaquen els «secessionistes»: Acadèmia 1993 (amb 29 innovacions), GAV 1983[76] (amb 26) i Fullana-Miedes 1985 (amb 25); després trobem Valor 1988 (amb 22 innovacions), Miedes 1983 (amb 20), GAV 1980, Brúixola 1990 i Pascual 1987 (amb 15), F. Pastor 1977 (amb 3) i Galmar 1977 (sense innovacions).

Si considerem els setanta mots, la classificació és, si fa no fa, la mateixa: Acadèmia (amb 83 innovacions), GAV 1984 (amb 48), Miedes 1983 (amb 37), Valor 1988 (amb 29), Pascual 1987 (amb 21), Brúixola 1990 (amb 19), GAV 1980 (amb 15), F. Pastor 1967 (amb 6), Galmar 1977 (amb 1).

Acadèmia 1993 (83 innovacions al *DFa*; marquem amb + els mots no enregistrats al *DCVB*): *gayola, baso*; *+estrelot*;[77] *purnejar, plovitejar, borrimar*,[78] *gotinyar*;[79] *arruixada, arruixo, aiguarrada, aiguaço, sacso d'aigua, alaboc, garrotada*,[80] *+diluviada, chapallo; rosa, esclafito, clocha* 'palomita de maíz'; *brullo, formaget*

(74) Potser en la decisió de l'Acadèmia d'excloure *inflar* han pesat les dades incompletes del *DCVB*, que dóna *unflar* com a propi de les tres capitals provincials (Cast., Val., Al.), i *inflar* com a no valencià (or., occ., bal.). Com a parlant alcoià puc assegurar que els valencians de la regió d'Alcoi només usen *inflar*, i senten *unflar* com a forma vulgar i forastera.

(75) Potser també en aquest cas les dades incompletes del *DCVB*, que dóna *cua* només de València i Alcoi, i *coa* de Vinaròs, Benassal, Alcalà de X., Elx, Val., Gandia, Morella i el Forcall, han influït en la decisió de l'Acadèmia d'excloure *cua*.

(76) Dels quals 21 ja apareixen al vocabulari de Fullana (1921); només hi manquen: *creïlla, estómec, fiçó* i *follí*.

(77) *Estrelot* 'lucero, lucero del alba' i el modisme *eixir-li a u l'estrelot* 'tardar mucho en hacer algo'.

(78) *Borrimar* (Castelló, *DCVB*).

(79) *Gotinyar* (Alcoi, *DCVB*).

(80) *Garrotada* (Guardamar, *DCVB*).

'requesón'; *rellomello*; *farto*, *+fartana* (n.e. Acadèmia 1992), *fartera*, *engaldidor*; *galduf*, *galdirot*, *galafre*, *galima*, *gulo*,[81] *+buco* (llegiu *bucó*); *creïlla*; *pebrera*; *nyispro*; *chimenera*; *folli*, *sollim*; *fritera*; *aïna*; *tomello*; *romer*; *junça*; *hedra*; *fleix*; *gesmil*; *arraïl*, *raïl*; *sarvacho*; *tito*; *oroneta*, *orandella*; *palput*; *currica* 'jauría'; *desfilarjar* <sic>; *estomec*; *bascoll*; *cosquerelles*, *picoretes*, *sorelles* [*coscanelles*, *cosquelles*, *cosquinogues*, *cassonies*, *cassonigues*]; *balsem*; *gobanella*; *parpall*, *parpalla*; *chicotet*, *petorri*; *llarguerut*; *esvarar*, *esgollar*, *esguillar-se*,[82] *eslliçar*;[83] *esguitar*; *espentar*, *espento*; *unflar*; *emponar-se*, *aponar-se*; *joguet*.

Fullana-Miedes (54 innovacions al *DFa*): *colmena*; *estrela*; *pluinar*; *chapallo*, *arruixo*; *abaejo*, *abadeig*; *rellomello*; *formaget*; *galduf*, *farto*; *quereguilla*, *creïlla*, *criadilla*; *peminto*, *pimento*; *nyespla*; *xumenera*, *ximenera*; *sollim*; *aïna*, *ferrage*; *tomello*; *junc*; *junça*; *hedra*; *freix*, *fleix*; *gesmil*, *gesmiler*; *arrail*; *carranc*; *tito*; *palput*; *paypay*, *palmito*; *bascoll*; *cosconelles*, *cosquerelles*; *balsem*; *gobanilla*, *monyica/munyica*; *parpall*; *chicotet*, *petorri*; *llarguerut*; *esvarallar*, *esvarar*; *esguitar*; *espenyer*, *espentejada*; *unflar*, *abotinflar-se*; *joguet*; *flamerada*.

GAV 1983 (48 innovacions al *DFa*): *aiguafort*, *aiguaço*; *formaget* 'requesón'; *rellomello*, *sanguisolet(a)* '*solomillo*'; *creïlla*, *pataca*; *pimento*, *pebrereta*; *nyesple* <sic>, *nyespro*; *ximenera*; *solli(m)*; *ahina*; *perilla*; *tomello*; *romer*; *junça*; *hedra*; *fleix*; *gesmil*; *rail*, *arrail*; *tito*; *oroneta*; *palput*, *putput*; *bascoll*; *cosconelles*, *cos(i)golles*, *cosquerelles*, *coscarrelles*, *sigolletes*, *cuscanelles* [*cosigolles*]; *balsem*; *gobanilla*; *parpalla*, *parpall*; *petorri*; *llarguerut*; *esvarallar*, *esvarar*; *esguitar*; *espentar*; *unflar*; *emponar-se*; *joguet*.

Miedes 1983 (37 innovacions al *DFa*): *purnejar*; *arruixo*; *borles*, *roses* 'palomitas de maíz'; *brullo* 'requesón'; *llomello*; *creïlla*; *peminto*, *pimento*; *fraura*; *xumenera*, *ximenera*; *escuradora*; *aïna*; *perilla*; *tomello*; *romer*; *junça*; *freix*; *gesmil*, *gesmiler*; *tita*; *oroneta*; *desfilagarchar*; *bascoll*; *cosconelles*; *balsem*; *gobanilla*; *parpall*; *chicotet*; *esvarar*, *esguillar*, *esvarallar*; *esguitar*; *espentar*; *unflar*; *joguet*.

Valor 1988 (29 innovacions al *DFa*): *plovitejar*,[84] *arruixada*; *llomell* o *llomello*; *fartó*; *creïlla*; reg. *pebrera*, reg. *pimentó*; *follí*; *romer*; *junça*; *gesmí*, *gesmil*; reg. *raïl*; *oroneta*; *jama* 'jauría'; *estómec*; *bascoll*; *cuscanelles*, *cosquerelles*; *bàlsem*; fam. *xicotet*; *llarguerut*; *esvarar-se*; *esquitar*; *espentar*; *unflar*; *emponnat*; *joguet*.

Pascual 1987 (21 innovació al *DFa*): *arruixada*; *rosa*, *roseta* 'palomita de maíz'; *unflar*; *creïlla*; *ximenera*; *follí*; *romer*; *gesmil*, *gesmiler*; *oroneta*; *cosquerelles*; *esvarar* (o *esvarar-se*); *esguitar*, *esquitar*; *espentar*; *unflar*; *emponnat* (o *aponat*); *joguet*.

Brúixola 1990 (19 innovacions al *DFa*): *borla* 'palomita de maíz'; *fartó*; *creïlla*; *pebrera*; *pataca*; *nyespla*; *carlota*; *romer*; *gesmil*; *oroneta*; *cosquerelles*; *xicotet*; *esvarar*; *esguitar*; *espentar*, *espényer*; *emponnar-se*, *emponnat*; *joguet*.

(81) *Guló* 'golut' (Guardamar, *DCVB*).

(82) *Esguillar-se* (pir. or., eiv., *DCVB*).

(83) *Eslliçar* (Maestrat, Maella, *DCVB*).

(84) *Plovitejar* és propi del Maestrat i de la Plana de Castelló (*DCVB*).

GAV 1980 (15 innovacions al *DFa*): *purnejar*; *creïlla*; *fraura*; *escuradora*; *perilla; gesmil, gesmiler; bascòll; gobanilla; chicotét; esvarar; esguitar; espentar; unflar; joguet.*

F. Pastor (6 innovacions al *DFa*): *aiguarrada*; *creïlla*; *pebrera* (sentit n.e. Fabra); *ximenera*; *llavaplats*; *llarguerut.*

Galmar (1 innovació al *DFa*): llavaplats.

**Riquesa lèxica i variació sinonímica**

No tots els vocabularis valencians tenen les mateixes dimensions. Ja en els títols alguns ens avisen del seu caràcter *bàsic, elemental* o *fonamental*, i els més ambiciosos adopten el títol de «diccionaris». Tanmateix, donat el caràcter bàsic de la majoria dels setanta mots triats, crec que serà il·lustratiu comparar els mots que manquen en les diverses obres estudiades. El menys complet és el bàsic del GAV (1980), amb 32 omissions,[85] algunes per ventura explicables per les dimensions reduïdes de l'obra; sorprenents són les onze omissions –ben bàsiques– de Fullana-Miedes, obra de gran format: *bocadillo, palomitas de maíz, lavaplatos-lavavajillas, bandeja, basura, bombilla, lagarto, jauría, agujetas, pequeño, en cuclillas*; les deu omissions de Valor també deceben una mica: *bocadillo, palomitas de maíz, lavaplatos-lavavajillas, bandeja, basurero, abanico, deshilar, deshilachar, calzoncillos, empujón, llamarada*; i també les set de Miedes: *tragón-glotón, aguijón, murciélago, abubilla, jauría, larguirucho, en cuclillas*; més normals són les quatre de F. Pastor: *palomitas de maíz, níspero* 'fruto', *agujetas, en cuclillas*; GAV 1983: *palomitas de maíz, lavaplatos-lavavajillas, jauría, agujetas*; i Brúixola: *lavaplatos-lavavajillas, junco, juncia, jauría*; quasi complets resulten Galmar (amb dues): *palomitas de maíz, níspero* 'fruto'; Acadèmia (dues): *lavaplatos-lavavajillas, agujetas*; i Pascual (una): *deshilachar.*[86]

El vocabulari més ambiciós pel que fa a la variació sinonímica és, sens dubte, el de l'Acadèmia. Encertats i originals em semblen els cinc sinònims aplegats per a traduir cast. *lloviznar*: *purnejar, plovitejar, borrimar, gotejar, gotinyar*; tenim on triar entre els onze que tradueixen cast. *chaparrón*: *arruixada, arruixo, aiguarrada, aiguaço, sacso d'aigua, alaboc, ramassada, glopada, garrotada, diluviada, chapallo*; o entre els tretze de *tragón-glotón*: *farto, fartana, fartera, engaldidor, galduf, gloto, galdirot, golut, golos, galafre, galima, gulo, buco.* Però no sempre és tan ric, de vegades es limita a un sol lexema, per exemple en els casos de *aladre, estrela, creïlla, pera* 'bombeta', *gesmil, coa, carranc, saltamartí* o *esguitar.*

(85) Vet-les ací: *arado, chaparrón, cucurucho, palomitas de maíz, requesón, solomillo, tragón-glotón, pimiento, níspero* 'fruto', *zanahoria, hollín, tomillo, romero, juncia, hiedra, fresno, aguijón, cangrejo, lagarto, murciélago, pavo, jilguero, abubilla, jauría, deshilar, deshilachar, calzoncillos, cosquillas, agujetas, párpado, larguilucho, en cuclillas.*

(86) I encara trobem *esfilagarsar* 'deshilachar' en l'invers.

## Consideracions finals

**Vocabularis de F. Ferrer Pastor**. Després del vocabulari valencià-castellà (1960) publicà el castellà-valencià (1967); uns anys després (1972) publica tots dos plegats –amb alguna reducció escadussera–, i al 1977 publica la que serà versió definitiva dels seus vocabularis –reproduïda en facsímil fins a les edicicions dels anys norantes–. De l'acarament de les edicions de 1967 i 1977 –pràcticament idèntiques– amb el vocabulari castellà-català de S. Albertí (1961) es desprèn que F. Pastor va aprofitar aquesta obra. En quasi la meitat dels setanta mots estudiats es pot rastrejar l'empremta del model: en 21 casos –el 30%– F. Pastor ha copiat literalment les traduccions d'Albertí;[87] en d'altres casos elimina algun dels sinònims d'Albertí –números 11, 29, 41, 52, 57–, en selecciona alguns –nos 2, 4, 5, 12–, i de vegades n'afig algun de nou –nos 1, 14, 53–. Malgrat el caràcter excessivament decantat cap a les solucions lèxiques de Catalunya, cal reconèixer als vocabularis de F. Pastor el mèrit de ser els primers que en els temps difícils oferien bona ajuda als qui decidien mampendre l'aprenentatge del valencià.

**Diccionari Galmar**. Publicat primer relligat i amb el títol de *Diccionari Castell* (1977), aparegué uns anys després en rústica i amb el nou títol de *Diccionari Galmar*.[88] El nostre estudi mostra que Galmar (1977) és reedició quasi idèntica d'Albertí (1961) –amb permís o sense, en tot cas no en diuen res–: en els setanta mots estudiats només hi ha sis canvis respecte a Albertí (1961) i segurament són deguts al fet que es degué partir d'una edició posterior.[89] No és estrany, doncs, que siga el vocabulari més catalanitzant. Només hi he trobat una innovació «valenciana»: *llavaplats*. Té el «mèrit» de ser el primer vocabulari «català» venut com a «valencià»...

**Vocabulari Bàsic del GAV (1980)**. De caràcter volgudament bàsic –hi manquen 32 dels 70 mots estudiats– i «escissionista», aquest diccionari és prou original i acompleix bé la seva intenció de no introduir cap mot no valencià, car en els setanta mots estudiats no n'hem trobat cap que no ho siga i de vegades dóna alguna solució encertada, com ara *escuradora* 'rentavaixella'.

**Vocabulari del GAV (1983)**. Encara que no es diu en la introducció, es tracta d'una adaptació «secessionista» –sobretot ortogràfica– del vocabulari valencià-castellà de Fullana (1921).[90] Podria haver tingut el mèrit –si s'hagués fet bé– d'oferir-nos l'arreplega de Fullana ordenada per sinònims, però els autors del vocabulari no han treballat gaire fi, car en només setanta mots trobem les errades següents: *sanguisolet* 'solomillo' per *sanguisoleta* (Escrig, Fullana); *solli* (errada per *sollim*); *englatina* <sic>; *mucilac* 'murciélago' (ja en Fullana); *esfilagarrar* <sic> 'deshilachar' (és errada evident per *esfilagarsar*); *depir* <sic> per *clapir*; *cosgolles*, errada per *cossigolles*; *espejar* 'empujar' errada per *espitjar*.[91] Aquest diccionari no acompleix gaire bé la seva intenció de no introduir cap mot no valencià, car en els setanta mots estudiats hem trobat els catalanismes següents: *espetec*;[92] *mato* 'requesón' (sentit n.e.

(87) Vegeu els números 3, 8, 19, 20, 21, 26, 30, 31, 33, 34, 42, 43, 45, 46, 47, 48, 49, 51, 58, 62 i 65.

(88) Vegeu-ne més explicacions a les referències bibliogràfiques.

(89) Vegeu els números 4, 7 i 38 (amb innovacions), 44 i 63 (amb canvis d'ordenació) i 14 (amb pèrdua); en els números 11, 35 i 69 la coincidència és amb l'edició reduïda Albertí (1975).

(90) Sobre aquest vocabulari vegeu Casanova (1989).

(91) Ja en Fullana *espejar* 'empujar'; n.e. ni Escrig, ni Llombart, ni MGadea; deu ser errada per *espitjar* 'espentar' (Mall., *DCVB*; també a la Safor i a la Marina Alta).

(92) Ja en Fullana: *espétec* 'crujido, chaparrón'; Llombart: *espetech* 'cruijido'; Martí i Gadea: *espetec* 'cat(alanismo). crujido'.

Fullana); *pebrot, pebrotera*:[93] *eura*[94] i *gessami*. Quan el mot de Fullana no els ha interessat, no han tingut inconvenient d'excloure'l, com ara *petit* –viu a Alcoi–.

**Vocabulari Elemental d'Emili Miedes (1983).** Sembla que té com a font bàsica Fullana (1921). Aquest diccionari acompleix bé la seva intenció de no introduir cap mot no valencià, car en els setanta mots estudiats només n'hem trobat un que no ho siga: *hèura*. No és gaire acurat, car hi hem trobat les errades següents: *pimiento* (errada per *pimento*); *bandeja* 'font / batea' <sic>; *ornella* <sic> (errada per *oronella*); *palput* 'cuclillo' <sic>. La mala grafia, les errades i la curtedat de mires de l'autor fan inutilitzable aquesta obra.

**Diccionari de Fullana-Miedes** (1985). És adaptació reconeguda –al contrari del GAV (1983)– i bastant respectuosa de Fullana (1921).[95] En aquesta obra –singular dins la lexicografia «secessionista»– Emili Miedes supera de bon tros el seu treball de 1983 –gràcies, és clar, a Fullana–. És l'obra «anticatalanitzant» que en més casos admet sinònims considerats pels seus col·legues com a massa «catalans», car admet mots com ara *bacallar, farigola, cua, cranc, pessigolles* i *defensar*. No es caracteritza per una cura especial, car hi trobem faltes d'ortografia com ara *esbarar*, badades com *salto* 'saltamartí', *mucilac* 'rata-penada' o *rosmari* 'romer', i errades com *ornella* 'oroneta' i *joquina* 'joguet'.

**Diccionari Gregal de Vicent Pascual (1987).** És sense dubte el més complet, el més útil i el més treballat de tots els vocabularis valencians. Ha sabut triar un bon model, el millor diccionari castellà-català fins a l'hora d'ara, el de l'Enciclopèdia Catalana (1985), l'ha copiat sense complexos,[96] l'ha adaptat amb un criteri ben encertat i l'ha completat amb nombrosíssimes formes valencianes. El seu criteri –compartit amb Brúixola i qualque vegada amb Valor i amb Fullana-Miedes o l'Acadèmia– és donar sempre en primer lloc la forma valenciana i després d'altres sinònims –no sempre estrictament valencians; així trobem *paloma* i *papallona*, *bacallà* i *abadejo* o bé *joguet* i *joguina*, *romer* i *romaní*.

**Vocabulari Fonamental d'Enric Valor (1988).** Té un caràcter genial i alhora una mica desconcertant. De vegades enregistra mots i variants que manquen en quasi tots els altres vocabularis, com ara *plovitejar* o *fiçó*; altres voltes oblida mots ben fonamentals com ara *bocadillo, palomitas de maíz, lavaplatos-lavavajillas, abanico, calzoncillos, empujón* o *llamarada*. De vegades és el més catalanista –és el campió, com hem vist adés, dels equilibrats però amb preferència per la forma catalana–, d'altres el més valencianista –*estómec, bàlsem, llarguerut*–. Valor es mostra enemic dels mots acabats en *-o*: *abadeig*,[97] *llomell*. Presenta l'originalitat[98] d'assenyalar el caràcter regional o dialectal d'alguns mots: *aladre*, reg. *arada*; *pebrot*, reg. *pebrera*, reg. *pimentó; sutja*, reg. *estalzim*, reg. *follí*; [*jonc*, dial. *junc*], *gesmí*, reg. *gessamí*, dial. *gesmir, gesmil; arrel*, reg. *raïl; papallona*, reg. *paloma; gall dindi*, reg. *titot*, reg. indià, reg. *indiot; canell*, reg. *monyica*. En un parell de casos es veu que ha tingut davant –i l'ha copiat– el diccionari de Moll (1977): vegeu *buc d'abelles, nispro* i *esbraonament* 'agujetas'.[99]

(93) *Pebrot* i *pebrotera* ja en Fullana i català oriental (*DCVB*).

(94) El *DCVB* dóna *heura* com a català occidental i valencià, però les dades de l'*ALDC* mostren una partició clara entre *heura* (cat. or.) i *hedra* (cat. occ.), veg. Veny 1988.

(95) Però en els casos més cridaners no té inconvenient de trair Fullana, com ara en el cas de *petit*, que resol –no solament com el GAV 1983, que l'eliminen– eliminant també l'entrada castellana *pequeño*...

(96) N'hi ha prou per a comprovar-ho amb acarar els dos diccionaris per qualsevol plana: tipografia, introducció de noms propis –de persona i de lloc–, fraseologia... Pel que fa a les nostres setanta paraules, hi ha coincidència total en els números 2, 21, 45, 46, 52 i 58, coincidència amb canvi d'orde dels sinònims en els números 4, 27 i 42 i coincidència més innovació en els números 11, 13, 18 i 39.

(97) També *abadeig* en GAV 1983 i *fardaig* en Miedes 1983.

(98) També es troba aquest tipus de marca alguna vegada en DEC 1985.

(99) Sobre els diccionaris de Valor vegeu F. Bernat i X. Favà (1992).

**Vocabulari Brúixola (1990).** Comparteix amb Pascual el criteri encertat de donar sempre preferència a la forma valenciana, sense oblidar els altres geosinònims. Presenta la novetat de donar la pronúncia dels mots –en general ben correcta–. Original, molt pràctic i prou complet: un bon vocabulari.

**Diccionari de la Real Acadèmia de Cultura Valenciana (1992-1993).** Esforç malaguanyat. El que podria haver estat el millor vocabulari valencià de tots els temps, restarà com a obra només útil als estudiosos de la llengua. Hom ha realitzat un gran esforç per desfer l'obra magna del *DCVB*: pacientment s'han dedicat a anar extraent tots els mots valencians d'aquest gran diccionari. En principi, el projecte era legítim i valuós –és el mateix que va fer el mestre Moll en el seu magnífic vocabulari mallorquí-castellà (1965)–, però ací –adoptant una grafia «separatista»– s'ha pretès anul·lar tota la tasca feta des dels anys trentes.[100]

JORDI COLOMINA I CASTANYER
*Universitat d'Alacant*

REFERÈNCIES BIBLIOGRÀFIQUES

Acadèmia (Real Academia de Cultura Valenciana) (1992) *I. Diccionari Valencià-Castellà*, València, ed. Del Senia al Segura, 530 pàgs.

— (1993) *II, Diccionari Castellà-Valencià*, V., ed. Del Senia al Segura, 407 pàgs.

ALBERTI, Santiago (1961) *Diccionari Castellà-Català* i *Català-Castellà*, Barcelona, 1182 pàgs.

— (1975) *Diccionari (Petit) Castellà-Català, Català-Castellà*, B.. 2ª ed., B., 1978, 338 pàgs.

BERNAT, Francesc & Xavier FAVÀ (1992) 'La llista de la RTVV (1990): Estudi lexicogràfic' dins Joan Solà (ed.) 1992, pàgs. 75-121.

BRÚIXOLA (1990) *Brúixola. Vocabulari Castellà-Valencià, Valencià-Castellà*, amb la col·laboració de Joan Ramon Borràs Hernandis i Ricard Arias Ramon, sense lloc d'edició, ed. Bibliograf, 254 + 235 pàgs.

CASANOVA HERRERO, Emili (1989) 'La faceta lexicogràfica del pare Fullana', *Boletín de la Sociedad Castellonense de Cultura*, LXVII, pàgs. 415-441.

COLOMINA I CASTANYER, Jordi (1992) 'Morfologia i ortografia: a propòsit d'algunes grafies controvertides', *A Sol Post*, 2, Alcoi, Ed. Marfil, pàgs. 129-153.

— (1993) 'Propostes valencianes per a l'estàndard oral', *I Jornades de Sociolingüística: «La llengua estàndard» (Alcoi 1992)*, Alcoi, Ajuntament, pàgs. 11-36.

— [en premsa] 'A propòsit d'una llei fonètica catalana poc observada', *Miscel·lània Germà Colón*, B., PAM.

(100) Vegeu la ressenya de Colón (1993).

COLÓN DOMÈNECH, Germà (1993) 'Crítica al Diccionario de la Academia de Cultura Valenciana: 1, Un propósito disgregador; 2, Un enredo ortográfico, i 3, El valenciano enfermizo', *El País*, ed. de València, 6, 7 i 8 d'abril de 1993.

DAg (Marian AGUILÓ i FUSTER) (1914-1934) *Diccionari Aguiló*. Materials lexicogràfics aplegats per M. A. i F., revisats i publicats sota la cura de P. Fabra i M. de Montoliu, 8 volums, B.

DCor (Joan COROMINES) (1980-1991) *Diccionari Etimològic i Complementari de la Llengua Catalana*, 9 volums, Barcelona.

DCVB (A. Mª ALCOVER & F. de B. MOLL) (1926-1968) *Diccionari Català-Valencià-Balear*, 10 volums, Palma de Mallorca.

DEC (Diccionaris Enciclopèdia Catalana) (1982) *Diccionari de la Llengua Catalana*, B., Enciclopèdia Catalana, 1679 pàgs.

— (1985) *Diccionari Castellà-Català*, B., Enciclopèdia Catalana, 1341 pàgs.

DFa (Pompeu FABRA) (1932) *Diccionari General de la Llengua Catalana*, B., 5ª ed., 1968, 1762 pàgs.

Escrig (Josep ESCRIG) (1851) *Diccionario Valenciano-Castellano*, V.: P. Aguilar, 900 pàgs. [Use la 2ª ed. V.: P. Aguilar, 1871, que sembla que és idèntica a la primera].

FERRER PASTOR, Francesc (1960) *Vocabulari Valencià-Castellà*, V., Sicània, 223 pàgs.

— (1967) *Vocabulari Castellà-Valencià*, V., Sicània, 477 pàgs.

— (1972) *Vocabulari Castellà-Valencià i Valencià-Castellà*, V., Gran Enciclopedia de la Región Valenciana, 596 pàgs.

— (1977) *Vocabulari Castellà-Valencià i Valencià-Castellà*, V., L'Estel, 1075 pàgs.

— (1985) *Diccionari General*, V. 1985, 1599 pàgs. 2ª ed., V., 1992, 1599 pàgs. i un apèndix verbal.

FERRER PASTOR, Francesc & Josep GINER I MARCO (1956) *Diccionari de la Rima*, V., Editorial Frederic Doménech, LXVII + 1122 pàgs.

— (1980) *Diccionari de la Rima*, V., 1797 pàgs. [Segona edició una mica ampliada de l'obra de 1956; hom ha eliminat l'ortografia i la morfologia que servien d'introducció, i també el nom de Josep Giner i Marco].

FULLANA MIRA, Lluís (1921) *Vocabulari Ortográfic Valenciá-Castellá*, V., ed. Edeta, XLII + 637 pàgs. Reed. facs., V., GAV, 1979.

FULLANA MIRA, Lluís & Emili MIEDES BISBAL (1985) *Vocabulari Valencià-Castellà i Castellano-Valenciano*, V., ed. Asval S. A., XIV + 823 + XXIV (entre pàgs. 438 i 439) + XV pàgs.

GALMAR (1977) *Diccionari GALMAR. Castellà-Valencià, Valencià-Castellà*, «equipo de redacción dirigido por» Rosendo CASES i Diego LISBOA, amb una «Gramàtica valenciana» per Lorenç <sic> Riba i Albert Forns, i «Diagramación y presentación» de Horacio Silva, sense numeració. [Sembla que aquesta edició aparegué cap a 1980 i és una reedició facsímil en rústica del *Diccionari CASTELL Castellà-Valencià, Valencià-Castellà*, V., 1977, del qual s'ha eliminat el «Resumen de Historia del País Valenciano» de R. Ventura Melià].

GAV (Colectiu de Mestres de la Secció de Pedagogia del Grup d'Acció Valencianista) (1980) *Vocabulari Basic de la Llengua Valenciana. Castellà-Valencià, Valencià-Castellà*, treball coordinat per Antoni Fontelles i Mª Pilar Hervàs, V., 246 pàgs.

— (1983) *Vocabulari Valencià-Castellà, Castellà-Valencià*, treball coordinat per J. Torán, G. Lacassaña i A. Roselló, V., 923 pàgs.

Llombart (Constantí LLOMBART) (1887) *Diccionario Valenciano-Castellano de J. Escrig*, 3ª ed. corregida y aumentada, V., P. Aguilar, 1231 pàgs.

MASCLANS I GIRVÈS, Francesc (1954) *Els noms vulgars de les plantes a les terres catalanes*, B., IEC, 251 pàgs.

MGad (Joaquim MARTÍ I GADEA) (1891) *(Novísimo) Diccionario General Valenciano-Castellano*, V.: Josep Canales Romà, 1975 pàgs.

MIEDES BISBAL, Emili (1983) *Vocabulari Elemental de la llengua Valenciana. Valencià-Castellà, Castellà-Valencià*, V., ed. Valencia 2000, 411 pàgs.

MOLL, Francesc de B. (1965) *Vocabulari Mallorquí-Castellà amb inclusió de les paraules típiques de Menorca i Eivissa*, Palma de Mallorca, 327 pàgs.

— (1977) *Diccionari Castellà-Català*, Palma de Mallorca, 389 pàgs.

PASCUAL, Vicent (1987) *Diccionari Gregal. Valencià-Castellà, Castellà-Valencià*, 512 + 509 pàgs.

— (1990) *Diccionari Tabarca. Valencià-Castellà, Castellà-Valencià*, V., 512 + 509 pàgs. [És segona edició (facsímil) del *Diccionari Gregal* de 1987].

RECIO, Carles (1985) *Diccionari Diferencial Valencià-Català, Català-Valencià*, V., Ateneu de la Cultura Valenciana, 336 pàgs.

SOLÀ, Joan (editor) (1992) *Sobre lexicografia catalana actual*, B., Empúries, 305 pàgs.

VALOR, Enric (amb la col·laboració d'Enric Ll. VALOR) (1988) *Vocabulari Fonamental*, B, Plaza & Janés editors, 501 pàgs.

VENY, Joan (1988) 'Noves dades sobre la partió català occidental-català oriental', *V Col·loqui d'Estudis Catalans a Nord-Amèrica*, B., pàgs. 11-41.

JAUME CORBERA

# LA LEXICOGRAFIA CATALANA MODERNA A LES ILLES BALEARS

Es pot dir que la lexicografia catalana moderna comença a les Illes Balears amb l'aparició del primer fascicle del *Diccionari Català-Valencià-Balear*, d'Antoni Ma. Alcover i Francesc de Borja Moll, l'any 1926, amb el qual tenim la primera obra feta amb rigor científic i un plantejament tècnic lexicogràfic adequat. Abans, durant aquest segle, s'havien publicat uns quants vocabularis d'interès dialectològic (els podeu veure referits a *Lexicografía Catalana (I)*), un vocabulari destinat als turistes que volguessen visitar Eivissa[1] i alguns *vocabularis* fets amb la finalitat de facilitar l'aprenentatge del castellà.[2]

En el present treball no tractaré ni del DCVB ni de cap d'aquests vocabularis. Del primer, ben conegut, en parlen extensament Colon-Soberanas (1986: 211-216); als altres, no m'hi referiré perquè no són pròpiament «lexicografia catalana»: el seu objecte no és servir un públic ample usuari actiu o passiu de la llengua catalana, sinó proporcionar unes dades a uns especialistes en dialectologia, o satisfer la curiositat d'uns viatgers descobridors d'un parlar «singular», o donar uns mitjans a uns escolars *necessitats* d'aprendre la *lengua nacional* per sortir del seu analfabetisme.

Les obres descrites i analitzades seran, doncs, aparegudes a les Balears amb posterioritat a l'acabament del DCVB (1962) i ja plenament inserides dins el moviment

(1) «Vocabulario de las dicciones del léxico ibizenco que no tienen analogía ni semejanza con el castellano y su significado», inclòs a Arturo PÉREZ-CABRERO, *Ibiza. Arte, arqueología, agricultura, comercio, costumbres, historia, industria, topografía. Guía del turista*. Barcelona, 1909; pàgs. 141-153.

(2) *Nociones de lenguaje. Vocabulario Mallorquín-Castellano. Grado Preparatorio*. Libro del Alumno. Por Miguel PORCEL Y RIERA; Palma de Mallorca, 1921 (8ª ed.; la primera, que no he pogut localitzar, deu ser de 1910/11). [N'hi ha una versió *«catalán-español»*, de la qual conec la setena edició (Palma de Mallorca, 1931).]

de recuperació lingüística empès els darrers vint anys.[3] Sobretot, si bé no estrictament de manera exclusiva, l'anàlisi es referirà al reflex que s'hi pugui apreciar de la particularitat dialectal balear en relació a la normativa vigent; és a dir, fins a quin punt aquestes obres mostren una voluntat de reivindicació de les variants i paraules insulars acceptades tàcitament o explícitament pel marc normatiu. Qüestions relacionades amb altres aspectes de la realització lexicogràfica i la seva metodologia ja són tractades per altres especialistes en aquest mateix número de *Caplletra*.

Presentaré les obres classificades segons els criteris exposats per Haensch (1982).

## 1. OBRES DE CARÀCTER LINGÜÍSTIC

### 1.1. Del discurs individual

M. COLOM MATEU, *Glossari General Lul·lià*. Palma, Moll, 1982-1985.

Després del *Diccionari Català-Valencià-Balear* és l'obra més ambiciosa de la lexicografia balear. Ofereix, ordenat i documentat, el lèxic català de Ramon Llull o dels seus copistes antics, incloent el d'escrits adulterats o de dubtosa paternitat lul·liana. La presentació del material, amb totes les accepcions ben definides, exemplificades i degudament diferenciades, i amb l'ètim corresponent, en fa una eina indispensable per al coneixement de la llengua de Llull, i utilíssima en l'estudi i la documentació primerenca de moltíssimes de formes i paraules de la llengua general.

### 1.2. Del discurs col·lectiu

1.2.1. *De caràcter general*

1.2.1.1. Monolingües

1.2.1.1.1. Onomasiològiques

A. LLULL MARTÍ: *Vocabularis temàtics per a correcció de barbarismes i enriquiment del lèxic personal*. Ajuntament de Manacor, 1979.

Selecció d'uns 12.000 mots repartits en cinquanta vocabularis temàtics, els quals estan agrupats en dotze «temes bàsics» (*I. L'UNIVERS. ELS ASTRES, LA TERRA, GEOGRAFIA. II. LA VIDA. LES PLANTES, ELS ANIMALS.*, etc.). L'obra té una inten-

*Nociones de lenguaje. Vocabulario Mallorquín-Castellano. Grado Medio.* Libro del Alumno. Por Miguel PORCEL Y RIERA; Palma de Mallorca, 1912 (2ª ed.).

Els anys 30 encara es publicaren els següents:

*Vocabulario para la enseñanza del castellano en Mallorca*, arreglado por Un Maestro Nacional de Palma. Palma, s/d [1933].

*Vocabulario Mallorquín-Español*, redactado por Antonio VIVES GINARD. Palma, 1935.

(3) Baldament comparteixi els criteris exposats per Cabré-Lorente (1991: 36-38) per a la consideració d'una obra com a «lexicogràfica», no he exclòs de la meva anàlisi alguns documents que, d'acord amb tals criteris, no hauria d'haver tengut en compte, tots els quals són fruit de la inquietud que durant aquest període s'ha manifestat per conèixer i difondre els noms adequats als conceptes més diversos.

ció normativa i correctiva, i per això incorpora molts de tecnicismes i termes científics que l'autor considera poc coneguts pels parlants (sobretot els de 1979!). Com que no fa definicions, per facilitar la feina de correcció al lector, sempre que una paraula catalana pot ser-li dubtosa o poc coneguda s'hi assenyala devora, en negreta i entre parèntesis, la paraula castellana corresponent, amb la suposició que aquesta és ben entesa. Per exemple: *«fong (**hongo**), liquen, floridura (**moho**); molsa (**musgo**), herba, arbust, arbre»* (§ 8.1.).

Bon coneixedor del seu propi dialecte, al costat de les formes més generals l'autor inclou en aquests *Vocabularis* les variants i paraules mallorquines (*«NÚVOL o nigul»*, *«CALABRUIX, pedregada o calamarsa»*, *«terebint o (mall.) «arbre del pi»»*), encara que de vegades només n'hi ha una, la forana (*murtrera*, § 9.2.) o l'autòctona (*camamil·la*, § 9.3.). Fins i tot a vegades hi posa aquelles que són considerades vulgars no normatives (com *«gatosa (mall. vg. «gatova»)»*), amb la intenció de promoure la recuperació del parlar més col·loquial, ja greument amenaçat a la darreria dels 70.

L'obra, que inclou també noms geogràfics i gentilicis, no duu cap tipus d'informació gramatical sobre els mots ni d'explicació contextual o semàntica, excepte en molt comptades ocasions, ni identifica tampoc les espècies animals o vegetals amb el nom científic. A pesar de la seva senzillesa, però, és una obreta útil per ajudar aquelles persones amb dèficit de vocabulari a recuperar moltes denominacions oblidades o arraconades. En aquest sentit, segurament seria molt més eficaç si els vocabularis fossen acompanyats d'il·lustracions que ajudassen a captar molt més clarament el significat d'algunes de les paraules aplegades.

1.2.1.1.2. Semasiològiques

J. CORBERA POU: *Vocabulari de barbarismes del català de Mallorca*. Palma, Promotora Mallorquina de Mitjans de Comunicació, 1983 (2a. 1984).
J. CORBERA POU: *Nou vocabulari de barbarismes del català de Mallorca*. Palma, El Tall, 1993.[4]

Té per finalitat ajudar el lector a reconèixer aquelles faltes de llengua, principalment de vocabulari, imputables a la interferència d'un altre idioma –especialment el castellà–, fornint-li a la vegada la solució autòctona mallorquina correcta. Va sorgir de la necessitat de dotar el públic interessat en la correcció lingüística d'una obra que partís dels errors més comuns en mallorquí i que tengués presents les solucions genuïnes autòctones, sent les obres d'aquest tipus existents prèviament referides únicament al català central. La primera versió té una presentació més senzilla: una primera part de vocabulari general, on es mesclen tots els barbarismes de qualsevol origen; una segona part d'explicacions i comentaris sobre algunes de les faltes denunciades a la prime-

(4) A pesar dels dubtes raonables expressats per Mercè Lorente (1993), no em sembla desencertada la consideració dels vocabularis de barbarismes com a obres lexicogràfiques monolingües, ja que són fonamentalment repertoris de paraules usades parlant una llengua determinada i elaborats amb la mateixa tècnica que els diccionaris generals, encara que hi hagi la possibilitat, com en aquest cas, que el llibre inclogui una sèrie d'explicacions discursives sobre temes gramaticals.

ra; una tercera, més breu, que crida l'atenció sobre determinats fets d'evolució de la llengua deguts sense dubte a la influència forastera; i una quarta sobre els excessos injustificats en la correcció en què sovint, amb la idea de fugir del castellà considerat malèfic, cau el parlant. Al final hi ha una bibliografia de lèxics especialitzats perquè els interessats puguin completar pel seu compte les seves mancances terminològiques. La segona versió fa un plantejament més didàctic del contengut: reparteix el lèxic general (al qual hi ha algunes supressions –no justificades enlloc– i algunes addicions) en diversos apartats, segons l'origen del barbarisme: «Barbarismes directes», «Mots mallorquins alterats formalment o semànticament» i «Calcs diversos del castellà»; afegeix alguns capítols a la segona part, i refà parcialment la segona i la tercera. La bibliografia del final hi ha desaparegut, havent-hi ja tants de títols publicats i popularitzats arreu, sobretot gràcies al TERMCAT. El *(Nou) Vocabulari* és una obra agosarada en la reivindicació que fa de les formes autòctones mallorquines, moltes absents del «Diccionari Normatiu» i bastants d'altres quasi desconegudes per les darreres generacions.

1.2.1.2. Bilingües

F. de B. MOLL: *Vocabulari Mallorquí-Castellà amb inclusió de les paraules típiques de Menorca i Eivissa.* Palma, Moll, 1965.

Vet ací, amb paraules de l'autor, el seu propòsit en fer el llibre:

> *No pretenem registrar «tots» els mots usats en el llenguatge de les Balears, sinó els més importants i aquells dels quals més interessa conèixer l'equivalència castellana* (pàg. VII).
>
> *És indubtable que molt sovint aquest Vocabulari, en lloc de servir per a aprendre l'equivalència castellana dels mots mallorquins, haurà de servir per a aprendre, a base del castellà, la manera d'escriure el vocable mallorquí corresponent: exercirà, doncs, també les funcions d'un diccionari ortogràfic, que per a molts de baleàrics és un article de primera necessitat* (pàg. XI).

L'obra, per complir aquest propòsit, es centra sobretot en el lèxic propi de les Illes:

> *el material lingüístic que hi figura alfabetitzat és, fonamentalment, el que s'usa en el llenguatge parlat com a vernacle a l'illa de Mallorca... cada vocable va seguit de l'equivalència castellana. Però com que el llenguatge de Mallorca és, en més del 90 per cent, idèntic al de Menorca i Eivissa, crec que interessa incloure en el registre present els pocs centenars de paraules típiques d'aquestes dues illes...* (ibidem).

> *he distribuït la matèria del Vocabulari en aquesta forma:*
> *A) Una sèrie general de paraules que formen el lèxic que podríem anomenar «bàsic» del llenguatge de les Balears».*
> *B) Intercalades entre les d'aquella sèrie general, les paraules que no s'usen normalment en el llenguatge viu de les Illes, però que són o han estat usades pels escriptors baleàrics pel fet de pertànyer al llenguatge literari general.*
>
> *La sèrie A... conté dues classes de vocables: els «hereditaris» o populars..., i els «adoptats» o cultismes i neologismes de ciències, arts, tècnica i deports... Dins aquesta sèrie A, faig una distinció que crec interessant: indic amb signes especials aquells mots o significats que són típics de Menorca o d'Eivissa en relació a Mallorca ...* (pàg. IX).
> *... Per distingir bé aquests mots* [els de la sèrie B] *no usuals en el llenguatge parlat de les Balears, van impresos en una vertical diferent de la de la sèrie general, entrats prop d'un centímetre cap a la dreta de l'extrem esquerre de la columna* (pàg. X).

A la macroestructura hi trobam moltes variants considerades no normatives (*corema*, *coremer*, *paupar*, *roada*, *roegó*, *soiar*...), amb la indicació «V.» [«vegeu»] que remet a la forma correcta.

En conjunt, és un *vocabulari* molt simple (qualque equivalència és, per massa simple, errada: *COLLBLAU pato*), sense gaires indicacions gramaticals: només, si és «verb reflexiu». Tampoc no explica, generalment, les diferències entre les diverses traduccions possibles, corresponents a altres tantes accepcions.

Des del punt de vista que més ens interessa, però, sí que reflecteix plenament, com es dedueix de les citacions anteriors, la particularitat lèxica balear.

F. de B. MOLL: *Diccionari català-castellà. Diccionari castellà-català*. Palma, Moll, 1978.

Ens centrarem en l'edició conjunta d'aquests dos diccionaris que el 1977 havien estat publicats per separat. De fet, no hi ha diferències entre la primera versió i la segona, si no és en la introducció («instruccions» i «criteris»): a la primera part (*català-castellà*) és la mateixa que primerament havia sortit al volum solt *castellà-català*, paraula per paraula, amb els mínims retocs indispensables per adaptar-la al plantejament invers i l'edició conjunta. A la segona part (*castellà-català*), la introducció és una adaptació-traducció de la primera, repetint les mateixes idees i explicacions, però canviant els exemples.

Vegem, novament amb les seves pròpies paraules, quin és el propòsit de l'autor amb aquests diccionaris:

> *Dins les evidents limitacions d'espai pròpies d'un llibre manual, volem oferir a tota la gent de parla catalana tots aquests elements d'informació:*
> *a) Un aplec aproximat de 60.000 vocables amb llurs equivalències en l'altre idioma.*
> *b) Una idea clara dels diversos significats dels mots que en tenen més d'un, amb la corresponent distinció semàntica en l'altra llengua.*
> *c) Una quantitat considerable de sinònims dintre de cada accepció dels mots, de manera que hi figurin representats els respectius vocables de cada contrada dels Països Catalans* (pàg. VIII).

Tècnicament, la presentació del diccionari és bona. (Vegeu-ho als paràgrafs ***Disposició dels articles***, ***Estructura dels articles*** i ***Indicació dels significats***.)

A les pàgines XII-XIII justifica els criteris de selecció pels quals admet molts de «dialectalismes» (*«mots o variants de mot que en el llenguatge parlat s'usen en una porció relativament petita del territori on es parla un idioma»*). Els criteris són teòricament i en general ben raonats i acceptables, però ens semblen inadequats en casos com els de *roella* i *roada* (a la primera part recull ***roella*** com a forma secundària de ***rosella***, però aquella no surt a la segona part com a traducció del castellà ***amapola***; en canvi, ***roada*** surt a totes dues parts), o *acuar*, tengudes sempre per «vulgars». Formes com *homo*, *noltros*, *revull* (per *reüll* < *rull*, adjectiu), que trobam als articles corresponents, tradicionalment rebutjades per la llengua literària balear, no són justificades enlloc. Com tampoc no ho és, com a traducció possible de la frase castellana ***«¡chúpate ésta!»***, l'expressió *«esclafa-la an* [sic] *aquesta!»*. En el seu desig de donar sinònims dialectals, l'autor sovint no té en compte, a més, diferències de registre o cronològiques. Així, com a equivalents del castellà ***achicar*** (en el sentit d'«humiliar»), posa els verbs catalans *«acovardir, acollonir»*; o, encara més desconcertant, tradueix ***«prepucio»*** per *«prepuci, cap de fava* [!, i és inexacte!], *capulla»*. I, per altra banda, a l'article ***lepra*** trobam els equivalents catalans *«lepra, llebrosia, mesellia»*, mot, aquest darrer, absolutament desusat avui en dia. Normalment, els sinònims hi són sense indicacions de pertinença dialectal; en el cas de l'entrada ***guisante***, però, s'hi especifica: *«pèsol* (en Cataluña y Mallorca), *fesol* (en Menorca), *estiregassó, xítxero* (en Mallorca)». I ens quedam amb les ganes de saber com ho diuen a València, les Pitiüses i l'Alguer... En un darrer exemple de la confusió metodològica d'aquesta obra, vegem que tradueix el castellà ***buenas noches*** pel català *«bona nit, bon vespre»*, sense cap aclariment sobre l'ús diferenciat que s'ha de fer de les dues fórmules.

A les pàgines XIII-XVI explica els criteris que justifiquen l'admissió d'algunes formes generalment considerades fins aleshores «barbarismes» i, per tant, absents del diccionari Fabra i de la majoria dels altres diccionaris. Entre les formes que admet amb arguments més o manco ben fomamentats, hi ha *enxufar* i *enxufament*, *entregar*, *triumf* i *triumfar*, *servici*, *sèptim*, *dècim*, *vigèsim* (ordinals)... En canvi, no admet *petroler*, que hauria de ser *petrolier*.[5]

(5) Vegeu altres característiques d'aquest diccionari a Beaulieu - Farriol - Rosés (1992).

F. de B. MOLL, N. MOLL MARQUÈS, *Diccionari escolar català-castellà, castellà-català*. Palma, 1984.

És una versió reduïda de l'anterior. La reducció afecta tant la macroestructura com la microestructura, que ha estat simplificada a molts d'articles, llevant-ne sinònims, especificacions, explicacions i algunes accepcions (v. p. ex. l'article ***pardal***). Té la mateixa pretensió de ser útil als usuaris de tots els Països Catalans, per la qual cosa incorpora, dins les limitacions que s'ha marcat, mots de tots els dialectes.

Manté els errors i les propostes discutibles de l'origen, i fins els augmenta: hi trobam *home, homo* com a equivalents del castellà ***hombre***, així com *rosada* i *roada* pel castellà ***rocío***. Però mentre *homo* té entrada també a la part català-castellà, *roada* n'és absent.

1.2.2. *De caràcter parcial*

1.2.2.1. Vocabularis dialectals

P. MELIS PONS, *La llengua de Menorca. Recull de paraules menorquines (I)*. Ciutadella, Col·lectiu Folklòric de Ciutadella, 1988.
P. MELIS PONS, *La llengua de Menorca. Recull de paraules menorquines (II)*. Ciutadella, Col·lectiu Folklòric de Ciutadella, 1990
P. MELIS PONS, *La llengua de Menorca. Recull de paraules menorquines (III)*. Ciutadella, Col·lectiu Folklòric de Ciutadella, 1992.

Després d'una introducció històrica a la llengua de Menorca, l'autor estudia una sèrie de paraules típiques de l'illa o que hi tenen un sentit particular: n'explica el significat, l'origen, la vitalitat... Els mots són presentats per ordre alfabètic, i amb els tres volums apareguts fins ara arriba a la lletra «J». No hi ha distinció d'allò que són paraules absolutament autòctones (com *abadeu, abotinar-se, aderg, afitó, allambrinat*...), de les que ho són només semànticament (*agatinyar, amarinar, bacó, banyadura*...), o de les que són simples variants formals (*àbit* 'api', *abulló* 'albelló', *beerols* 'abellerols', *cobròmbol* 'cogombre', *dixar* 'deixar'...). En alguns casos, la paraula o la variant no és típica únicament de Menorca, i la inclusió hi és justificada per raons diverses (*abatzer, agró, andoiar*...). En general, l'autor segueix els estudis de F. de B. Moll i J. Coromines per explicar els mots (i, per als anglicismes, també els de V. Ortells i X. Campos), acceptant-ne a voltes també els errors (per exemple, en relacionar *bòrn* 'medusa' amb «borm» 'malaltia del bestiar' i fer-los derivar tots dos del llatí MORBU; *bòrn* ve del grec BROMOS 'pudor' [Corbera: 1993, 523]); però també

adesiara aporta la seva pròpia interpretació i etimologia, quasi sempre ben raonada (per exemple, *abarbatar*). En conjunt, és una aportació de síntesi ben útil per conèixer la singularitat lèxica del dialecte menorquí, que hauríem de veure conclosa ben aviat.

M. PESCE VICH, *Paraules de les Pitiüses quasi oblidades*. Eivissa, 1986.

Fulletó de 27 pàgines, de mida quartilla, amb un vocabulari d'equivalències «eivissenc-castellà» i un altre de «castellà-eivissenc», sense cap indicació gramatical ni explicacions. L'autor pretén fer un aplec de mots i expressions típics dels parlars eivissenc i formenterer que actualment són poc usats o quasi desconeguts. Sorprèn que els dos vocabularis no són equivalents, sinó que paraules que trobam a un no són a l'altre, i viceversa. La grafia catalana no s'atén sempre a la normativa, com no no s'hi atén tampoc globalment la curta «Introducció». Tot plegat, fa la impressió d'un recull «folklòric», més que no d'una obreta amb voluntat reivindicativa.

E. RIBAS I MARÍ: *Aportació Pitiüsa al DCVB*. Eivissa, I.E.E., 1991.

Una interessantíssima i utilíssima ampliació del DCVB (més de tres mil sis-centes entrades) a càrrec d'un eivissenc molt bon coneixedor del seu dialecte. L'arreplega es basa en un bon nombre de documents escrits i informacions directes d'un grapat de persones d'Eivissa i Formentera.

La presentació del material es fa seguint el model del DCVB, el qual convé consultar per comprendre bé el valor de la forma aportada. Moltes d'aquestes són simples variants de mots ja recollits, altres són ampliacions semàntiques pròpies de les Pitiüses i altres, potser en nombre menor, són paraules del tot noves absents del recull d'Alcover-Moll.

Un únic retret s'hi pot fer, potser, a aquesta important *Aportació*: la inclusió sobrera de castellanismes documentats antigament, dels quals no en dóna constància actual (per exemple, *alvedrio, bailia, cautiveri, cavalgadura, cerda, consuelo*...). Tenen interès per a la història del lèxic, però no es poden considerar «formes pitiüses», sent purament ocasionals. També es podria haver prescindit de grafies simplement «errònies», com *absota* per «absolta», de valor poc més que anecdòtic.

1.2.2.2. Vocabularis terminològics

M. FULLANA, *Diccionari de l'art i dels oficis de la construcció*. Palma, Moll, 1974.

Obra mestra de la lexicografia especialitzada. El seu autor (aparellador de professió) coneix perfectament l'ofici de la construcció i coneix també amb profunditat la llengua. El *Diccionari* es basa sobretot en la terminologia mallorquina, acollint les variants formals pròpies de l'illa, si bé amb una certa hesitació (així, per exemple, *DEPÒSIT o DIPÒSIT, EMBÚS o EMBÓS, MORRALLÓ o MORRIÓ, MOTLLE o MOTLE, MOTLLURA o MOTLURA*; però *EMMOTLLAR, EMMOTLLAT*), però contenint igualment nombrosos noms propis dels altres dialectes. Els articles són molt complets, amb la definició ben clara del terme i la traducció castellana, en totes les accepcions en què s'usa. Allò que en fa una obra utilíssima és, però, molt especialment la gran quantitat d'il·lustracions que acompanyen les definicions, d'una qualitat i claredat excepcional. L'addició d'un «Glossari ideològic» i d'un «Glossari castellà-català» acaben de donar a l'obra el seu valor extraordinari. Amb el DCVB i el *Glossari General Lul·lià*, i considerant la diferència essencial entre les tres obres, és una de les realitzacions més importants de la lexicografia balear moderna.

J. SERRA I COLOMAR, *Diccionari de carrer*. Maó, Patronat Municipal de Cultura, s/d.
*Lèxic dels establiments alimentaris*. Col·lecció Lèxic Núm. 1. Consell Insular de Menorca, 1989.
*Lèxic dels esports*. Col·lecció Lèxic Núm. 2. Consell Insular de Menorca, 1989 (2ª 1992).
*Lèxic de l'habitatge*. Col·lecció Lèxic Núm. 3 Consell Insular de Menorca, 1990.

Les institucions menorquines són a les Balears les que més han fet per promoure la llengua catalana. Dins les seves accions s'emmarquen aquests 4 fulletons, de plantejament senzill i contengut reduït. El *Diccionari de Carrer* fa referència «*bàsicament a les coses i accions més elementals que formen part efectiva del carrer: edificis i vies de comunicació sobretot*» (pàg. 3). Cada entrada, sense cap indicació gramatical, duu la traducció castellana i una definició curta i simple, a vegades reduïda a un sinònim. Al final hi ha un vocabulari de barbarismes, sense definicions. Encara que es basa sobretot en el lèxic propi menorquí, hi trobam també qualque mot estrany a l'illa (com *atzucac*).

Els tres títols de la col·lecció «Lèxic» tenen en comú la seva simplicitat: són llistes de termes catalans agrupats per camps conceptuals («mesures», «forns de pa i pastisseria», «material esportiu divers», «jardineria», «tipus d'habitatge», «fusteria i mobiliari», etc.). El segon i el tercer duen també l'equivalència castellana, però el primer només la duu en els mots considerats més mals d'entendre (per ex.: *Queviures (cast. «víveres»), Codonyat... (cast. «membrillo»)*. No hi ha, tampoc, cap indicació gramatical. Tot i sent fets a Menorca, aquests lèxics no proposen purament les formes o els mots menorquins, quan són particulars, sinó que aquests solen anar acompanyats de sinònims considerats més generals: *«ROVELL D'OU; VERMELL D'OU»*, *«MELICOTÓ, melicotons; PRÉSSEC/s»*, *«TOMBARELLA/es; TOMBESSA/es»*, *«TORNESCRÚ (*anglicisme) destornillador TORNAVíS tornavisos destornillador»*, *«AMPOLLA... botella BOTELLA... botella BòTIL... botella»*. En alguns casos, però, el lèxic s'oblida inexplicablement de la paraula o forma autòctona i en proposa només una de forana: *«RASPALL/s cepillo»* (en menorquí *ESPALMADOR*), *«ÀNEC, ànecs (cast. «pato»)* (en menorquí *ÀNEDA*). Per altra banda, vacil·len en l'adopció de variants formals pròpies normativament correctes, acceptant-les a voltes sense restriccions (per exemple, *ESPATLA*, *GUÀTLERA* [al lèxic núm. 1]) i rebutjant-les altres vegades (per exemple, *AMETLLA / AMETLLAT, motllo* [al lèxic núm. 1], *ESPATLLA* [al lèxic núm. 2], *MOTLLE* [al lèxic núm. 3]). La vacil·lació es manifesta també en altres formes: per una part hi trobam *BAGUL* en lloc de l'autòctona i més etimològica *baül*, i per l'altra hi llegim *GARNERA* i *GARNERETA*, variants considerades «vulgars» de les originals i normatives *granera* i *granereta*.

*Especejament del Bestiar Boví, Ovi i Porcí.* Consell Insular de Menorca, 1987.

És un cartell plegat en quatre fulls amb la terminologia de les peces del cos del bestiar boví, ovi i porcí amb valor comercial. Mostra per una cara la canal de cada espècie amb les peces que en surten i el seu nom corresponent, mentre que per l'altra cara hi ha un vocabulari català-castellà d'aquests noms. La intenció del «Servei d'Assessorament Lingüístic» del Consell Insular de Menorca, en editar-lo, és la de promoure l'ús del lèxic propi del sector, molt acastellanat, com tota l'activitat comercial. En aquest sentit, si bé vol reflectir, escrivint *coa, espatla*, l'especificitat dialectal menorquina, dins el marc de la normativa catalana comuna, introdueix termes no autòctons com *xai, palpissos, rellom*, etc. Inexplicablement, substitueix la denominació balear de *xulla* per la continental de *cansalada*.

*Suplement al «Vocabulari de bars i restaurants-bar» i al «Vocabulari de restaurants»*. Fundación Barceló, s/l, s/d.

Es tracta de dos fulls solts, en mida de foli, sense referència d'autor. A un, per una sola cara, hi ha una llista de «Plats i productes típics de Mallorca que no figuren al «Vocabulari de bars i restaurants-bar», ni al «Vocabulari de restaurants»». A l'altre, per les dues cares, hi ha una llista de «Sinònims i variants emprats a Mallorca d'algunes formes del «Vocabulari de bars i restaurants-bar», i del «Vocabulari de restaurants»». Les llistes són en general encertades, i són a més un testimoni de la diferència important en gastronomia que hi ha entre les diverses regions catalanes, i de la gran dificultat de trobar una terminologia estàndard en aquest camp. Alguns retrets, però, s'hi poden fer, com la discordança entre una i altra que mostren els mots *Rubiols* («Plats i productes típics...») i *Robiolets* («Sinònims i variants ...»); segurament la segona grafia és més adequada a la pronúncia general, atès que l'ètim de la paraula és incert. En qüestions de grafia, també és discutible el *Tombet*, dit pertot *tumbet*, i seria preferible *pelaia* (< PELAGICA) a *Palaia*; igualment, les grafies *ametlla* i *guàtllera* no reflecteixen les variants mallorquines, ben legítimes, *ametla* i *guàtlera*; *fessols* i *fava perada* pareixen una simple errada d'impremta. Quant a les variants proposades, *fregida* és tan irreal com *Freginada* (la forma mallorquina és *fritada*) i *porc singlar* és ben sobrera. Finalment, sembla prescindible el castellanisme *Amargos*, perfectament substituït per *Panellets*, i és innecessària l'addició de *llobina* i *llop*, sent *llobarro* també denominació mallorquina.

J. ROSSELLÓ LLITERES, *Diccionari de la menestralia llatí-català*. Palma de Mallorca, Arxiu Diocesà de Mallorca, 1991.

Fulletó d'equivalències llatí-català, de només 8 pàgines sense numerar (de 24 x 17 cm), de termes referits sobretot a la menestralia, però també a altres estaments ciutadans de l'edat mitjana. Sense cap pretensió filològica o lexicogràfica, hi manca qualsevol informació gramatical.

M. MASSUTÍ, *Glossari de terminologia pesquera*. Palma, Conselleria d'Agricultura i Pesca, 1991.

No és aquesta una vertadera obra lexicogràfica, sinó més tost una espècie de «guia informativa» sobre la pesca a les Balears presentada per ordre alfabètic d'uns termes

fonamentals. Així, al costat d'entrades d'interès lingüístic (*abissal, anòxia, anticicló, aqüicultura, arqueig, astacicultura, balisa - balisament, batimetria, batipelàgic...*), n'hi ha d'altres d'interès sobretot cultural (*administració de l'aliment en cultius, alimentació en cultius (pinsos composts), arts d'arrossegament, autòtrofs vegetals, avaluacions d'estocs...*) i una gran majoria d'informació jurídica, burocràtica o administrativa (*acta d'infracció, administracions pesqueres, aigües interiors, ajudantia de Marina, ajudes a l'aqüicultura, Alcúdia (Confraria d'Alcúdia), associació d'armadors de vaixells de pesca...*). Com és lògic, als articles hi falta qualsevol informació complementària gramatical. Té un interès (socio)lingüístic, així i tot, l'article *normalització lingüística*, que reprodueix el decret 65/1989, del 22 de juny, del Govern Balear, pel qual es regula i es normalitza la nomenclatura popular de les espècies marines amb interès pesquer a la Comunitat Autònoma de les Illes Balears.

B. PALOS I VADELL: *Recull de terminologia muntanyenca relacionada amb l'excursionisme*. Palma, Ed. Moll, 1992 (2ª). (1ª ed., Conselleria de Cultura, Educació i Esports - Ed. Moll, 1991.)

Recull de termes i expressions que un excursionista necessita saber per caminar per les terres catalanes: accidents geogràfics, fenòmens atmosfèrics, esports de muntanya, construccions rústiques tradicionals... Seguint el seu propi criteri, ben fonamentat, l'autor ha fet la selecció a partir de dues obres magnes de la lexicografia catalana: el DCVB i la GEC. No s'ha contentat, però, de copiar simplement els articles d'aquestes obres, sinó que quan li ha semblat convenient hi ha afegit exemples i explicacions propis, o ha simplificat les definicions. Encara que a vegades hi trobam noms definits amb termes que no són al *Recull* (vegeu, per exemple, *llenegada, llenegadís*), altres hi trobam comentaris sobrers i/o inexactes (per exemple: ***Tarda** f castellanisme...*), i altres equivalències que, per a un mallorquí, no són ben encertades (com *«**Mac** m Còdol»*; a Mallorca un *mac* és una pedra de la mida, més o menys, del puny, no necessàriament arrodonida per l'aigua), en conjunt és un recull útil que pot ajudar molt a recuperar els termes adequats en el coneixement del paisatge terrestre. Una bona col·lecció d'il·lustracions hi ajuden molt, a aquesta finalitat.

SEMINARI PERMANENT DE CIÈNCIES SOCIALS DEL CEP DE PALMA DE MALLORCA, *Vocabulari bàsic de la Reforma*. Centre de Professors de Menorca, 1992.

Com el *Glossari de terminologia pesquera*, més que una obra lexicogràfica és una introducció als conceptes nous aportats per la reforma de l'ensenyament primari i secun-

dari. Els articles tenen tots, en aquest sentit, un plantejament tècnic, amb explicacions llargues i esquemes i diagrames il·lustratius. L'únic valor lingüístic que comporta és l'intent de fixació d'una terminologia adequada a les noves formulacions didàctiques.

ABRINES, B., i altres, *Brulla. Vocabulari bàsic per al cicle inicial.* Palma, Conselleria de Cultura, Educació i Esports, 1987.
ABRINES, B., i altres, *Canya. Vocabulari bàsic per al cicle mitjà.* Palma, Conselleria de Cultura, Educació i Esports, 1988.
ABRINES, B., i altres, *Blat. Vocabulari bàsic per a l'ensenyament secundari obligatori.* Palma, Conselleria de Cultura, Educació i Esports, 1990.

Són tres vocabularis «bàsics» per als distints nivells de l'ensenyança obligatòria, amb propostes didàctiques per als professors. Ofereixen unes llistes de termes, per camps conceptuals, sense cap tipus d'explicació o definició, amb els quals els mestres i alumnes han de poder treballar la llengua a cada nivell. El vocabulari és acompanyat d'il·lustracions, abundants i fonamentals per a la seva utilització a *Brulla*, reforçadores de la grafia a *Canya*, i útil complement a *Blat*. No entrarem a valorar els tres títols des del punt de vista didàctic, car no és aquest el nostre objectiu, sinó que farem algunes remarques sobre el contingut.

Fets per a les escoles balears, pareix que haurien de proposar sobretot les formes i variants autòctones (en els casos minoritaris de discrepància respecte als dialectes continentals) més que no altres de foranes. Això es compleix en molts de casos (per exemple, *mostel*, *escarabat*, *esclata-sang / pebràs*, *garrova*, *nora*, *picapedrer*, *desset*, *devuit*, *denou*...) Però en molts altres hi trobam vacil·lació i, sovint, un oblit inexplicable de les maneres de dir illenques. Hi llegim, per exemple, *ametler* i *ametller*, *ametlat* i *ametllat*, *batle* i *batlle*, *espatlla*, *motlle*, *revetla*, *rotlo*, *vetlar*...; *desbarat* és al costat de *dipositar* i *disfressa*; a «Canya» i «Blat» trobam *cua* exclusivament, que alterna amb *coa* a «Brulla»; hi ha *llenguado* en lloc del nom oficial balear *palaia*, *blat de moro* per *blat de les Indies / blat d'India* o *dacsa*, *peixateria* hi fa oblidar *pescateria*... *Calàpet* alterna amb *gripau*, *aiguamoll* amb *albufera*, *núvol* amb *nigul*, *pepa* amb *nina*; *nina*, *boixa* i *filleta* amb *nena*, *mostatxos* amb *bigotis*, *granera* amb *escombra*, *poal* amb *galleda*... L'*espalmador* només hi és com a objecte per «espolsar», mentre que per a «pentinar» o «netejar-se les dents» cal usar el *raspall*. *Calces* es dóna com a sinònim de *mitges* i de *bragues*. En canvi, alguns mallorquinismes s'imposen a les altres illes (*granot*, *grampó*) i hi ha *gri* per *grill*, forma aquella tenguda sempre per no normativa. Finalment, no hi manca qualque castellanisme innecessari, com *cutxilla* i *empotrat/ada*; o qualque ultracorrecció, com *fava pelada*.

P. TOMÀS PALMER, G. VICENS MIR, *Diccionari de biologia general*. Universitat de les Illes Balears - Conselleria de Cultura, Educació i Esports, 1994.

Recull exhaustiu de termes de biologia, definits i traduïts al castellà, útil tant per als ensenyants de nivells secundari i superior, com als mateixos alumnes, com als traductors i correctors que treballen en aquest camp, com a qualsevol investigador i estudiós que en pugui fer ús.

Pel seu caràcter estrictament tècnic, no aplega sinó els termes cultes propis del llenguatge científic, per la qual cosa hi són pràcticament absents termes corrents a la llengua col·loquial, no havent-se hagut de plantejar, doncs, els autors la necessitat de tria entre diversos possibles sinònims dialectals. Així i tot, en les poques vegades en què veus populars han estat recollides com a termes precisos, els autors han mostrat la seva preferència pels sinònims o per les variants no balears –si n'hi ha hagut–, més concretament per les del català central, baldament aquelles siguin igualment acceptades per la normativa vigent. Així, *capgròs* per *granotet*, *cargol* per *caragol* i *cua* per *coa*. *Gavatx*, però, es posa com a sinònim complementari de *pap*.

## 2. OBRES DE CARÀCTER ENCICLOPÈDIC

*Gran Enciclopèdia de Mallorca*. Volums I - X. Palma, Promomallorca, s/d.

Obra de caràcter purament enciclopèdic, sense pretensions de diccionari de llengua: *«és una obra de caràcter analític i d'àmbit local, ordenada alfabèticament, on s'ha donat entrada a les coses, fets i personatges més rellevants de Mallorca»* (I, 8). Respectant les normes gramaticals i ortogràfiques de l'I. E. C., acull les variants mallorquines tradicionals (per exemple, *ametler*, preferent a *ametller*), exceptuant-ne *bolletí*, substituïda normalment per *butlletí*. En toponímia, es permet la inclusió de formes dialectals no normatives, com *«Avenc de s'Aigo»*, seguint un criteri que consideram discutible.

## CONCLUSIONS

La primera conclusió que en podem treure és que la lexicografia catalana a les Balears ha estat a l'època moderna poc productiva, si n'exceptuam l'obra magna del *Diccionari Català-Valencià-Balear* i el recull monogràfic del *Glossari General Lul·lià*. No hi ha cap diccionari general amb pretensions normatives, i la seva absència ha estat

suplida pels diccionaris bilingües de Francesc de Borja Moll, que durant anys han servit de pauta de correcció per a tots aquells qui es volien mantenir en la tradició literària balear.

Una segona conclusió és que tant les obres d'en Moll com els vocabularis més recents, ja sien terminològics ja sien amb finalitat didàctica (i prescindint dels que pretenen solament aportar algunes dades al coneixement dialectal), mostren una vacil·lació important a l'hora d'incorporar variants i mots o expressions particulars de les Illes, fins i tot en el cas que es tracti de formes explícitament o implícita acceptades per la normativa. Aquesta vacil·lació –deguda, en el cas de Moll, a un excés de permissibilitat respecte a les variants illenques, mancat d'un criteri previ; i deguda, en els altres casos, a un exagerat i malentès respecte per la normativa, sovint identificada amb el català central, i també a un desconeixement real de les pròpies alternatives– és la causa d'incoherències importants en la presentació del contingut, amb la conseqüent desorientació al lector o a l'usuari que això comporta. Només els *Vocabularis temàtics...* i el *(Nou) Vocabulari de barbarismes...* presenten, en aquest sentit, una definició ben clara cap a la recuperació de les maneres de dir autòctones de Mallorca, dins el respecte per la normativa catalana unitària.

Es pot dir que la lexicografia catalana moderna a les Balears espera encara l'obra nova, metodològicament avançada, de plantejaments *nacionals* i sense complexos provincians. En aquest aspecte, sense menysprear l'aportació que per la seva sola presència representen, els *diccionaris*, *vocabularis* o *lèxics* existents no poden complir adequadament la funció d'eina útil per a la recuperació global de la llengua que, en aquests moments, també els és exigible.[6]

JAUME CORBERA
*Universitat de les Illes Balears*

## REFERÈNCIES BIBLIOGRÀFIQUES

BEAULIEU, H.; FARRIOLS, A.; ROSÉS I ALBIOL, F. (1992) «Comparança de tres diccionaris bilingües». A: Solà, J. (1992).

CABRÉ, M. T.; LORENTE, M. (1991) *Els diccionaris catalans de 1940 a 1988*. Barcelona, Universitat de Barcelona.

COLON, G.; SOBERANAS, A.-J. (1986) *Panorama de la lexicografia catalana*. Barcelona, Enciclopèdia Catalana.

(6) He d'agrair al Gabinet de Terminologia de la UIB, centre col·laborador del TERMCAT, la seva ajuda en l'arreplega de totes les obres ressenyades.

CORBERA POU, J. (1993) «La influència de l'italià i dels parlars sards i itàlics en les denominacions algueresesdels animals marins». A: *Actes del Novè Col·loqui Internacional de Llengua i Literatura Catalanes*. Alacant-Elx, 9-14 de setembre de 1991, II, 513-527. Montserrat, P. A. M.

HAENSCH, G. (1982) «Tipología de las obras lexicográficas». A. Haensch, G.; Wolf, L; Ettinger, S.; Werner, R. (1982).

HAENSCH, G.; WOLF, L; ETTINGER, S.; WERNER, R. (1982) *La lexicografía. De la lingüística teórica a la lexicografía práctica*. Madrid, Gredos.

«Lexicografía Catalana» (I). A: *Anthropos* 44, 1984, 22-32.

LORENTE, M. (1993) «Funció lexicogràfica i contacte de llengües». A: *Actes del IX Col·loqui Internacional de Llengua i Literatura Catalanes*. Alacant-Elx, 9-14 de setembre de 1991, III, 269-281. Montserrat, P.A.M.

SOLÀ, J. (ed.) (1992) *Sobre lexicografia catalana actual*. Barcelona, Empúries.

JAUME CORBERA

# LA LEXICOGRAFIA CATALANA A L'ALGUER

Els treballs d'arreplega del lèxic alguerès són escassos i, fins fa pocs anys, ben magres. A finals del XIX o a principis del segle actual hi ha els petits reculls fets per Guarnerio, Ciuffo i Palomba (vegeu-ne les referències a Colon-Soberanas 1986: 184-185), als quals són freqüents els errors de transcripció, d'afiliació etimològica o, fins i tot, de significació. Més endavant tenim els de Griera (1922), més extens però també més erroni, i el de G. Serra (1927), curt però molt fiable. I no és ja fins a una època molt recent que apareix el d'E. Blasco (1984), limitat a l'estudi de 339 conceptes (encara que dins el conjunt de l'obra apareixen molts més mots algueresos), on també trobam algunes equivocacions importants; el de L. Interlandi (1986), molt específic; i els de Francesc Manunta (1988), original però insegur, i Francesc Manunta i Pere Català i Roca (1992), d'utilitat mínima per la manca d'identificació dels noms amb les espècies. Hi ha també els de R. Caria (1990), i els més recents de J. Corbera (1993, 1992) i de R. Caria (1993), tots centrats en uns camps conceptuals restringits. La contribució de J. Bruguera (1978) no afegeix res de nou als treballs anteriors, dels quals assumeix, sense haver-los sotmès a una revisió crítica, els errors que contenien.

Amb defectes o virtuts, però, totes aquestes no són, al capdavall, més que aportacions més o menys interessants o útils per al coneixement del dialecte, destinades sobretot als especialistes. No són diccionaris fets amb tècniques o pretensions lexicogràfiques, i per això no en tractarem en aquesta contribució. Ens ocuparem només d'aquelles obres que es presenten expressament sota l'etiqueta de «vocabulari» o «diccionari» i que, a més a més, utilitzen la presentació típica d'aquestes obres.

*Caplletra 17* (Tardor 1994), pp. 195-200

Dit això, no s'ha de creure tanmateix que la producció lexicogràfica algueresa sigui abundant: just hi ha editats 2 títols que s'ajusten a les nostres pretensions. Hi afegiré, però, una obra inèdita feta per G. Palomba, de molta més envergadura que l'al·ludida al principi, a la qual fan referència, sense conèixer-la directament, Colon-Soberanas (1986: 185).

Giovanni PALOMBA, *Lessico*. Manuscrit a una llibreta de mida de foli, numerat en dues parts. La primera duu el subtítol *Raccolta dei nomi più usati*, i comprèn les pàgines 1 a 144. La segona té el subtítol *Raccolta delle frasi più comuni*, numerada de la pàgina 1 a la 6.[1]

Es tracta d'un vocabulari d'equivalències a tres bandes, alguerès - català - italià, sense definicions ni explicacions complementàries de cap classe, llevat d'en casos excepcionals, en què aquestes apareixen en nota a peu de pàgina (per exemple, l'alguerès *Lus duendus* duu la nota: *«Dal Castigliano duende - Spirito che il volgo crede che vive nella casa»*). Els mots, el gènere dels quals és assenyalat per l'article, són agrupats per camps conceptuals (*«De Deu»*, *«De l'iglesia»*, *«Dignità ecclesiastiche»*, *«Delle feste principali»*, *«Degli elementi»*, *«Degli astri o funzioni della natura»*, etc.). Les paraules algueresses són escrites amb ortografia pròxima a la catalana (tenint en compte que aquesta és «pre-fabriana», si bé no excessivament divergent de les *Normes* de 1913), de la qual dissenteix sobretot per l'addició d'uns diacrítics que volen marcar l'especificitat de pronúncia algueresa: *ė* [a], *ö* [u], *ṛ* [l], *ḷ* [r], *ḍ* [r]; i per l'ús de *ć*, en lloc de *x* o *tx*, i de *ś*, en lloc de *x* o *ix* (bé que no sempre). Així, escriu *Tres pėrsonas* ([tres palsónas]), *Lu tronö* ([lutrònu]), *Lus candelobras* ([luskandaróbras]), *La viuḍansa* ([laviuránsa]), *Un nićö* ([úńnítʃu]; però *la trixa o tricа*), *las cuxas o cuśas*. Omet, a més a més, les *hh* inicials: *oma* («home»), *eroe* («heroi»), *Olanda*. El *Lessico* comprèn unes 3.500 veus, entre les quals hi ha noms de països, de regions d'Itàlia, gentilicis, topònims de l'Alguer i expressions diverses, a més d'unes 150 frases «més usades», i constitueix per això una font molt rica de primer ordre per conèixer la particularitat del parlar alguerès, especialment del vocabulari.

Antonio ARCA, *Lèxic Militant. Vocabolario di base della lingua italiana proposto in catalano-algherese*. L'Alguer, La Celere - Llibreria Antic Taller, 1985.

L'autor pretén donar als algueresos el vocabulari de base en català-alguerès a partir del corresponent italià. Pren aquest vocabulari de la proposta feta per Tullio De Mauro

(1) Una fotocòpia d'aquest manuscrit em fou facilitada pel Sr. Pasqual Mel·lai, de l'Alguer, l'any 1984. Gràcies a ell, doncs, m'és possible ara fer aquesta ressenya de l'obra inèdita de Palomba.

al seu llibre *Guida all'uso delle parole* (Roma, 1980). Són, així, 6.690 paraules teòricament suficients i necessàries per fer-se entendre d'un auditori mitjanament alfabetitzat.

Per traduir els mots italians, l'autor s'ha servit sobretot del *Diccionari Italià-Català Català-Italià* de J. Fornas Prat (Barcelona, 1982), les mancances del qual aquell repeteix i augmenta. Així, per exemple, *abitazione* és traduït per *habitació*, *avere* per *haver/tenir* (com si fossin ben equivalents!), *bandiera* per *senyera*, *domandare* per *preguntar, demanar* (sense cap altra especificació), *dovere* per *haver, deure* [sic], *due* per *dos*, *costume* per *vestit, costum* (!), *ghiaccio* per *glaç*... Per altra banda, contra el que podríem esperar del seu títol, l'alguerès hi apareix només subsidiàriament i quasi folklòricament. A moltíssims d'articles la paraula algueresa diferenciada no hi és, mentre que a altres hi és donada com a forma vulgar grafiada arbitràriament, aparentment indigna de ser incorporada a la norma catalana; pareix com si fos considerada una simple curiositat. En alguns altres casos, però, l'autor no ha sabut trobar traducció catalana a la veu italiana (absent al diccionari de Fornas) i es limita a donar l'italianisme propi de l'alguerès col·loquial i una paràfrasi explicatòria («AUTOBLINDO f. *autoblindo* (vehicle blindat)», «AUTOBOTTE f. *autobotte* (vehicle per al transport d'aigua)», «AUTOREVOLE agg. *autorevol* (qui té ascendent sobre un altre)», «BATTIPANI m. *battipanni* (estri per a netejar tapissos)»...).

Amb l'italià com a metallengua, hem de dir que aquesta obra res aporta a la lexicografia catalana en conjunt, i que és, a més a més, ben inútil en l'empresa de facilitar als algueresos una eina de recuperació de la seva pròpia parla. Més que sentir-s'hi identificats, s'hi senten *ridiculitzats*.

Josep SANNA, *Diccionari Català de l'Alguer*. Amb la col·laboració de Josep Subirats i Emili Pascual. L'Alguer / Barcelona, 1988.

A pesar del retret *avant la lettre* que li fan Colon-Soberanas (loc. cit.), absolutament injust i injustificat, el diccionari de Mn. Sanna no és purament dialectal i localista, sinó que, recollint les particularitats lèxiques, fonètiques i morfològiques algueresses, pretén integrar-les a la normativa catalana. Aquesta tasca no era gens fàcil per a l'autor, sobretot pel fet que moltes de les paraules aplegades són desconegudes en els altres dialectes i calia primer de tot saber-ne l'etimologia –estudi que encara no s'havia fet– per establir-ne l'ortografia més apropiada.

Per afrontar millor les dificultats previsibles a l'hora de concretar la feina, l'autor i els seus col·laboradors varen proposar uns «Criteris pràctics en la preparació del Diccionari», als quals, però, trobam una certa confusió, especialment en considerar fenòmens equivalents allò que són particularitats dialectals genuïnes algueresses i allò que són influències sardes o, sobretot, italianes. Partint, doncs, d'aquesta confusió inicial

i fonamental (que sembla agreujada pel prejudici que tot el que no es diu igual que en català central no pot tenir una forma gràfica pròpia i acordada amb els principis de l'ortografia catalana), la «normalització» resultant de l'ortografia algueresa és incoherent, desconcertant i, en molts de casos, *falsa*: reflecteix una forma inexistent a la realitat.

Un cas molt clar n'és el d'aquelles paraules sorgides d'un ètim en T'L o D'L. La solució inicial algueresa, com la balear i la sud-occidental, és [ll] (>[l]), pronúncia que té una solució gràfica perfectament establerta per la tradició literària: TL. Però com que aquesta solució, tot i havent estat recollida per Francesc de B. Moll i Manuel Sanchis, no constava explícitament al *Diccionari General de la Llengua Catalana* en el moment de redactar l'obra de què tractam, hom no va saber què fer-ne, de la pronúncia algueresa, arribant fins a escriure-la de tres maneres diferents: **espatla** (la grafia adequada), **mela** (grafia impròpia per a «**(a)metla**») i **motlle** (grafia adequada als parlars nòrdics peninsulars).

En alguns altres casos no s'aprecia la validesa de les formes algueresses, modificades innecessàriament segons la norma central: **dipòsit** (per «**depòsit**»)[2] i els seus derivats, **ordre** i **desordre** (per «**orde**», «**desorde**»), **butlletí** (per «**bolletí**»).

Per una altra banda, davant fenòmens com les metàtesis, freqüents en alguerès, però conegudes també en tots els dialectes, la solució adoptada és irregular: s'hi prefereix **craba** a **cabra**, també inclosa; s'hi troba exclusivament **aquidrar**, en lloc de **(a)cridar**, però s'hi rebutgen **dromir**, **prebe** o **probe**. I és inexplicable que paraules catalanes vives igualment en altres dialectes, l'ortografia de les quals no hauria d'oferir dubtes, s'hi trobin escrites en grafia dialectalitzant: **càntara** (per «**càntera**»), **palalla** (per «**palaia**» o –millor– «**pelaia**»), **velmarí** (per «**vellmarí**»)...

Però allà on la feina oferia, sense dubtes, més dificultats era en l'adaptació dels nombrosos forasterismes entrats en l'alguerès, provinents sobretot dels parlars sards veïnats, dels dialectes itàlics més pròxims i de l'italià oficial; i el fet és que davant tota aquesta munió de paraules originalment no catalanes no hi ha una solució regular i coherent per a tots els casos. Respecte als italianismes, alguns hi són amb la grafia original (a vegades, amb l'addició de qualque accent), altres hi són parcialment catalanitzats i altres hi són catalanitzats (gràficament) del tot. Qualcun té una grafia inexplicable (**faige** < FAGGIO). Quant als mots de procedència sarda o italiana dialectal, el desconcert és semblant. D'un costat, als acabats originalment en **-u** se'ls ha volgut donar aparença catalana canviant-ne la vocal final per **-o**, canvi injustificat que no respon a cap pronúncia (l'alguerès els fa en **-u**) i desfigura innecessàriament l'origen: **crabofigo** (< logudorès KRABUFIGU), **pàlago** (< cors PARAGU), **piguraro** (< logudorès PIUGU LADU), **bromo** (< sicilià BROMU)... D'un altre costat, davant el desconeixement, en la majoria de casos, de l'ètim adoptat per l'alguerès, en lloc d'obrar amb prudència i escriure el mot de la manera més aproximada a la pronúncia, les grafies s'han establert per pura intuïció, amb resultats tan arbitraris com **donademulo** (per

(2) Recordem que aquesta, més etimològica, havia estat acceptada a la primera edició del *Diccionari General de la Llengua Catalana*, i que en va ser suprimida a la segona per decisió unipersonal del Sr. Ramon Aramon. Vegeu-ho al DECLC, de Joan Coromines, s.v. PONDRE.

«**dona de muru**», = logudorès DONNA DE MURU), **capdullo** (per «**cabodullo**», < napolità CAPODUOGLIO), **bodroni** (per «**budroni**», < logudorès BUDRONE), **bocó** (per «**bucó**», < campidanès BUKKONI), **catarà** (per «**català**», = logudorès KADELANU), **fromígola** (per «**frumígula**», < logudorès FRUMMIGULA), etc.

Un diccionari, però, no és només la forma, sinó també el contingut. I en aquest aspecte, cal reconèixer que la feina de Josep Sanna té un gran valor. Una replega de prop de 25.000 mots usats en l'alguerès contemporani és, sense discussió, un tresor lingüístic incomparable. A més a més, el recull és presentat amb tots els requisits perquè sigui útil: la forma gràfica, la categoria gramatical, l'equivalència en català normatiu, la pronúncia, la definició (a vegades, potser massa simple), el nom científic (quan n'hi correspon; però n'hi ha qualcuns d'errats!) i els exemples necessaris.

En conclusió, es tracta, en conjunt, d'una obra desigual. És insegura en la presentació del corpus lexical, per tot el que hem vist i perquè no acaba d'adoptar una actitud decididament reivindicativa de les formes genuïnes enfront de les forasteres (admetent, per exemple, els adjectius en **-bil** o en **-ivo**: **improbàbil**, **dubitativo**). Però és, també, una obra rica i ben estructurada, amb una quantitat extraordinària d'informació sobre el parlar alguerès.

En definitiva, pot perfectament complir l'objectiu per al qual va ser confeccionada: *«ajudar los algueresos en aquesta batalla per la sobrevivència contra el temps i contra les coses»*, en paraules del mateix malaguanyat Mn. Josep Sanna i Useri, desgraciadament traspassat ben poques setmanes més tard després que el *Diccionari* vés la llum.

JAUME CORBERA
*Universitat de les Illes Balears*

## REFERÈNCIES BIBLIOGRÀFIQUES

BLASCO FERRER, E. (1984) *Grammatica Storica del Catalano e dei suoi dialetti con speciale riguardo all'Algherese*. Tübingen, Gunter Narr.

BRUGUERA, J. (1978) «El català a Sardenya». Dins *Nationalia, III. Segones Jornades del CIEMEN. Còrsega i Sardenya per les reivindicacions nacionals*, 95-142. Publicacions de l'Abadia de Montserrat.

CARIA, R. (1990) *Il mondo del Càlic - Studi di toponomastica e lessicografia algherese*. Sàsser.

CARIA, R. (1993) «La vegetació espontània del camp de l'Alguer. Aproximació a l'estudi onomàstic». A: *Miscel·lània Jordi Carbonell*, 5, 205-229. Publicacions de l'Abadia de Montserrat.

COLON, G.; SOBERANAS, A.-J. (1986) *Panorama de la lexicografia catalana*. Barcelona, Enciclopèdia Catalana.

CORBERA POU, J. (1993) «La influència de l'italià i dels parlars sards i itàlics en les denominacions algueres dels animals marins». A: *Actes del IX Col·loqui Internacional de Llengua i Literatura Catalanes*. Alacant-Elx, 9-14 de setembre de 1991, II, 513-527, Montserrat, P.A.M.

— (1992) «Els noms dels ocells terrestres en alguerès». A: *Miscel·lània Jordi Carbonell*, 4, 265-278. Publicacions de l'Abadia de Montserrat.

GRIERA, A. (1922) «Els elements sards en el català d'Alguer». A: *Butlletí de Dialectologia Catalana*, X, 133-139.

INTERLANDI, L. (1986) «L'estructuració del lèxic mariner alguerès». A: *Actes del VII Col·loqui Internacional de Llengua i Literatura Catalanes*, 511-522. Publicacions de l'Abadia de Montserrat.

MANUNTA, F. (1988) *Cançons i líriques religioses de l'Alguer Catalana*. L'Alguer, La Celere.

MANUNTA, F.; CATALÀ I ROCA, P. (1992) «Noms de peixos en alguerès». A: *Segon Congrés Internacional de la Llengua Catalana*. IV. Àrea 3. Lingüística Social, 97-107. Palma, U.I.B.

SERRA, G. (1927) «Aggiunte e rettifiche algheresi all'Atlas Lingüístic de Catalunya (cc. 1/586)». A: *L'Italia dialettale*, III, 197-216.

M. T. CABRÉ / M. BOZZO / A. EGEA
M. JUNCADELLA / C. SANTAMARIA

# EL *NOU DICCIONARI NORMATIU DE LA LLENGUA CATALANA*

## 1. EL PROJECTE

### 1.1. Justificació

La llengua catalana ha tingut fins ara en el *Diccionari general de la llengua catalana* (DGLC) de Pompeu Fabra el diccionari de referència normativa que necessita qualsevol llengua de cultura normalitzada per orientar l'ús correcte que de la llengua han de fer en determinades situacions els seus parlants.

Els esdeveniments històrics i polítics produïts al llarg del segle XX, però, han fet que aquest diccionari no seguís el procés d'evolució que segueixen els diccionaris normatius arreu del món. Efectivament, la primera edició del DGLC va aparèixer el 1932 per iniciativa del mateix Fabra, aleshores president de la Secció Filològica de l'Institut d'Estudis Catalans, per fornir els parlants d'una eina de referència, mentre no fos publicat el gran *Diccionari* de l'Institut d'Estudis Catalans.[1]

I van transcórrer vint-i-dos anys abans que no es pogués fer la segona edició del DGLC (1954), que Carles Riba, president en aquell moment de la Secció Filològica i autor del pròleg, presenta com una reedició, ja que els canvis respecte a la primera edició són mínims i fruit, moltes vegades, de la intervenció de la censura del moment.

(1) La provisionalitat d'aquesta obra l'explicita Pompeu Fabra en el pròleg del *Diccionari*, on el presenta com a «canemàs del futur diccionari de l'Institut».

Del 1939, any en el qual l'IEC fou desposseït del seu patrimoni, fins que va ser novament reconegut oficialment l'any 1976, l'Institut d'Estudis Catalans va viure en semiclandestinitat i va seguir les seves tasques amb precarietat gràcies a la voluntat investigadora dels seus membres i a les aportacions econòmiques de ciutadans i institucions privades. Aquest sistema de vida i de funcionament va repercutir, evidentment, en les edicions del DGLC a partir dels anys seixanta, de les quals s'encarregà Josep Miracle, que va introduir al text de Fabra força noves entrades i modificacions de les entrades existents, aprovades per la Secció Filològica, i addicions i modificacions recollides personalment. Miracle, però, per respecte a l'obra de Fabra i a l'autoritat de la Secció Filològica, va marcar explícitament en el text tot allò que era fruit de la seva aportació personal.

En la sessió plenària d'octubre de 1990, la Secció Filològica, conscient de la seva responsabilitat en la normativització de la llengua catalana i sabent la importància que té per als parlants de disposar d'un diccionari normatiu que respongui a les necessitats de la llengua moderna, va decidir d'abocar els seus esforços i els de les Oficines Lexicogràfiques en l'elaboració del *Nou diccionari normatiu*, sobre la base que havia de ser la revisió i l'actualització del diccionari de Fabra i amb la intenció d'anar-lo mantenint al dia en futures edicions. L'aprovació de la Llei sobre l'autoritat lingüística de l'IEC per part del Parlament de Catalunya el 24 d'abril de 1991 no va fer més que ratificar políticament, si calia, l'impuls del projecte d'elaboració del *Nou diccionari normatiu de la llengua catalana* (NDNLC). Aquest projecte, que inicialment s'havia de dur a terme en un període de sis anys, s'ha hagut d'accelerar i readaptar a un nou calendari de treball, d'acord amb els compromisos del mateix Institut.

El projecte d'elaboració del NDNLC parteix de sis supòsits fonamentals:

- És el diccionari de l'institut d'Estudis Catalans, i doncs, ha de ser assumit pel Ple de la Institució, tant en el sentit de la seva acceptació, com en el de la seva elaboració. Es planteja, per tant, com un treball col·lectiu de l'Institut (i no com un treball d'una secció concreta) i plenament institucional.
- Serà considerat, quan s'editarà, l'obra de referència normativa fins que l'Institut no disposi d'un nou diccionari (el gran *Diccionari* de l'Institut), que s'haurà d'elaborar a partir dels materials que es constitueixen en el marc del projecte del Corpus Textual del Català Contemporani.
- És una actualització i regularització del *Diccionari general* de Pompeu Fabra, i, per tant, no s'elabora a partir de zero, sinó sobre la base del «canemàs» del *Diccionari general*, i es fonamenta, sempre que això respon a les exigències de la lexicografia contemporània, en els criteris que es van aplicar durant els anys trenta en la confecció del DGLC.
- Però no obstant aquesta fidelitat, ha d'acoblar-se a les necessitats expressives i comunicatives del món d'avui i, per tant, ha d'actualitzar-ne la nomenclatura i el contingut d'acord amb l'estat actual de la llengua, i, si cal, ha de posar el dia els criteris de presentació dels articles.

- És diccionari normatiu,[2] i, per aquesta raó, només contindrà els mots de la llengua que l'Institut considera plenament correctes i identificats com a mots catalans. No pot contenir, doncs, mots que, tot i ésser correntment usats pels parlants, no es considerin ni catalans[3] ni correctes.
- Ha de respectar el concepte que una llengua correcta no és una llengua monosistemàtica, sinó plural, tant pel que fa a la seva extensió com a la diversitat dels seus registres.[4]

## 1.2. Organització del treball

L'organització dels treballs del NDNLC ha estat (i continua essent), atesos els seus supòsits fonamentals, d'una gran complexitat, tant pel que fa referència a la distribució i assumpció dels àmbits de responsabilitat com a l'organització mateixa de les tasques, si tenim presents el calendari d'acceleració fixat per l'IEC i el nombre de col·laboradors que hi intervenen.

La Secció Filològica va encomanar a Joan Bastardas, aleshores director de les Oficines Lexicogràfiques i president de la Comissió Lexicogràfica de la Secció Filològica, la direcció del projecte, i a M. Teresa Cabré, actualment directora de les Oficines Lexicogràfiques, la seva coordinació.

En aquesta línia, les tasques i responsabilitats en l'elaboració del diccionari s'han distribuït ordenadament entre les diferents unitats i òrgans que componen l'Institut d'Estudis Catalans:

- Les Oficines Lexicogràfiques es responsabilitzen de l'organització general dels treballs i de la preparació dels materials per al diccionari.
- La Comissió Lexicogràfica, composta per membres de la Secció Filològica especialistes en lèxic, té la responsabilitat de revisar els materials de les Oficines, d'aprovar les formes proposades i els criteris d'elaboració del diccionari, i de presentar els treballs a la Secció Filològica.
- La Secció Filològica, el Ple de la qual és qui té l'última paraula sobre els criteris que cal seguir en l'elaboració del NDNLC, dóna el vist-i-plau al text del diccionari.
- Les seccions de l'IEC i les societats filials adscrites, a les quals pertanyen científics de les diverses àrees del saber, són responsables de la selecció de la terminologia especialitzada de cada àmbit temàtic i de l'atribució de les definicions que hi corresponen.
- Finalment, és el Ple de l'Institut d'Estudis Catalans qui ha d'aprovar el *Nou diccionari normatiu de la llengua catalana*.

(2) Un diccionari és normatiu si el conjunt dels mots que conté està integrat pels mots considerats com a correctes per la norma (en el cas de la llengua catalana, per la norma acadèmica). Un diccionari és descriptiu si recull la totalitat dels mots usats efectivament pels parlants, amb independència de la seva correcció. La majoria de diccionaris generals que no han estat elaborats per una institució normativa es troben a cavall entre un i altre tipus: inclouen el conjunt dels mots acadèmicament correctes, però n'integren d'altres d'una freqüència d'ús molt alta i d'abast geogràfic generalitzat. Però, això no obstant, no perden de vista una funció orientadora de l'ús correcte.

(3) Cal precisar que un mot manllevat d'una altra llengua també pot ser considerat català, si el sistema l'ha assumit com a propi i amb independència de la seva adaptació gràfica o formal. Un mot com *espagueti* és avui dia un mot català, en aquest cas adaptat gràficament (incorporació de la *e* inicial, simplificació de la doble *t*) i morfològicament (el seu plural és *espaguetis*); però si el diccionari normatiu conté *allegretto*, tot i no haver estat plenament adaptat en la forma gràfica, significa que també és un mot correcte.

(4) Cal no perdre de vista que una llengua amb una única opció expressiva està destinada, a la llarga, a ser substituïda per una altra llengua concurrent que abasti la diversitat funcional de necessitats comunicatives i expressives.

### 1.3. Procés de treball

Posar al dia un diccionari (i sobretot un diccionari que no ha estat revisat periòdicament) respectant els supòsits que s'han explicitat més amunt, requereix una organització complexa de feines i funcions, que en el cas del NDNLC s'han centrat en dos punts:

- La incorporació al diccionari de paraules generades pels diversos usos socials i pels avenços de la ciència i la tècnica.
- L'actualització dels continguts (definicions, exemples, accepcions i criteris lexicogràfics) del DGLC i l'elaboració dels articles corresponents a les paraules noves.

En aquesta perspectiva, mirant alhora de rendibilitzar aquest treball per a utilitzacions posteriors de caràcter divers, s'ha distribuït el treball en tres etapes.

En una primera etapa s'ha inclòs la **constitució d'una base de dades general informatitzada**[5] **(BDOL)**[6] que conté, a més del DGLC, paraules de lèxic comú i de lèxics *marcats*[7] recollides de fonts diverses (diccionaris, vocabularis especialitzats, reculls dialectològics, treballs inèdits, etc.).[8]

Paral·lelament a aquesta tasca, també s'ha inclòs en la primera etapa un estudi del *Diccionari general* de Pompeu Fabra amb dues finalitats: *a)* l'anàlisi del text en ell mateix, marcant els continguts que avui dia poden ser considerats obsolets (i, alhora, detectant possibles errades del diccionari); i *b)* l'anàlisi dels criteris implícits amb què Fabra va elaborar el *Diccionari general.*

La segona etapa consisteix en la **confecció del *Nou diccionari normatiu de la llengua catalana.*** En aquesta etapa es fixen els criteris lexicogràfics d'elaboració del diccionari, a partir de l'estudi del DGLC a què acabem de fer referència i d'estudis de qüestions lingüístiques conflictives; se seleccionen les entrades, les subentrades i les altres informacions pertinents que han de figurar en el diccionari;[9] es revisen els articles antics, i s'elaboren articles nous.

La tercera etapa comporta la **creació de la base de dades específica del NDNLC (NOUDIC)**, un cop ha finalitzat el procés d'elaboració de tots els materials. Aquesta base de dades conté el material complet i organitzat en articles del *Nou diccionari normatiu de la llengua catalana.*

## 2. LA BASE DE DADES GENERAL (BDOL)

### 2.1. Constitució

La constitució de la *base de dades de les Oficines Lexicogràfiques* (BDOL) s'ha dut a terme amb l'objectiu principal de disposar d'un conjunt de materials lexicogràfics

(5) La base de dades constituïda, si és actualitzada permanentment, permetrà que la Secció Filològica disposi de materials prou importants per fer descripcions acurades i justificades de determinats aspectes del lèxic de la llengua.

(6) BDOL significa 'base de dades de les Oficines Lexicogràfiques'.

(7) Es considera *lèxic marcat* tots els subsistemes de mots caracteritzats dialectalment o funcionalment per alguna especificitat restrictiva: l'àmbit geogràfic en què s'usen, el seu caràcter històric, la pertinença a un determinat registre temàtic (la terminologia) o formal (el grau de formalitat quant als nivells expressius: formal, informal, vulgar, etc.).

(8) Els criteris que han determinat la selecció dels diccionaris font han estat purament quantitatius: han entrat a BDOL els diccionaris que contenen un 15 % d'informació *diferent* en relació amb la que conté el DGLC.

(9) En aquest punt cal fer esment que la responsabilitat de la selecció d'entrades noves està distribuïda entre la Secció Filològica i la resta de seccions de l'Institut. La Secció Filològica es fa càrrec de la selecció d'entrades noves de lèxic comú, de dialectalismes i de termes argòtics. Les altres seccions de l'IEC (Secció Històrico-Arqueològica, Secció de Filosofia i Ciències Socials, Secció de Ciències i Tecnologia, i Secció de Ciències Biològiques) s'ocupen de la selecció de la terminologia especialitzada i de l'atribució de definicions als termes seleccionats.

en un format que permeti, d'una banda, realitzar tota mena d'estudis lèxics i, de l'altra, fornir dades per a qualsevol projecte lexicogràfic.

Fins ara ha servit per establir els criteris en què es basa el DGLC i els que s'han aplicat en l'elaboració del nou diccionari, i, d'altra banda, ha permès de disposar de dades per a la redacció del NDNLC.

### 2.2. Estructura

Els camps en què s'estructura cada registre d'aquesta base de dades lexicogràfica es poden dividir en quatre blocs ben diferenciats segons el tipus d'informació que contenen:

- En la part del registre, corresponent a la **identificació** del mot, hi consten l'*entrada*, que és el mot de referència per a totes les fonts; un *codi numèric* d'identificació del mot, que és assignat automàticament pel programa informàtic, i un seguit de *dades complementàries* comunes a totes les fonts.
- En la part corresponent a la **informació procedent de les fonts**, hi figuren diversos camps, que inclouen informació específica procedent d'una font. En aquests camps s'hi especifiquen: la *font* de procedència de les informacions i totes les *informacions* contingudes en aquesta font.
- La tercera part, que correspon a les **propostes lèxiques**, reprodueix en forma de camps l'estructura dels articles del DGLC, i conté tota la informació que sobre un mot proposen les Oficines Lexicogràfiques per al NDNLC. Aquesta proposta pot coincidir amb les informacions que consten al DGLC, però també pot ser parcialment o totalment nova respecte al DGLC, quan es tracta d'una entrada nova o de la modificació d'un article del diccionari Fabra.
- La darrera part de cada registre dóna raó de la **gestió** de les dades, que indica l'etapa d'elaboració en què es troba cada fitxa lexicogràfica (definitiva / provisional / pendent d'estudi, etc.) i les etapes per les quals ha passat.

### 2.3. Consultes

L'estructura de la base de dades dóna la possibilitat de recuperar selectivament les informacions introduïdes, la qual cosa ha servit per detectar errors i incoherències de tota mena que hi pogués haver en el DGLC, per elaborar criteris lexicogràfics, per homogeneïtzar entrades de característiques semblants, etc.

La recuperació de dades de BDOL pot tenir dues formes: en línia, a través dels terminals de l'ordinador, o bé en suport imprès en forma de llistats. La selecció d'un o altre format depèn del tipus de consulta que es vol fer, del temps que s'inverteix en el procés i de la quantitat de dades que es pretén recuperar.

## 3. L'ELABORACIÓ DEL NDNLC

L'elaboració pròpiament dita del *Nou diccionari normatiu de la llengua catalana* ha estat distribuïda en cinc fases, tal com sintetitza el quadre annex al final d'aquest article. Cada fase implica un conjunt de treballs i un seguit de decisions que confluïxen al final de tot el procés i que constitueixen un pas important en la confecció del text del diccionari. En aquestes fases es tracten els temes següents: l'elaboració de criteris lexicogràfics, l'actualització de la nomenclatura del DGLC, l'elaboració d'articles, l'aprovació del text del diccionari i la publicació de l'obra.

### 3.1. Fase primera: elaboració de criteris lexicogràfics

Aquesta fase, que culmina amb la fixació dels criteris lexicogràfics que s'han de seguir en l'elaboració del *nou diccionari*, està estretament lligada, com ja s'ha dit, a l'estudi del DGLC, perquè l'elaboració del *Nou diccionari normatiu de la llengua catalana* s'ha de basar, sempre que sigui possible i racional, en els criteris utilitzats en el *Diccionari general de la llengua catalana* de Pompeu Fabra. Per aquest motiu, un estudi aprofundit del DGLC permet:

- Explicitar els criteris d'elaboració implícits en el DGLC.
- Determinar quins d'aquests criteris han estat modificats posteriorment per la Secció Filològica.
- Decidir en quins casos cal seguir el criteri de Fabra, en quins altres el de la Secció Filològica, o bé si cal refusar el criteri inferit i establir-ne un de nou.
- Establir criteris, en els casos en què no s'ha pogut detectar cap sistematicitat en el *Diccionari general*.

L'estudi dels criteris de Fabra, que és el fonament d'aquesta primera fase, comença amb l'anàlisi de totes les informacions que pot contenir un article del DGLC, encara que no apareguin de manera sistemàtica en cada un dels articles. També explicita com es representen determinades informacions i en quin punt de l'article se situen.

A més de l'estudi sistemàtic, es fa una lectura del text del DGLC per detectar els errors que pot contenir i una correcció posterior dels errors detectats.

Es pot sistematitzar l'estudi de les informacions que dóna el DGLC en tres apartats, que fan referència als tres eixos d'una obra lexicogràfica: les entrades, els articles i l'estructura general.

#### 3.1.1. *Entrades*

L'estudi de les entrades abasta les entrades pròpiament dites, que encapçalen l'article i que en el DGLC estan escrites en negreta, i les formes alternatives, que gene-

ralment segueixen l'entrada i tenen representacions tipogràfiques diferents (amb lletra negreta separades de l'entrada per una coma o amb lletra cursiva i entre claudàtors).

En l'estudi de les entrades es tenen presents dos tipus d'informacions:

D'una banda, informacions que fan referència al segment entrada i tracten del tipus d'entrades (paraules, locucions, afixos), de la seva representació gràfica (qüestions ortogràfiques, formes alternatives) i dels aspectes morfològics (formes flexionades, forma dels afixos).

D'altra banda, informacions que fan referència a les relacions entre les entrades, i que, per tant, serveixen per estudiar les entrades com a integrants de la nomenclatura del diccionari (proporció d'entrades de lèxics marcats, tractaments dels homògrafs, etc.).

A partir dels estudis elaborats s'estableixen els criteris següents:

- Criteris d'inclusió d'entrades procedents de lèxic comú, de les diverses llengües d'especialitat, dels diferents dialectes geogràfics, i d'inclusió de termes d'argot i de manlleus.
- Criteris d'inclusió de noms propis, sigles, acrònims, marques registrades, locucions, i també prefixos, sufixos i formes prefixades i sufixades.
- Criteris d'inclusió de derivats aspectius i de derivats adverbials en *-ment*.
- Criteris d'inclusió d'entrades de noms i adjectius amb una flexió diferent de la general: entrades en plural, entrada d'una sola forma (masculina o femenina) en els adjectius, etc.
- Criteris per determinar la conveniència d'establir entrades diferents en el cas d'homògrafs, formes alternatives, formes femenines d'alguns substantius, formes del participi dels verbs, etc.
- Criteris per fixar la grafia dels mots compostos, dels manlleus, dels cultismes i d'altres casos vacil·lants.
- Criteris de representació de les formes flexionades regulars en els noms i els adjectius.

### 3.1.2. *Articles*

L'estudi dels articles abasta totes les informacions que conté o pot contenir un article del DGLC després de l'entrada. Les informacions sobre el contingut, que són molt diverses, es poden agrupar en quatre blocs específics: el primer fa referència als aspectes complementaris de l'entrada, el segon bloc tracta del contingut de les accepcions, el tercer estudia les subentrades i el quart abasta les relacions entre els continguts d'articles diferents.

*a.* **Informacions complementàries a l'entrada**: Inclou l'estudi de les informacions sobre la procedència o la forma de l'entrada (informacions etimològiques i morfològiques). Aquests estudis permeten establir els criteris següents:

- Criteris d'inclusió d'informació sobre l'etimologia d'una entrada, sobre la inclusió d'informacions adjuntes a l'etimologia: llengua d'origen, forma gràfica i pronúncia del mot original.
- Relació de les llengües d'origen que han de constar en el *Nou diccionari*.
- Criteris d'inclusió i de representació de formes irregulars de plurals, en els noms, i de formes flexionades, en els verbs.

*b.* **Informacions de les accepcions**: Aquest bloc d'estudis sobre l'article se centra en les informacions que contenen cada una de les accepcions i subentrades de l'article. Cada article conté una accepció o diverses accepcions, i tant en un cas com en l'altre cal veure les informacions de cada accepció. Aquestes informacions poden ser sintàctiques, semàntiques i pragmàtiques:

- Les **informacions sintàctiques** fan referència a la categoria gramatical general i a la subcategorització dels mots. L'estudi del DGLC serveix per determinar les categories gramaticals que apareixen en els articles, com i en quins llocs de l'article es representen, i quantes accepcions afecten.
- Les **informacions semàntiques** apareixen en les definicions i en algunes marques valoratives, representades en el DGLC per abreviatures com ara *fig.*, *per ext.*, etc., que contenen els articles. De les definicions convé analitzar-ne els tipus en relació amb els diversos grups morfològics i amb els grups semàntics, tenint en compte el nombre d'accepcions dels articles.
- Les **informacions pragmàtiques** inclouen les marques d'ús, representades al DGLC per abreviatures com ara *pop.*, *fam.*, *vulg.*, etc., i els exemples. El seu estudi fa palès l'abast dins l'article de les marques d'ús, l'esquema dels exemples i el tipus d'entrada que inclou exemples.

Els estudis que acabem d'esmentar, com en el cas de les entrades, serveixen per elaborar un seguit de criteris per al NDNLC:

- Relació de les categories gramaticals i de les possibles combinacions de categories per al *Nou diccionari*, amb les abreviatures corresponents.
- Criteris sobre l'obligatorietat d'informació de les categories gramaticals i el seu abast, és a dir, quan una categoria afecta el conjunt de l'entrada, quan s'aplica a una o més d'una accepció concreta i quan s'aplica només a un exemple o a una subentrada concreta.
- Criteris per determinar en quins casos el subjecte i els complements del verb definit formen part de la definició i com es representen.
- Relació de les valoracions semàntiques que s'han de fer servir en el nou diccionari, amb les abreviatures corresponents.
- Criteris d'inclusió de valoracions semàntiques i l'abast que han de tenir dins l'article.
- Criteris referents a la definició, que en determinen els trets bàsics i els trets co-

muns a diversos grups d'entrades, atenent a la categoria gramatical, als marcadors morfològics de l'entrada i als camps semàntics.
- Relació de les marques d'ús que s'han d'emprar en el diccionari, amb les abreviatures corresponents.
- Criteris d'inclusió de marques d'ús i de l'abast d'aquestes marques dins de l'article.
- Criteris per determinar què s'entén per exemple en el NDNLC.
- Criteris d'inclusió d'exemples en una accepció i del tractament d'exemples inadequats (obsolets, ofensius per a algun grup social, etc.) que hi pugui haver en el DGLC.

*c.* **Subentrades**: Les subentrades són les informacions lèxiques que contenen els articles i que, juntament amb les entrades, formen la nomenclatura del diccionari. Com que el diccionari no dóna com a entrades els sintagmes travats, els articles contenen en el seu interior la part de la nomenclatura del diccionari que correspon als sintagmes.

L'estudi de les subentrades inclou l'establiment de la tipologia de les expressions (locucions, sintagmes i frases fetes) i la detecció dels criteris d'inserció d'una subentrada en un article.

D'aquests estudis, se'n desprenen els criteris següents:
- Criteris per determinar què s'entén per subentrada en el NDNLC.
- Criteris d'inclusió de subentrades referits a l'entrada a la qual van lligades i sobre la seva ordenació dins l'article.
- Criteris sobre les informacions adjuntes opcionals i obligatòries de les subentrades.

*d.* **Remissions**: Els articles del diccionari es relacionen entre si a través de les remissions, i, d'aquesta manera, palesen les seves relacions sinonímiques. L'estudi d'aquestes relacions permet establir els criteris pels quals Fabra suposadament prioritzà una entrada per sobre d'una altra i controlar si són aplicats sistemàticament en el DGLC. Permet, també, d'establir els criteris per determinar l'ús de les remissions i la seva representació en el *Nou diccionari normatiu.*

### 3.1.3. *Estructura*

En l'estudi del DGLC cal buscar, també, els criteris referents a l'estructura:
- L'estudi de la macroestructura serveix per analitzar l'ordenació de les entrades i els aspectes tipogràfics que afecten el conjunt del diccionari.
- L'estudi de la microestructura del diccionari serveix per fer un inventari de les informacions que pot contenir un article, analitzar-ne l'opcionalitat, l'ordre d'aparició i la representació tipogràfica.

A partir dels criteris detectats en l'estudi s'estableixen els criteris següents:
- Criteris sobre les informacions que ha de contenir una accepció, sobre la situació, l'ordre i les relacions que mantenen entre si les diverses accepcions.
- Criteris d'ordenació de les diverses accepcions (semàntics, cronològics, gramaticals, etc.) i d'ordre de prioritats entre aquests diversos criteris.
- Criteris sobre l'ús dels diferents tipus de separadors d'accepcions (barra, doble barra i punt i a part).

### 3.2. Fase segona: actualització de la nomenclatura

La constitució de qualsevol base de dades lexicogràfica, un cop confeccionats els programes que han de sostenir-la, requereix de seleccionar les fonts que, seguint criteris explícits, han de fornir els futurs diccionaris d'entrades lèxiques. Un cop fixada la selecció, cal preparar les dades convenientment, introduir-les a la base de dades seguint un protocol rígidament sistemàtic i, finalment, verificar-les per garantir-ne la fidelitat a les fonts.

Així, de cara a l'elaboració de la nomenclatura del *Nou diccionari normatiu*, es procedeix a una selecció de les entrades que constitueixen la base de dades lexicogràfica (que provenen tant del DGLC com d'altres fonts seleccionades) i a l'aprovació de la selecció feta.

#### 3.2.1. *Fonts de BDOL*

Les fonts que es tenen en compte a l'hora de constituir BDOL són de diversos tipus: algunes són de lèxic general i altres de lèxics marcats o específics. Els tipus de fonts més rellevants han estat els següents:

*a.* **Diccionaris de llengua**: Tenint en compte que l'objectiu principal d'aquesta base de dades és l'actualització del DGLC, és lògic que la font principal sigui el mateix DGLC (2a edició), completat amb les addicions, les modificacions i les supressions tant d'entrades com d'accepcions aprovades per la Secció Filològica i publicades en els *Documents de la Secció Filològica, I* i *II*.
Una font d'importància destacada és el *Diccionari de la llengua catalana* d'Enciclopèdia Catalana, per tal com ha estat avalat per l'IEC i ha pogut fer, així, una funció de suplència del diccionari normatiu actualitzat.
D'entre diversos diccionaris monolingües i plurilingües, es seleccionen els que presenten diferències superiors al 15 % respecte al DGLC. Les diferències poden consistir en addicions o modificacions tant d'entrades com d'accepcions.

*b.* **Diccionaris especialitzats, vocabularis i reculls**: En la selecció de les fonts de terminologia i d'argot, a més dels diccionaris de lèxic general, es tenen en

compte diversos diccionaris especialitzats, els mots aprovats pel Consell Supervisor del Termcat, altres vocabularis, glossaris d'obres generals, articles diversos i reculls publicats en revistes.

Per a la terminologia, abans d'escollir els vocabularis especialitzats que serveixen per fornir la base de dades, s'estableixen setanta àrees temàtiques, sobre la base de l'anàlisi de les àrees fixades pel Termcat, la *Gran enciclopèdia catalana*, la classificació decimal universal per a biblioteques i el nomenclàtor temàtic de la UNESCO i els suggeriments de les seccions de l'IEC.

*c.* **Materials inèdits**: Una font important d'informació són materials inèdits de procedències diverses, alguns dels quals són sol·licitats per les Oficines Lexicogràfiques i d'altres corresponen a l'oferiment espontani de les persones que els han elaborat.

Dels materials no publicats cal destacar principalment:

- Un primer bloc constituït per neologismes de tres procedències: *a*) la premsa (neologismes de l'Observatori de Neologia),[10] *b*) obres literàries (recull de neologismes literaris elaborat per Júlia Butiñà a partir d'un buidatge de «Les Millors Obres de la Literatura Catalana») i *c*) les aportacions fetes pel Servei de Correcció de l'IEC.
- Un segon bloc constituït pel material fornit pels mitjans de comunicació (inventaris procedents de serveis lingüístics de publicacions periòdiques i d'emissores de ràdio i televisió).
- Un tercer bloc que afecta la terminologia i que està integrat per inventaris d'algunes àrees temàtiques que el Termcat proporciona a les Oficines partint dels materials que conté la seva base de dades.

*d.* **Col·laboracions personals**: Una part del material de la base de dades sorgeix de col·laboracions personals, entre les quals es poden destacar les següents:

- Llistes elaborades per membres de la Secció Filològica, parlants d'alguna de les variants diatòpiques del català, que contenen mots representatius dels diferents àmbits dialectals.
- Propostes de noves entrades i noves accepcions i de correcció d'errors detectats en el DGLC, fets per persones molt diverses al llarg dels anys i recollits per les Oficines Lexicogràfiques.

Un cop les fonts són seleccionades, es fa un treball de preparació per a la introducció de les informacions a BDOL. Es segueixen dos procediments diferents per introduir les dades:

- L'escanneització de textos serveix per introduir el text del DGLC, font principal de BDOL. S'introdueix en primer lloc el text de la primera edició del DGLC, i s'hi incorporen posteriorment les diferències detectades entre la primera i la segona edició.[11]

(10) L'Observatori de Neologia acull un projecte de recerca subvencionat per la Secció Filològica de l'Institut d'Estudis Catalans que consisteix en la recopilació en suport automàtic de les paraules noves que des de 1989 van apareixent a la premsa editada a Barcelona.

(11) L'escanneització del DGLC s'ha realitzat en col·laboració amb el projecte del Corpus Textual Informatitzat de la Llengua Catalana, del mateix IEC, i amb el Servei de Tractament de Textos Catalans de la Universitat de Barcelona.

– Les dades procedents d'altres fonts són introduïdes a BDOL manualment.[12]

3.2.2. *Selecció de nomenclatura*

Amb les informacions procedents de les diferents fonts contingudes a BDOL, es fa una primera selecció de dades amb l'objectiu de delimitar la nomenclatura del *Nou diccionari normatiu de la llengua catalana*.[13]

*a.* **Criteris**

Els criteris bàsics per a la selecció d'entrades de BDOL s'estableixen tenint en compte quatre elements:

- Que el text del DGLC, que és el text de base del NDNLC, s'havia de respectar al màxim possible. En la selecció, doncs, s'han recollit (amb algunes esmenes) totes les entrades que conté el DGLC.
- Que l'objectiu del treball és confeccionar un diccionari de llengua general (i no pas un diccionari especialitzat). Aquest objectiu ha fet que s'hagi hagut de prescindir dels termes massa especialitzats que el parlant mitjanament culte no empraria espontàniament.
- Que els parlants dels diferents dialectes del català s'hi han de sentir representats. Tot i que no es tracta d'un diccionari dialectal, s'han de recollir les formes pròpies dels grans dialectes catalans (tenint presents els criteris restrictius sobre els mots dialectals aprovats i publicats per la Secció Filològica als *Documents de la Secció Filològica, II*) perquè el diccionari sigui equilibrat des del punt de vista de les variants geolingüístiques.
- Que l'extensió del NDNLC no ha de sobrepassar en més d'un 30 % l'extensió del DGLC.

*b.* **Procés de selecció**

Pel que fa a la nomenclatura ja existent al DGLC, el procés de selecció ha consistit en la proposta de supressió d'algunes entrades del *Diccionari general* a partir de la lectura acurada que s'ha fet del text de Fabra i els criteris aprovats per la Secció Filològica.[14]

El procés de selecció de les entrades noves s'ha fet sobre els llistats obtinguts de BDOL. A cada entrada dels llistats esmentats se li ha atribuït un codi numèric que indica el grau d'importància per ser incorporat al NDNLC.

Els llistats s'han distribuït, segons el tipus d'unitats lèxiques, en quatre blocs per tal de facilitar el treball als diversos especialistes i assessors:

*a*) **Lèxic comú.** Les entrades corresponents a aquest bloc de llistats han estat seleccionades pels membres de la Secció Filològica.

(12) Aquesta tasca l'han dut a terme becaris col·laboradors, tots estudiants de filologia catalana.

(13) Aquesta selecció de dades ha estat necessària, tot i que les informacions ja havien estat introduïdes de manera selectiva, a causa de la diversitat de les fonts que s'han incorporat a BDOL, de la multiplicitat d'especialistes que hi han intervingut i de l'extensió del període temporal en què s'ha realitzat la introducció de les dades.

(14) Les supressions que s'han dut a terme solen ser de fitxes que inclouen entrades poc documentades, absoletes, massa especialitzades en qualsevol aspecte (tècnic, dialectal, argòtic), barbarismes innecessaris, etc., o, simplement, que no s'adequen als criteris fixats per la Comissió Lexicogràfica (criteri sobre la supressió dels noms propis com a entrades de diccionari, criteris sobre els gentilicis, criteri sobre els participis adjectivats, etc.).

*b*) **Terminologia**. De la selecció de termes, se n'han fet càrrec els especialistes de les diverses àrees del saber, membres de les seccions de l'Institut d'Estudis Catalans o d'alguna de les seves societats filials.

*c*) **Dialectalismes.** La selecció de dialectalismes ha estat fruit d'un recull que han fet membres de la Secció Filològica parlants de les diferents varietats dialectals, revisats a continuació per dos lingüistes també parlants de la mateixa varietat dialectal.

*d*) **Argot**. La selecció de l'argot l'han feta lingüistes de generacions diferents, per tal d'aconseguir que els termes fossin compartits per una part presumptament representativa de la societat.

Les Oficines Lexicogràfiques, un cop lliurats els diferents llistats, han acarat les diverses qualificacions donades a cada un dels mots i han elaborat una proposta d'acceptació. Aquesta proposta ha passat sistemàticament a la consideració de la Comissió Lexicogràfica i de les seccions de l'IEC.

*c.* **Selecció de manlleus**

Paral·lelament i complementària a les tasques de selecció que acabem de veure, les Oficines Lexicogràfiques han preparat la documentació de base per a la discussió sobre la incorporació de manlleus al *Nou diccionari normatiu.*

Per facilitar la selecció dels manlleus als membres de la Comissió Lexicogràfica i als membres de la Secció Filològica, atès que es tracta d'un tipus de mots especialment polèmic, s'ha elaborat un estudi detallat de cada cas a partir d'una graella de criteris:

*a*) La documentació del mot en les fonts de referència [1) *Diccionari general de la llengua catalana*, 2) *Diccionari de la llengua catalana*, 3) *Diccionari català-valencià-balear* o *Diccionari etimològic i complementari de la llengua Catalana*, 4) Sense cap documentació].

*b*) El grau de necessitat d'aquest mot, per a la qual cosa s'ha indicat si hi havia un recanvi genuí que pogués expressar la mateixa idea.

*c*) La indicació del registre del mot, per comprovar si hi ha o no canvi de registre respecte al mot que la llengua ja posseeix.

*d*) El grau d'adaptació del mot al sistema lingüístic.

### 3.3. Fases tercera i quarta: elaboració d'articles i elaboració del text del diccionari

Els criteris d'elaboració dels articles del NDNLC aprovats per la Secció Filològica són aplicats en el text del DGLC i serveixen de pauta per a l'elaboració dels articles nous.

Per dur a terme aquesta tasca, les Oficines Lexicogràfiques han redactat uns documents de treball que permeten transformar els criteris elaborats en la fase anterior en instruccions clares i precises. Aquests documents d'instruccions permeten que tot el conjunt de col·laboradors dugui a terme la regularització del text del DGLC i l'elaboració d'articles nous d'una manera ràpida, coherent i segura.

### 3.3.1. *Regularització del text del DGLC*

Un cop establerts els criteris definitius, i com a primer pas en l'elaboració del text del NDNLC, s'han d'aplicar aquests criteris al text del DGLC per tal de regularitzar-lo d'acord amb els nous criteris. Aquest treball es realitza a partir dels llistats extrets de BDOL, que contenen tota la informació del DGLC sobre el punt corresponent.

### 3.3.2. *Actualització dels continguts*

Regularitzat el text del DGLC, se n'actualitzen els continguts, fet que comporta modificacions i addicions respecte als articles del diccionari de Fabra, que es poden esquematitzar en els punts següents:

*a*) Supressió d'alguns dels significats del DGLC, perquè són erronis o massa especialitzats.
*b*) Modificació o supressió de definicions, exemples, sinònims i remissions del text del DGLC considerats obsolets, incomplets o marcats ideològicament.
*c*) Elaboració de les noves accepcions, que recullen tota la informació pertinent continguda en les fitxes.
*d*) Incorporació de les marques d'ús corresponents als mots.
*e*) Incorporació d'exemples que il·lustren els usos de les entrades.
*f*) Incorporació de subentrades que completen la nomenclatura del diccionari.

### 3.3.3. *Elaboració d'articles nous*

L'actualització de la nomenclatura del DGLC comporta, finalment, la incorporació de tots els neologismes de cada tipus de lèxic (lèxic comú, dialectalismes, argot i terminologia),[15] que han estat seleccionats pels diferents especialistes. Per a aquests casos, s'elaboren els articles complets del diccionari seguint els criteris aprovats per la Secció Filològica i les informacions recollides de les diverses fonts. L'elaboració d'aquests articles inclou les quatre operacions següents:

*a*) L'establiment de la forma de l'entrada.

(15) En aquest cas es consideren neologismes tots aquells mots que no tenien entrada en el DGLC.

*b*) La selecció de les informacions complementàries de l'entrada (morfològiques i etimològiques).
*c*) La selecció d'informacions sintàctiques, semàntiques i pragmàtiques per als nous articles.[16]
*d*) La selecció de subentrades i l'elaboració de la definició per a cada una de les subentrades.[17]

3.3.4. *Revisió*

El procés d'elaboració del *Nou diccionari normatiu* es tanca amb tres fases de revisió del text final: la primera revisió la duen a terme les Oficines Lexicogràfiques; la segona, les diverses seccions de l'Institut; i la darrera, la Secció Filològica.

La primera revisió del text del *Nou diccionari* consisteix en una lectura del text final amb dues finalitats bàsiques:

*a*) Comprovar que els criteris lexicogràfics aprovats s'han aplicat uniformement i sistemàtica.
*b*) Comprovar que les informacions que conté cada un dels articles no són repetitives ni incoherents.

La primera revisió del text ben segur que detecta algunes incongruències i porta a alguns canvis en les propostes que en el seu moment han fet els especialistes i assessors que han elaborat i revisat els articles per blocs temàtics. Aquest fet pot provocar ambigüitats o errors, sobretot en la terminologia.

En la segona revisió, per tal d'abordar els possibles errors fruit de la primera revisió, es tramet a les seccions de l'IEC el llistat dels canvis que s'han fet en cada cas, perquè comprovin que no s'hi hagi introduït cap equivocació en les definicions de conceptes científics i tècnics.

La tercera i última revisió, que duen a terme la Comissió Lexicogràfica i la Secció Filològica, consisteix a fer una lectura del text del *Nou diccionari normatiu* amb la finalitat de comprovar que les revisions anteriors s'han fet correctament i poder així presentar el text al Ple de l'Institut d'Estudis Catalans perquè sigui aprovat.

L'actualització del *Diccionari general de la llengua catalana* és un projecte institucional de l'Institut d'Estudis Catalans, que té el mandat de la societat de regular i normativitzar la llengua. Cal, doncs, que l'Institut aprovi en sessió plenària el text del *Nou diccionari normatiu de la llengua catalana.*

### 3.4. Fase cinquena: publicació del *Nou diccionari*

El projecte d'actualització del DGLC culmina amb la creació d'una base de dades específica només per al text del *Nou diccionari normatiu*, a partir de la qual s'iniciarà la cinquena fase del projecte: la publicació del NDNLC.

(16) La multiplicitat de fonts que constitueix BDOL fa que en moltes ocasions hi hagi informacions contradictòries d'un mateix mot o d'un mateix article. Per solucionar el problema quan hi ha divergència, es demana als especialistes que seleccionin les informacions pertinents, i les Oficines Lexicogràfiques apliquen els criteris del NDNLC a les informacions seleccionades.

(17) Com en el cas de la selecció de la nomenclatura i de l'actualització dels continguts, intervenen dos grups d'especialistes en l'elaboració dels articles:

– Per al lèxic comú, els dialectalismes i l'argot, els articles són elaborats a les Oficines Lexicogràfiques partint de les informacions recollides en les diverses fonts. És la Comissió Lexicogràfica qui els dóna el vist-i-plau i en proposa l'aprovació a la Secció Filològica.

– Per a la terminologia, són els membres de les seccions o de les societats filials de l'IEC els qui elaboren les definicions dels termes partint, també, del material contingut a BDOL. Les Oficines Lexicogràfiques completen la resta dels articles.

La publicació de l'obra inclou, d'una banda, totes les tasques pròpies de l'edició d'un llibre (elaboració d'un programa d'edició, correcció d'originals, galerades i paginades, etc.), i, de l'altra, l'edició del text del diccionari en suport informàtic.[18]

És evident que tota la feina exposada en aquest article fóra debades si el *Nou diccionari normatiu de la llengua catalana* no és assumit per les institucions del país com a instrument fonamental per a la normalització lingüística i si la societat catalana no l'acull com a eina fonamental de treball i de consulta.

Estiu 1994

M. TERESA CABRÉ
MARIA BOZZO
ÀNGELS EGEA
MARTA JUNCADELLA
CAROLINA SANTAMARIA
*Oficines Lexicogràfiques de l'Institut d'Estudis Catalans*

(18) La feina, però, no s'acaba amb la publicació del *Nou diccionari normatiu*. La societat i la llengua evolucionen constantment i aquesta evolució ha de recollir-se sistemàticament en els materials de les Oficines Lexicogràfiques de cara a les noves edicions del diccionari.

Cal tenir en compte també que el període en què s'ha preparat la primera edició del diccionari no ha bastat per tractar sistemàticament tota la seva informació. Les noves edicions hauran d'anar superant progressivament les anteriors.

El fet de disposar d'una base de dades en suport automàtic (BDOL) facilitarà aquests treballs de revisió i reactualització del diccionari. Una base de dades d'aquesta mena pot esdevenir el punt de partida fonamental per a qualsevol projecte lexicogràfic i per a qualsevol estudi sobre lèxic. Per aquest motiu, la voluntat de les Oficines Lexicogràfiques és que, en el futur, la base de dades del diccionari sigui una eina de consulta a l'abast de tota la comunitat investigadora i punt de suport per a la normalització de la llengua.

## REFERÈNCIES BIBLIOGRÀFIQUES

BASTARDAS, J.; CABRÉ, T.; ALSINA, V.; BOZZO, M.; JUNCADELLA, M.; SANTAMARIA, C. (1992) «Normativa lèxica del català. L'actualització del *Diccionari normatiu*». A: *XXe Congrès International de Linguistique et Philologie Romanes*. Tome IV, Section VI-Lexicographie/Iberoromania. Zuric.

FABRA, P. (1932) *Diccionari general de la llengua catalana*. Llibreria Catalònia, Barcelona.

— (1982) *Diccionari de la llengua catalana (1a ed.)*. Enciclopèdia Catalana, Barcelona.

— (1993) *Diccionari de la llengua catalana (3a ed. ampliada i actualitzada)*. Enciclopèdia Catalana, Barcelona.

— (1991) *Documents de la Secció Filològica, I*. Institut d'Estudis Catalans, Barcelona.

— (1993) *Documents de la Secció Filològica, II*. Institut d'Estudis Catalans, Barcelona.

RAFEL, J. (1987) «El "Diccionari del Català Contemporani"». *Serra d'Or*, núm. 333, pàg. 428-431.

**ELABORACIÓ DEL NOU DICCIONARI NORMATIU DE LA LLENGUA CATALANA**

**1.**

| TREBALLS | | | RESPONSABLES |
|---|---|---|---|
| CRITERIS LEXICOGRÀFICS | | | |
| estudi DGLC | mètode | entrades | Oficines Lexicogràfiques<br>Comissió Lexicogràfica<br>Secció Filològica |
| | | articles | |
| | text | presentació | |
| elaboració dels criteris del NDNLC | | | |

**2.**

| TREBALLS | | RESPONSABLES |
|---|---|---|
| ACTUALITZACIÓ DE LA NOMENCLATURA | | |
| selecció de fonts d'actualització | lèxic comú | Secció Filològica |
| preparació | terminologia | seccions de l'IEC |
| verificació<br>aprovació | dialectalismes | Secció Filològica |
| autoconstitució | argot | Secció Filològica |

FIXACIÓ DELS CRITERIS D'ELABORACIÓ DEL NDNLC

LLISTA D'ENTRADES NOVES DEL NDNLC

**3.**

| TREBALLS | | RESPONSABLES |
|---|---|---|
| ELABORACIÓ D'ARTICLES | | |
| aplicació dels criteris al DGLC | regularització de formes i presentació | Oficines Lexicogràfiques<br>Comissió Lexicogràfica<br>Secció Filològica |
| | actualització de continguts | |
| elaboració d'articles nous | | |
| revisió | | |

| TEXT DEL *NOU DICCIONARI NORMATIU DE LA LLENGUA CATALANA* | | |
|---|---|---|
| | entrades | DGLC |
| | | entrades noves |
| | articles | DGLC regularitzat |
| | | articles nous |

**4.**

| TREBALLS | RESPONSABLES |
|---|---|
| APROVACIÓ DEL TEXT DEL NDNLC | Ple de l'IEC |

**5.**

| TREBALLS | RESPONSABLES |
|---|---|
| PUBLICACIÓ | |
| correcció del text | Servei de Correcció |
| preparació de l'edició | |

JOAQUIM RAFEL

# UN CORPUS GENERAL DE REFERÈNCIA DE LA LLENGUA CATALANA

## INTRODUCCIÓ

En altres llocs i ocasions hem donat informació sobre l'origen i el desenvolupament del projecte de l'Institut d'Estudis Catalans anomenat Diccionari del Català Contemporani (DCC), dins el qual es realitza el Corpus Textual Informatitzat de la Llengua Catalana (CTILC).[1] No repetiré ací qüestions relacionades amb la justificació, l'inici, l'execució i el desenvolupament d'aquest projecte, que hom pot trobar en algunes d'aquestes publicacions, sinó que en aquest treball em referiré exclusivament a alguns aspectes de la caracterització del CTILC, i, sobretot, donaré compte del seu estat actual i de les possibilitats d'explotació, que, malgrat no haver-lo pogut completar encara, són ja una realitat.

Només recordaré breument en aquests paràgrafs introductoris, per a destacar-ne la motivació bàsicament lexicogràfica, que el projecte DCC té com a objectiu principal la futura elaboració d'un diccionari descriptiu del català d'acord amb els principis metodològics i tecnològics que la lexicografia ha assumit en els darrers anys com a conseqüència dels avenços de la ciència lingüística i de les possibilitats que la tècnica moderna posa a la nostra disposició. Un dels principis metodològics fonamentals d'aquest projecte és que l'elaboració d'un diccionari d'aquestes característiques ha

(1) «Cap a un diccionari del català contemporani», *Segon Congrés Internacional de la Llengua Catalana, IV, Àrea 3: Lingüística Social* (celebrat a Palma del 30 d'abril al 4 de maig de 1986), Palma, 1992, pàgs. 589-595. «El "Corpus textual automatitzat de la llengua catalana"» (en col·laboració amb J. M. Solanellas). *Actas de las II Jornadas Españolas de Documentación Automatizada. 20-22 de Noviembre de 1986. Ponencias y comunicaciones* (Torremolinos-Málaga, 1986), pàgs. 147-161; «El "Diccionari del català contemporani"». *Serra d'Or*, 1987, pàgs. 428-431; «El "Corpus textual informatitzat de la llengua catalana" i el "Diccionari del català contemporani". Un proyecto del Institut d'Estudis Catalans», *Anthropos*, 1988, Documentación cultural e información bibliográfica, pàgs. V-VII; «El "Diccionari

d'utilitzar com a font principal un corpus textual suficientment representatiu de la llengua que vol reflectir, en comptes de fonamentar-se, com s'ha fet tradicionalment en lexicografia, en els diccionaris preexistents i en un mètode de treball basat en la intuïció o en la mera competència dels lexicògrafs. Entre altres molts avantatges que ara no comentaré, això permet d'establir el significat dels mots a partir de l'ús que se n'ha fet en la llengua, en comptes d'utilitzar mètodes apriorístics o intuïtius; permet també d'exemplificar amb frases preses dels textos reals els diferents valors i usos de cada mot i les diferents estructures sintàctiques de què pot formar part, i permet, encara, la utilització d'informació de caràcter estadístic, que pot arribar a tenir una gran importància en l'estudi científic d'una llengua.

Cal dir, però, que, per bé que la idea d'aquest corpus sorgí d'una necessitat concreta i amb una finalitat específica, des del primer moment de la seva elaboració s'ha tingut en compte d'estructurar-lo de tal manera que no en resultin restringides les possibilitats d'utilització; s'han pres les precaucions metodològiques i tècniques necessàries perquè el corpus pugui ésser emprat amb profit per a qualsevol tipus d'investigació lingüística que requereixi la utilització de dades procedents de la llengua escrita. L'esforç humà i els recursos necessaris per a portar a terme un projecte d'aquesta naturalesa són molt importants; això ens féu pensar des del primer moment que el resultat d'aquests esforços havia de poder ésser utilitzat en múltiples aplicacions d'una manera indefinida.

## CARACTERÍSTIQUES DEL CORPUS TEXTUAL INFORMATITZAT DE LA LLENGUA CATALANA (CTILC)

### *Característiques generals*

La caracterització general de qualsevol corpus té en compte factors de naturalesa temporal, de naturalesa qualitativa i de naturalesa quantitativa; és a dir, el període de temps que abraça, el tipus de text que conté i l'extensió total de text que l'integra.

Des del punt de vista temporal, el CTILC abraça des de 1833,[2] com a data simbòlica de la represa de l'ús literari de la llengua catalana en l'època moderna, fins al moment actual.[3] Per una altra banda, els textos tinguts en compte per a formar part del corpus són textos publicats, sigui en forma de llibre o d'opuscle, full solt, diari, revista, etc.[4] Els textos seleccionats han estat introduïts d'una manera íntegra, amb les poques excepcions d'algunes obres que, per la seva extensió, no era possible introduir totalment; és important també tenir en compte que el text font, que s'utilitza com a referència definitiva, és sempre la primera edició de l'obra, i que el seu contingut és introduït al corpus sense fer-hi cap correcció ni modificació; això significa que el corpus incorpora tota la diversitat de formes i grafies que trobem, sobretot, en textos no normalitzats.

del català contemporani": treballs realitzats i previsions de futur». *Llengua & Literatura*, Revista anual de la Societat Catalana de. Llengua i Literatura, 5, 1992-93 [1994], pàgs. 733-737; «Le "Diccionari del català contemporani" et le "Corpus textual informatitzat de la llengua catalana". Brève description du projet et état des travaux», *Actes du XX Congrès International de Linguistique et Philologie Romanes, Université de Zurich (6-11 avril 1992)*, IV, Tübingen-Basel, 1993, pàgs. 816-821. D'altra banda, en les diferents memòries d'activitats de l'Institut hom pot trobar informació sobre l'estudi previ, l'aprovació, l'inici i el desenvolupament progressiu d'aquest projecte: *Memòria d'activitats. (octubre 1982-desembre 1983)*, Barcelona, 1984, pàgs. 131 i 135-137; *Memòria d'activitats 1984*, Barcelona, 1986, pàgs. 136 i 137; *Memòria d'activitats. Curs 1988-89*, Barcelona, 1990, pàgs. 25-28; *Memòria d'activitats. Curs 1989-90*, Barcelona, 1991, pàgs. 137-139; *Memòria d'activitats. Curs 1990-91*, Barcelona, 1992, pàgs. 147-159 i *Memòria d'activitats. Curs 1991-92*, Barcelona, 1993, pàgs. 165-177. Un altre títol que forma part de la història del projecte és la publicació interna *Investigación lingüística 1989-92*, Generalitat de Catalunya, Departament de Presidència. Comissió Interdepartamental de Recerca i Innovació Tecnològica (CIRIT). «Programes de recerca i desenvolupament de la Generalitat de Catalunya», 3, (s. d.) [Anexo nº 3. Diccionario del catalán contemporáneo. Corpus textual informatizado de la lengua catalana. Desarrollo general y estado actual del proyecto (43 pàgs.). Anexo nº 4. Diccionari del català contemporani. Informe sobre la fase de lematización. Descripción y métodos (34 pàgs)].

(2) Hi ha alguna obra anterior a aquesta data, però a efectes del còmput de grups cronològics prenem l'any 1833 com a data inicial.

Quant a la naturalesa de la llengua o a la temàtica dels textos, he de dir que el CTILC inclou textos de caràcter literari, que és el més habitual en corpus que existeixen per a d'altres llengües, però també textos de caràcter no literari, és a dir, textos que tracten de qualsevol dels temes possibles i des de qualsevol perspectiva, sense una finalitat artística o una intenció estètica; la proporció entre els textos literaris i els no literaris és del 40% per als primers i el 60% per als segons; aquests percentatges han estat establerts després d'estudiar a fons el conjunt de la producció escrita en català i tenint com a principal objectiu el de mantenir una representativitat adequada, proporcional, en la mesura possible, al volum de textos publicats i a la varietat tipològica.

Pel que fa a l'extensió del corpus, per bé que no està del tot acabat, com ja he dit, i com precisaré encara més endavant, podem donar una estimació molt acostada a la realitat, pel fet que hi ha introduïts a l'ordinador (i, per tant, comptabilitzats) pràcticament el 90% dels textos. El total de l'extensió prevista és d'entre cinquanta-dos i cinquanta-tres milions de mots. La repartició d'aquesta xifra, segons la tipologia dels textos i les seccions cronològiques que hem establert, és la següent:[5]

LLENGUA NO LITERÀRIA

| | Secció I (1833-1973) | Secció II (1874-1913) | Secció III (1914-1988) | TOTAL |
|---|---|---|---|---|
| 0. Correspondència | 2.597 | 40.252 | 77.750 | **120.599** |
| 1. Filosofia | 6.886 | 210.337 | 1.532.615 | **1.749.838** |
| 2. Religió i teologia | 235.218 | 533.478 | 2.203.350 | **2.972.046** |
| 3. Ciències socials | 202.995 | 909.723 | 4.494.161 | **5.606.879** |
| 4. Premsa | 109.686 | 462.242 | 2.998.680 | **3.570.608** |
| 5. Ciències pures i naturals | 23.715 | 372.237 | 1.828.993 | **2.224.945** |
| 6. Ciències aplicades | 225.303 | 573.767 | 3.680.495 | **4.479.565** |
| 7. Belles arts. Divertiment. Jocs. Esports | 5.027 | 416.406 | 2.361.817 | **2.783.250** |
| 8. Llengua i literatura | 136.310 | 333.840 | 1.759.025 | **2.229.175** |
| 9. Història i geografia | 107.855 | 426.895 | 2.962.889 | **3.497.639** |
| **Total** | **1.055.592** | **4.279.177** | **23.899.775** | **29.234.544** |

LLENGUA LITERÀRIA

| | Secció I (1833-1973) | Secció II (1874-1913) | Secció III (1914-1988) | TOTAL |
|---|---|---|---|---|
| 1. Assaig | 0 | 381.164 | 2.670.096 | **3.051.260** |
| 2. Narrativa | 536.748 | 2.499.426 | 10.885.237 | **13.921.411** |
| 3. Poesia | 245.472 | 604.528 | 1.836.732 | **2.686.732** |
| 4. Teatre | 378.981 | 836.246 | 2.471.084 | **3.686.311** |
| **Total** | **1.161.201** | **4.321.364** | **17.863.149** | **23.345.714** |

(3) En realitat les darreres obres seleccionades corresponen a 1988, any en què es tancà la selecció quan es realitzà aquesta tasca, que prenia com a marc períodes de cinc anys. Això no vol dir que, una vegada completat, el corpus no pugui ésser actualitzat progressivament a partir d'un pla adequat i dels recursos necessaris per a portar endavant aquesta tasca.

(4) Només s'han introduït alguns textos no publicats, com cartes o escriptures notarials; en aquests casos, els textos font són manuscrits o mecanoscrits. Aquests han estat tinguts en compte per a completar el panorama de la llengua escrita amb uns textos que, si bé no han estat publicats, tenen la característica d'haver estat donats per definitius; la carta, perquè en el moment d'enviar-la és un text que s'ha donat per acabat, i les escriptures notarials, perquè la seva pròpia naturalesa els dóna la característica de textos perfets.

(5) Aquesta repartició s'ha d'entendre com a definitiva pel que fa a la llengua no literària, perquè aquesta part del corpus està definitivament acabada. En canvi, les dades corresponents a la llengua literària representen l'estimació que es pot fer en aquest moment, perquè encara no s'han acabat els treballs de constitució d'aquesta altra part del corpus; tanmateix, aquesta estimació és molt aproximada a la realitat.

*Els textos seleccionats*

A l'hora de seleccionar els textos per a formar part del corpus, l'objectiu principal ha estat d'assegurar el major grau de representativitat possible de tots els textos publicats dins els límits temporals establerts, i al mateix temps un equilibri raonable entre les diferents classes. Per aconseguir aquest objectiu s'ha dividit el total de l'extensió temporal del corpus (1833-1988) en vint-i-tres grups cronològics (vuit grups de deu anys cada un entre 1833 i 1913, i quinze grups de cinc anys cada un entre 1914 i 1988); a l'hora de fer la selecció, s'ha procurat que, no solament d'una manera general, sinó també dins cada un d'aquests grups, hi hagi una representació adequada de cada tipus de text, entenent per tipus cada un dels gèneres (narrativa, teatre, poesia i assaig) pel que fa a la llengua literària, i cada una de les deu àrees temàtiques establertes, subdividides, la major part, en diverses subàrees, també fins a deu, pel que fa a la llengua no literària (aquestes àrees temàtiques, que, per qüestions pràctiques, segueixen de prop els criteris de classificació decimal de les biblioteques, són: 0. Correspondència, 1. Filosofia, 2. Religió i teologia, 3. Ciències socials, 4. Premsa, 5. Ciències pures i naturals, 6. Ciències aplicades, 7. Belles arts, divertiments, jocs i esports, 8. Llengua i literatura, 9. Història i geografia). Entenem que la representació és adequada des del punt de vista cronològic i temàtic quan podem garantir fins a un cert punt una relació de proporcionalitat entre la quantitat de text seleccionat per a cada tipus i per a cada grup cronològic i el volum de text de cada tipus publicat dins cada un d'aquests grups. A aquests criteris generals se n'hi afegeixen uns altres de complementaris, com són la consideració de les diferències dialectals, els subgèneres, o bé la temàtica o l'estil diferents dins els textos corresponents a un mateix gènere literari, per exemple. La minuciositat d'aquesta operació ha tingut com a objectiu d'aconseguir el màxim grau de representativitat i d'equilibri des de diversos punts de vista, a fi que el CTILC pugui tenir la funció que hem apuntat de corpus general de referència de la llengua catalana; sembla obvi que les conclusions a què podem arribar partint d'un corpus equilibrat i representatiu tenen un valor més alt que aquelles que podríem deduir d'un corpus que no presentés aquestes característiques.

El total d'obres seleccionades és de 3.302, que es reparteixen en 2.296 corresponents a la llengua no literària i 1.006 a la llengua literària.

*Les dades del corpus*

En una part dels corpus, més o menys extensos, que existeixen per a diferents llengües, la unitat de treball a través de la qual es pot accedir a la informació que contenen és el mot entès com a unitat gràfica, tal com apareix en el text font; això té l'inconvenient de l'alt grau d'ambigüitat que poden arribar a presentar aquestes unitats, que

varia segons les llengües (en català, per exemple, la forma *fins* pot correspondre a una preposició, a un substantiu masculí, a un substantiu femení o a un adjectiu; *creuen* pot ésser una forma del verb *creuar* o una forma del verb *creure*; *causa* pot ésser un substantiu o una forma del verb *causar*, o bé pot fer part d'una locució prepositiva; *cap* pot ésser un substantiu, una preposició, un adjectiu o bé una forma del verb *cabre*; *deu* pot correspondre a un adjectiu numeral, a un substantiu femení, a un substantiu masculí, o, encara, a dues formes diferents del verb *dar* o a dues formes diferents del verb *deure*, etc.). Aquest fet pot esdevenir un inconvenient a l'hora de consultar les dades del corpus, perquè l'usuari té accés a una informació massa diversa a partir d'una mateixa unitat d'interrogació.

En altres corpus textuals, en canvi, s'ha dut a terme una primera anàlisi de caràcter lingüístic de les dades, que té com a un dels objectius de desfer el tipus d'ambigüitat a què m'acabo de referir (ambigüitat gramatical), i, a la vegada, relaciona entre elles les diferents formes pertanyents a una mateixa sèrie flexional; es tracta de l'operació anomenada *lematització*. D'acord amb això, doncs, des del punt de vista del grau d'interpretació de les dades que contenen, els corpus es poden classificar en no lematitzats i lematitzats. El CTILC és un corpus del segon tipus; el seu procés d'elaboració inclou una primera fase d'introducció i validació de les dades textuals i una segona fase en la qual es realitza la lematització de totes les ocurrències del corpus.[6]

A través de la lematització, s'aconsegueixen, doncs, dos objectius particulars, que són, de fet, un conseqüència de l'altre: cada una de les ocurrències de cada mot gràfic (és a dir, cada una de les seves aparicions al llarg del text) és categoritzada gramaticalment i associada a una forma de referència anomenada *lema* (que es correspon aproximadament amb allò que podem considerar una entrada de diccionari); per exemple, davant cada una de les ocurrències de la forma gràfica *nous*, a la vista de l'entorn contextual, es determina a quina de les possibilitats de categorització gramatical correspon de les quatre que pot tenir aquesta seqüència de caràcters: NOU *adj.*, NOU *subst. masc.*, NOU *subst. fem.*, i NOURE *verb* (en el primer cas és la forma de masculí plural; en el segon i en el tercer, la forma de plural; i en el quart, la forma de segona persona del singular del present d'indicatiu); o bé, davant les diferents ocurrències de la forma gràfica *cap* en el text, es determina si corresponen a CAP *subst. masc.*, a CABRE *verb*, a CAP *prep.* o a CAP *adj.* (en el primer cas és la forma de singular, en el segon la de tercera persona del singular del present d'indicatiu, i en el tercer i en el quart, no té categorització morfològica, perquè correspon al que correntment anomenem mots invariables. Així, per una banda es desambigüen gramaticalment les formes homògrafes, i per una altra s'agrupen sota un mateix lema els diferents components d'una mateixa sèrie inflectiva; és a dir, algunes de les ocurrències de *cap* s'agrupen amb *caps* sota el lema CAP *subst. masc.*, i certes altres ocurrències de la mateixa grafia es relacionen amb *caben, cabem, cabran, cabrien, càpiga,* etc., sota el lema CABRE *verb*; de la mateixa manera, algunes de les ocurrències de la grafia *nous* s'agrupen

(6) La lematització es realitza en el CTILC a través d'un procediment semiautomatitzat; l'ordinador elabora una proposta de lematització, que després ha d'ésser examinada i simplement donada per bona o bé completada manualment. El sistema (programes i fitxers de lemes i formes necessaris), que no puc comentar aquí amb detall, ha estat elaborat completament dins els treballs del CTILC.

amb *noves, nova, nou,* sota el lema NOU, *adj.*, altres s'agrupen amb *nou* sota el lema NOU *subst. masc.*, o bé sota el lema NOU *subst. fem.*, segons els seus contextos, i altres, juntament amb *noïa, nourà, nogui,* etc., sota el lema NOURE *verb.* En un corpus sense lematitzar un usuari podria demanar només informació a través de la grafia *cap*, o de la grafia *nous*, i la informació que obtindria sobre les diferents ocurrències d'aquestes grafies mitjançant aquesta forma d'interrogació seria indiscriminada des del punt de vista lingüístic; és a dir, rebria, sense diferenciar, les formes corresponents al substantiu masculí, les corresponents al verb, a la preposició i a l'adjectiu, en el cas de *cap*, o a l'adjectiu, al substantiu masculí, al substantiu femení i al verb, en el cas de *nous;* en canvi, en un corpus lematitzat d'acord amb aquests principis, l'usuari té accés a les dades agrupades adequadament a partir del substantiu masculí CAP, del verb CAURE, de la preposició CAP, o de l'adjectiu CAP, o bé de l'adjectiu NOU, del substantiu masculí NOU, del substantiu femení NOU o del verb NOURE, perquè la base de dades on es troben emmagatzemades les diferents ocurrències del corpus conté aquesta informació.[7]

Si reprenem l'exemple de *cap*, podem dir que, en principi, el lema CAP *subst. masc.* tindrà relacionades dues formes, la de singular *(cap)* i la de plural *(caps)*, i que les diferents ocurrències d'aquest lema que hi ha en el corpus estaran associades a l'una o a l'altra d'aquestes dues formes. Afegiré, però, alguns comentaris que ajudaran el lector a fer-se una idea de la manera com s'ha tractat l'extraordinària varietat formal que apareix en el corpus i alguna qüestió de caràcter morfològic. En primer lloc cal tenir en compte que el mateix procés de lematització, a més de relacionar les diferents formes d'una mateixa sèrie flexional, agrupa també les diferents variants gràfiques que poden aparèixer en els textos (així, en el cas del lema CAP *subst. masc.*, trobem formes com *cab, cáp, câp, cabs, cáps,* a part de les normatives; o, en el cas del lema FILL, *subst. masc.*, trobem *figl, fii, fill, fiy, fiís, fiys,* a més de *fill* i *fills*).[8] Per una altra banda, a l'hora de definir els criteris per a la lematització, decidírem classificar com a formes d'un mateix lema –el corresponent al mot primitiu– tots els derivats apreciatius (diminutius, augmentatius, pejoratius, intensius),[9] per als quals, doncs, no es creen nous lemes; així, en l'estat actual del corpus, apareixen com a formes del lema CAP *subst. masc.* els mots *cabet, caparró, caparronet, caparrot, caperrot, capet, cabets, caparrins, caparronets, caparrons, caparrots, caperrots, capets, cabot*; i, en el cas de GRAN *adj.*, trobem *regran, regrán, regrans, grandaçot, grandàs, grandet, grandissim, grandíssim, grandot, grandeta, grandísima, grandissima, grandíssima, grandota, grandassassos, grandassots, grandets, grandissims, grandíssims, grandolassos, grandots, grandòts, grandetes, grandísimas, grandíssimas, grandíssimes*; totes aquestes formes van codificades amb una marca que permet tractar-les conjuntament, a part de la resta, si convé, per a un tipus d'estudi específic. Cal també tenir en compte que, quan un mot apareix usat metalingüísticament en un text, es codificat com a tal, i, en conseqüència, es classifica com una forma diferenciada de la resta, encara que tingui la mateixa grafia i el mateix codi morfològic que alguna altra. Totes aquestes qüestions es poden veure en els exemples de les figures 3 i 4.

(7) No puc estendre'm en aquest treball en tota una sèrie d'aspectes relacionats amb la lematització. Cal dir, però, que hi ha hagut una certa confrontació de parers sobre si és preferible lematitzar els corpus o no lematitzar-los: l'actitud de no lematitzar es basa en l'opinió que qualsevol tipus d'intervenció sobre el text verge pot condicionar les conclusions ulteriors. Per a aquesta qüestió ara em limitaré a remetre el lector a dos treballs que resumeixen aquestes posicions contraposades: MAURICE TOURNIER; «Sur quoi pouvons-nous compter? Réponse à Ch. Muller», in *De la plume d'oie à l'ordinateur. Mélanges en l'honneur de Gérald Antoine*, [Responsables de la publication: GILBERT BOISSIER: DANIELLE BOUVEROT], Nancy, 1984, pàgs. 131-136; i CHARLES MULLER; «Lemmatisation et information», *Hommage à Pierre Guiraud* [Comité de réd.: CHARLES P. BOUTON; ÉTIENNE BRUNET, LOUIS-JEAN CALVET] *Annales de la Faculté des Lettres et Sciences Humaines de Nice*. 52, 1985, pàgs. 285-291. Al marge, però, de debats ideològics, els enormes avantatges pràctics d'un corpus lematitzat són actualment reconeguts d'una manera generalitzada: «Although there is nothing new in the idea of assigning individual words to a particular word-class, it is no trivial task to do this with every word in a body of text. But there are considerable benefits to having a corpus in which every word is accompanied by a label which indicates what type of word it is.» (GEOFFREY LEECH, STEVEN FLIGELSTONE; «Computers and Corpus Analysis», in CHRISTOPHER S. BUTLER (ed.); *Computers and Written Texts*, 1992, pàgs. 124-125).

D'altra banda, admetent la lematització com una manera adequada de categoritzar les unitats d'un corpus a fi de facilitar-ne una explotació més efectiva, es presenten una sèrie de problemes metodològics de no fàcil solució: cal prendre una sèrie de decisions, entre d'altres, sobre la manera de

Hi ha, però, casos en què la variació va més enllà de la purament gràfica (i no té res a veure amb els derivats apreciatius ni amb els usos metalingüístics a què m'acabo de referir); es tracta dels casos en què aquesta no és merament gràfica, sinó que comporta una diferència en l'estructura fonològica o morfològica del mot, que pot semblar suficient per a determinar lemes diferents, però amb una similitud prou forta com per a mantenir algun tipus de relació lògica que permeti tractar-los, si convé, d'una manera agrupada (ens referim a casos com *almorzar* i *armosar,* al costat d'*esmorzar*; o bé *abercoc, abricoc, albaricoc, bercoc* o *obrecoc*, al costat d'*albercoc*; *juriol,* al costat de *juliol; radere,* al costat de *darrere; rellonge, relonge* o *relotge,* al costat de *rellotge; esgolfa* o *gorfa,* al costat de *golfa,* etc.). Per a resoldre aquests casos prenguérem la decisió de crear dues categories de lemes, que resten relacionats jeràrquicament dins el sistema informàtic: els lemes principals i els lemes secundaris; això permet de tractar-los independentment, cada un amb les seves formes associades, o bé acumular les dades dels secundaris als principals, quan la naturalesa de la informació que cerquem ho requereix (a les figures 1 i 2 podem veure uns exemples d'aquest tipus d'estructura).

| LEMA | RANG | FORMA |
|---|---|---|
| ESMORZAR M | principal | esmorçar S |
| | | esmorsà S |
| | | esmorsar S |
| | | esmorzar S |
| | | esmorçars P |
| | | esmorsars P |
| | | esmorzars P |
| | | esmorzar MET |
| ALMORZAR M | secundari | esmorsar S |
| | | almorzar S |
| | | esmorzar MET |
| ARMOSAR M | secundari | armosar S |
| | | armosarot DS |

FIGURA 1. En aquesta taula podem observar la relació jeràrquica establerta entre el lema ESMORZAR, com a lema principal, i els lemes ALMORZAR i ARMOSAR, etiquetats com a secundaris; cada un està relacionat amb les formes corresponents. Això permet, amb la contribució dels programes adequats, tractar-los de manera separada, com a lemes independents, o bé de manera conjunta, com si es tractés d'un sol lema, segons l'objectiu de la recerca projectada, o, simplement, de la naturalesa de la informació que es vol obtenir.

categoritzar casos que permeten més d'una interpretació (conjunció o adverbi, substantiu o adjectiu, verb o substantiu –en el cas de certes ocurrències de l'infinitiu–, verb o adjectiu –en el cas de certs participis–, etc.); cal tenir en compte que els textos utilitzen de vegades la llengua fins als seus límits, i que aquest fet planteja una sèrie de problemes als quals cal trobar una solució, sovint qüestionable. Els autors que acabo de citar es refereixen d'una manera tan breu com contundent a aquest problema: «One problem is that what class a word belongs to turns out in many cases to be very difficult to decide.» (*Ibid.*, pàg. 125) A part d'això, cal decidir també el grau d'intervenció que hom vol practicar al llarg d'aquesta operació (per exemple, classificant per separat els mots que formen part de locucions o perífrasis, o bé agrupant-los en un sol lema complex), que pot ésser més suau o més severa. Per a totes aquestes qüestions és convenient d'establir criteris tan clars com sigui possible, perquè, al llarg de la seva aplicació, la lematització del corpus es faci d'una manera coherent i amb el màxim grau d'uniformitat: en el cas del CTILC, la resolució d'aquesta problemàtica ha donat lloc a un voluminós *Manual de lematització.*

(8) Hi ha casos bastant complexos. El verb veure, per exemple, en l'estat actual del corpus, té 639 formes diferents; de les quals 162 corresponen a usos metalingüístics.

(9) Aquestes formacions amb afixos no lexicals no solen ésser tingudes en compte en els diccionaris. Evidentment, no hi poden donar lloc a una entrada nova, però des del punt de vista de la descripció de l'ús lingüístic, és molt important de conèixer quines de les possibilitats virtuals es materialitzen realment, en quines circumstàncies o situacions ho fan, quina freqüència tenen en la realitat de la llengua, o amb quines bases lexicals concretes es combinen preferentment aquest tipus d'afixos.

| LEMA | RANG | FORMA |
|---|---|---|
| DONCS | principal | donç |
| | | donchs |
| | | dónchs |
| | | dònchs |
| | | doncs |
| | | dóncs |
| | | dòncs |
| | | dons |
| | | donchs MET |
| | | doncs MET |
| | | dòngks MET |
| | | dòngs MET |
| | | dons MET |
| | | dòns MET |
| DONC | secundari | donc |
| | | donch |
| DONCES | secundari | donças |
| | | donsas |
| DONQUES | secundari | doncas |
| | | donques |
| DÒS | secundari | dos |
| | | dòs |
| | | dôs |
| | | dòs MET |

FIGURA 2. Veiem en aquesta figura un altre exemple de lemes relacionats amb el criteri de principal i secundaris: principal DONCS; secundaris, DONC, DONCES, DONQUES, DÒS. Cada un agrupa les seves formes. Les formes corresponents al lema DONCS podrien haver estat agrupades d'una altra manera; a l'hora de prendre les decisions, però, es considerà que la diferència entre formes com *donç* o *dons*, d'una banda, i *doncs* o *donchs*, d'una altra banda, no era suficient per a donar d'alta lemes diferents (DONCS i DONS).

## LA BASE DE DADES TEXTUAL DE LA LLENGUA CATALANA (BDTLC)

Una vegada acabat el procés de lematització de cada una de les obres introduïdes prèviament en el sistema informàtic, i després d'haver estat donats per vàlids, els resultats d'aquesta operació passen a formar part d'una única base de dades, que conté tota la informació necessària per a una adequada explotació del corpus: és la Base de Dades Textual de la Llengua Catalana.

Aquesta base de dades conté, per a cada una de les ocurrències del corpus, les informacions següents: la forma gràfica, la localització de l'ocurrència (codi de l'obra, pàgina, línia i número d'ordre del mot dins la línia), el codi morfològic i el número de referència del lema que li ha estat atribuït. D'altra banda, la base de dades també conté una sèrie d'elements que permeten reconstruir el context corresponent a cada ocurrència; aquestes dades són, bàsicament, tots aquells codis que corresponen al que podríem considerar *grosso modo* signes de puntuació (punts, comes, punts i comes, punts suspensius, signes d'interrogació i d'exclamació, etc.), i també les parts del text originari que no han estat tingudes en compte durant el procés de lematització, ja que en la fase d'introducció havien estat codificades com a no analitzables (nombres expressats en xifres, citacions d'altres autors o en d'altres llengües, etc.), o bé com a noms propis, que no són tinguts en compte en el procés de lematització, però que cal recuperar a l'hora de reconstruir el text. Cal remarcar, per a acostar una mica el lector a l'estructura dels fitxers informàtics que contenen tota la informació del corpus i dels programes que la tracten, que el text en la seva forma seqüencial (tal com es troba en les obres de referència) no és emmagatzemat en cap lloc de la memòria de l'ordinador; fins i tot en els casos en què la màquina ens forneix un context, aquest és construït cada vegada mitjançant l'execució dels programes de consulta, que en cerquen els diferents components entre les unitats d'informació que conté la base de dades, a les quals acabo de fer referència.

Per a l'explotació de la base de dades han estat elaborats una sèrie de programes que permeten, d'una banda, l'obtenció de llistats diversos, i, d'una altra banda, la consulta interactiva.[10]

Pel que fa als llistats que hom pot obtenir en aquest moment, fonamentalment poden ésser de lemes, de lemes i formes o de formes; les classificacions poden fer-se per ordre alfabètic directe, per ordre alfabètic invers o per ordre decreixent de freqüències. Aquests llistats es poden efectuar de tot el corpus, o de parts del corpus definides prèviament d'acord amb criteris cronològics o tipològics; poden realitzar-se també d'una obra o d'un conjunt d'obres determinades; d'altra banda, s'ha dissenyat també un tipus de llistat es què es pot apreciar la repartició de la freqüència dels diferents lemes amb criteris cronològics o tipològics. En l'apèndix d'aquest treball (figures I-V) en podem veure algunes mostres.

(10) Aquests programes foren elaborats inicialment per a la consulta interna i el manteniment de la base de dades, però, demostrada la seva utilitat general, i després de completar-los en alguns aspectes, avui permeten que el corpus pugui ésser consultat per qualsevol usuari per a obtenir-ne la informació necessària.

El sistema de consulta interactiva permet accedir a la base de dades a través del *lema*, a través de la *forma* o a través de la *localització*; però aquesta darrera modalitat, i, en part, la segona, tenen un interès exclusivament intern per a la validació i el manteniment de la base de dades. Per una altra banda, el sistema permet accedir, o bé al conjunt de tot el corpus, o bé a una part de les dades, que poden ésser seleccionades, com un subcorpus, per l'usuari; aquest subcorpus pot ésser definit també a partir de criteris cronològics, a partir de criteris tipològics, o d'ambdós alhora, o bé a partir d'un autor o d'un grup d'autors, o d'una obra o un grup d'obres; a més, si la consulta afecta un lema o una forma que té una freqüència molt elevada en el corpus, i, per tant, fóra molt feixuc de consultar-ne tots els contextos, el programa permet de ferne una selecció prèvia, a través de l'aplicació d'unes taules de nombres aleatoris, prenent com a referència un tant per cent del total d'ocurrències, o un nombre concret, fixats per l'usuari.

A través del sistema de consulta interactiva, podem obtenir, per exemple, si entrem per un lema determinat, la freqüència total del lema, les formes que té associades i la freqüència de cada una (vegeu-ne uns exemples, tal com ens apareixen a la pantalla de l'ordinador, a les figures 3 i 4), i, si accedim a continuació a una d'aquestes formes, obtenim les diferents ocurrències concretes que presenta en el corpus (vegeu la figura 5); podem accedir també al context que correspon a cada una d'aquestes ocurrències, i a una sèrie de dades relacionades, com són, l'autor i el títol de l'obra on es troben, l'any de publicació, el tipus de text (literari/no literari i les seves subdivisions), la localització específica (pàgina, línia, número d'ordre dins la línia), i les dades morfosintàctiques de la forma i del lema; si el context ofert en primera instància (tres línies físiques de l'edició de referència) fos insuficient per a la finalitat de la consulta, l'usuari pot accedir de manera immediata al paràgraf sencer; a les figures 6, 7 i 8 podem veure tres exemples d'obtenció de context, amb el context breu que apareix en primer lloc, i el context més extens que l'usuari pot demanar a continuació, si li cal; la figura 9 ens mostra un exemple de context en el cas d'un derivat apreciatiu associat al lema CAP *subst. masc.* i, encara, la figura 10, un exemple de context en un ús metalingüístic.

Per a poder proporcionar tota aquesta informació, el programa de consulta accedeix inicialment a la BDTLC, que conté només la informació més indispensable, sense repeticions innecessàries, amb la finalitat de fer les consultes el més àgils i ràpides possible. A partir d'aquestes dades, el programa recorre a uns altres fitxers: els del Diccionari Bàsic Informatitzat (DBI)[11] i els del Repertori d'Autors i Obres (RAO).[12] El DBI inclou les dades corresponents a la grafia dels lemes i de les formes i la seva categoria gramatical, i el codi de procedència (que ens indica si el lema o la forma es troben en alguna de les fonts lexicogràfiques o gramaticals utilitzades per a construir el DBI, o bé si no són recollits en aquestes fonts i, per tant, procedeixen únicament dels textos introduïts); d'altra banda, el programa obté del RAO el nom de l'autor, el títol de l'obra, l'any de publicació, i altres dades relacionades amb l'obra (tipus de llengua, etc.).

(11) El DBI és un element clau en l'operació de lematització, que es realitza per un procediment semiautomatitzat; consta d'un inventari de lemes i d'un inventari de formes (les formes flexionals que corresponen als dits lemes), cada una de les quals està relacionada lògicament amb el lema o els lemes a què pot estar associada. Inicialment el DBI contenia 88.067 lemes i 631.182 formes; en el moment actual, després d'haver-hi incorporat els lemes i les formes nous que han sorgit dins els textos en el curs de l'elaboració del corpus, consta de 160.952 lemes i 883.083 formes.

(12) El RAO fou constituït amb la finalitat de facilitar la tasca de selecció de les obres que havien de formar part del corpus. Aquest repertori, que consta de 6.273 autors i 26.121 obres, permeté creuar dades, com les dates de publicació de les obres, els tipus de text i els noms dels autors, a fi de poder prendre les decisions de la manera més objectiva possible. En el procés de consulta s'utilitza per a recuperar *in extenso* les dades referents a les publicacions, que en la BDTLC apareixen codificades.

BDTLC RECONSTRUCCIÓ SEGONS L'ORIGINAL A PARTIR D'UN LEMA

Lema: cap — Cat. gram.: M
Codi: 37.921 — Freq. lema: 18.032

Selecció de la forma

| | Forma | C.M. | Freq. abs. |
|---|---|---|---|
| | cap | FS | 1 |
| | cab | S | 5 |
| | cap | S | 15.864 |
| | cáp | S | 6 |
| | câp | S | 2 |
| | cabs | P | 1 |
| X | caps | P | 2.006 |
| | cáps | P | 4 |
| | cabet | DS | 3 |
| | caparró | DS | 40 |
| | caparronet | DS | 3 |
| | caparrot | DS | 12 |
| | caperrot | DS | 1 |
| | capet | DS | 17 |
| | cabets | DP | 1 |
| | caparrins | DP | 1 |
| | caparronets | DP | 1 |
| | caparrons | DP | 7 |
| | caparrots | DP | 2 |
| | caperrots | DP | 1 |
| | capets | DP | 8 |
| | cab | MET | 1 |
| | cabet | MET | 1 |
| | cabot | MET | 1 |
| | cap | MET | 37 |
| | caparró | MET | 2 |
| | caps | MET | 4 |

FIGURA 3. Relació de formes corresponents al lema CAP M (substantiu masculí) tal com ens les mostra el programa de consulta interactiva de la base de dades. Hi podem apreciar, degudament codificades, les formes de singular (S) i de plural (P). També hi figuren els derivats apreciatius (diminutius, augmentatius, pejoratius o intensius), diferenciats amb el codi D, i les formes que apareixen en contextos metalingüístics (MET): per a cada una de les formes gramaticals podem observar també les variants gràfiques que han aparegut en els textos. Al costat de cada una de les formes diferenciades tenim la freqüència, és a dir, el nombre de vegades que apareix en el corpus. A la capçalera de la pantalla podem observar la suma d'aquestes freqüències, és a dir, la freqüència del lema. En aquesta imatge podem observar que hem seleccionat (X) la forma **caps** P, que té una freqüència igual a 2.006, a fi d'obtenir-ne més informació.

BDTLC RECONSTRUCCIÓ SEGONS L'ORIGINAL A PARTIR D'UN LEMA

Lema: gran
Codi: 42.814

Cat. gram.: AI
Freq. lema: 56.772

Selecció de la forma

| Forma | C.M. | Freq. abs. | Forma | C.M. | Freq. abs. |
|---|---|---|---|---|---|
| G. | S | 2 | grandassos | DMP | 1 |
| g. | S | 1 | grandassots | DMP | 1 |
| gran | S | 40.585 | grandets | DMP | 19 |
| grán | S | 33 | grandissims | DMP | 2 |
| grân | S | 2 | grandíssims | DMP | 7 |
| grant | S | 24 | grandolassos | DMP | 1 |
| gran | P | 3 | grandots | DMP | 1 |
| grands | P | 11 | grandòts | DMP | 1 |
| grans | P | 15.732 | grandetes | DFP | 5 |
| gráns | P | 1 | grandísimas | DFP | 1 |
| grants | P | 11 | grandíssimas | DFP | 3 |
| regran | DS | 3 | grandíssimes | DFP | 12 |
| regrán | DS | 2 | gran | MET | 30 |
| regrans | DP | 1 | grand | MET | 4 |
| grandaçot | DMS | 1 | grandàs | MET | 5 |
| grandàs | DMS | 1 | grandassa | MET | 2 |
| grandet | DMS | 30 | grandassàs | MET | 1 |
| grandissim | DMS | 8 | grandassot | MET | 1 |
| grandíssim | DMS | 52 | grandet | MET | 1 |
| grandot | DMS | 6 | grandísim | MET | 1 |
| grandeta | DFS | 9 | grandissim | MET | 1 |
| grandísima | DFS | 2 | grandíssim | MET | 3 |
| grandissima | DFS | 5 | grandot | MET | 2 |
| grandíssima | DFS | 134 | grandota | MET | 1 |
| grandota | DFS | 2 | grans | MET | 3 |
| grandassassos | DMP | 1 | grant | MET | 1 |

FIGURA 4. Aquesta imatge ens mostra el mateix pas del procés de consulta, però per a un altre lema, el lema GRAN AI (adjectiu invariable). La relació de formes, en aquest cas, és més llarga i complexa que la del lema CAP, com es pot apreciar a simple cop d'ull. Les formes G. i g. corresponen a casos en què el mot apareix abreujat: *G. A. U.* (Gran Arquitecte de l'Univers) en una obra de Josep Torras i Bages (*¿Què és la masoneria?*, 1884) i *g. fol.* (gran foli) en una obra de Jordi Rubió i Balaguer (*Com s'ordena i cataloga una biblioteca*, 1917), respectivament.

BDTLC RECONSTRUCCIÓ SEGONS L'ORIGINAL A PARTIR D'UN LEMA

Selecció de les localitzacions de la forma

| | | |
|---|---|---|
| Forma: caps | Codi morf.: P | Freq. abs.: 2.006 |
| Lema: cap | Cat. gram.: M | Codi lema: 37.921 |

| | Obra | Niv1 | Niv2 | Localitzacions | Obra | Niv1 | Niv2 | Localitzacions |
|---|---|---|---|---|---|---|---|---|
| | 1.110 | 2 | | 15,14,9 | 1.127 | | | 34,8,6 |
| | | | | 15,18,11 | | | | 34,24,10 |
| X | | | | 17,7,3 | | | | 34,28,1 |
| | 1.110 | 7 | | 1,24,7 | | | | 34,30,8 |
| | 1.111 | 4 | | 1,25,10 | 1.128 | | | 117,16,4 |
| | | | | 3,19,5 | 1.131 | | | 51,28,11 |
| | | | | 4,13,3 | | | | 80,9,3 |
| | 1.112 | 1 | | 3,18,3 | 1.134 | | | 22,23,7 |
| | 1.112 | 2 | | 20,36,7 | | | | 29,12,7 |
| | 1.114 | 3 | | 5,1,5 | | | | 30,27,8 |
| | 1.116 | | | 33,17,2 | | | | 83,28,2 |

FIGURA 5. Aquesta imatge ens mostra el pas següent del programa de consulta, que consisteix en la relació de totes les localitzacions corresponents a les ocurrències de la forma consultada. De fet aquí veiem només un petit fragment (les vint-i-dues localitzacions que caben en una pantalla) de la relació de les 2.006 ocurrències de la forma *caps*. Per a cada ocurrència tenim la localització específica expressada de la manera següent: codi de d'obra, número de la pàgina, número de la línia i número d'ordre del mot dins la línia; en algunes obres hi figura també una referència de nivell, que és una especificació intermèdia aplicada en els casos que dins una mateixa obra recomença la numeració de les pàgines (el cas més freqüent correspon a la premsa, en què no s'ha conservat la referència a la disposició física de la informació en l'original, sinó que s'ha sotmès a una reestructuració). En aquest cas hem seleccionat (X), a fi de demanar-ne la informació contextual, l'ocurrència que es troba a l'obra 110, codi 2 de nivell 1, pàg. 15, lín. 18, mot núm. 11.

BDTLC RECONSTRUCCIÓ DEL CONTEXT SEGONS L'ORIGINAL

Autor: Publicacions Periòdiques
Títol: Tele-Estel, $$179, Barcelona, 1970.
Obra: 1.110 N.1: 2 N.2: Localització : 17,7,3

| | | | |
|---|---|---|---|
| Any d'edició: | 1970 | Forma. | : caps |
| Tipus public.: | Premsa | C. morf. | : P Plural |
| Total d'ocurrències | : 39.705 | Lema . . | : cap |
| Total de lemes usats | : 5.366 | C. Gram. | : M Nom masculí |
| Total de formes usades | : 8.916 | Codi lema: | 37.921 |

CONTEXT

«d'aquella carrera històrica era la ciutat de Berlín. En aquell moment, però, els *caps* de les democràcies occidentals. amb l'exclusió de Churchill, no s'adonaren de l'envergadura del dilema. Roosevelt, molt»

Autor: Publicacions Periòdiques
Títol: Tele-Estel, $$179, Barcelona, 1970.
Obra: 1.110 N.1: 2 N.2: Localització : 17,7,3

CONTEXT

«La qüestió alemanya
Estats Units: En el moment de començar l'ofensiva dels aliats contra l'Alemanya de Hitler.es produí una autèntica carrera entre els exèrcits anglosaxons que atacaven des de l'Oest i l'Ejèrcit roig que venia de l'Est. El premi del que arribava primer a la meta podia molt bé ésser el predomini polític a l'Europa Central. La meta d'aquella carrera històrica era la ciutat de Berlín. En aquell moment, però, els *caps* de les democràcies occidentals, amb l'exclusió de Churchill, no s'adonaren de l'envergadura del dilema. Roosevelt, molt segur d'ell mateix, i convençut de la seva força d'atracció personal, (anomenava Stalin "oncle Joe", l'oncle Pep) estava convençut que el cap soviètic no tenia ambicions europees i que només volia ésser el primer a entrar a Berlín per qüestions de prestigi nacional. La història encara no ha posat en clar si la decisió de parar l'avanç de les forces nord-americanes per a permetre l'entrada primer que ningú de l'Exèrcit roig a la capital alemanya, fou d'Eisenhower o de Roosevelt. El fet, però, és que foren els soldats russos els que entraren victoriosament al Berlín hitlerià.»

FIGURA 6. En aquesta figura podem veure dues noves etapes del procés de consulta. En primer lloc, el programa ens mostra un context breu, que correspon sempre a tres línies físiques de l'edició de referència; aquest context va acompanyat, a la capçalera de la pantalla, per una informació detallada sobre l'obra a què correspon l'ocurrència seleccionada (en aquest cas, el núm. 179 de la revista *Tele-Estel*, publicat a Barcelona l'any 1970), el grup tipològic (Premsa), l'extensió de l'obra (39.705 ocurrències), el nombre de lemes que conté (5.366) i el nombre de formes (8.916). També hi veiem les especificacions gramaticals del lema i de la forma. En molts casos el context que se'ns mostra serà suficient per a l'objectiu que motiva la consulta; si no fos així, podem avançar un pas més per a obtenir el paràgraf sencer, que veiem en la segona pantalla reproduïda: aquesta darrera forma de representació, a diferència de l'anterior, manté la reproducció exacta del format original pel que fa a la divisió de línies.

BDTLC RECONSTRUCCIÓ DEL CONTEXT SEGONS L'ORIGINAL

Autor: Francesc Curet i Payrot
Títol: Visions barcelonines 1760-1860 VIII. Muralles enllà
Obra: 485 N.1: N.2: Localització : 207,20,12

| | | | |
|---|---|---|---|
| Any d'edició: | 1956 | Forma. . | : caps |
| Tipus public.: | Ciències Socials | C. morf. | : P Plural |
| Total d'ocurrències | : 99.013 | Lema . . | : cap |
| Total de lemes usats | : 8.016 | C. Gram. | : M Nom masculí |
| Total de formes usades | : 14.055 | Codi lema: | 37.921 |

CONTEXT

«Allò que no diuen els documents fidedignes, ho supleixen la tradició i la llegenda que omplen els buits, lliguen tots els *caps* i, prenent peu d'un petit indici, munten sovint un castell imaginari que»

Autor: Francesc Curet i Payrot
Títol: Visions barcelonines 1760-1860, VIII. Muralles enllà
Obra: 485 N.1: N.2: Localització : 207,20,12

CONTEXT

«Allò que no diuen els documents fidedignes, ho supleixen la tradició i la llegenda que omplen els buits, lliguen tots els *caps* i, prenent peu d'un petit indici, munten sovint un castell imaginari que hom arriba a prendre per real. En moltes tradicions hi ha, però, quelcom de veritat, inflada i deformada en trametre's de generació en generació. De les tradicions autèntiques hom pot fer-ne cas, si les estudia amb les degudes precaucions, perquè, almenys, ens expliquen la interpretació que donaven els avantpassats als fets ignots o poc coneguts. Quan la tradició és convertida en llegenda, esdevé ja una ficció rondallística o poètica que embelleix amb l'art de la imaginació, els motius en què és inspirada. Ben al contrari de les tradicions "a posteriori", o sigui les apòcrifes que hom inventa amb exuberància alarmant, i atribueix falsament als avantpassats, les quals, a més de llur absurditat, desorienten els qui de bona fe els donen crèdit.»

FIGURA 7. Una altra mostra del mateix estadi ens ofereix una nova ocurrència de la mateixa forma *caps:* en aquest cas de l'obra *Visions barcelonines 1760-1860,* de Francesc Curet i Payrot, corresponent al volum VIII, titulat *Muralles enllà,* publicat l'any 1956, i classificat com a Ciències socials. El context ens mostra un altre possible ús d'aquest lema.

BDTLC RECONSTRUCCIÓ DEL CONTEXT SEGONS L'ORIGINAL

Autor: Emerencià Roig i Raventós
Títol: La pesca a Catalunya
Obra: 530 N.1: N.2: Localització : 116,2,2

| | | | |
|---|---|---|---|
| Any d'edició: | 1927 | Forma. . | : caps |
| Tipus public.: | Ciències Aplicad. | C. morf. | : P Plural |
| Total d'ocurrències | : 25.038 | Lema . . | : cap |
| Total de lemes usats | : 2.843 | C. Gram. | : M Nom masculí |
| Total de formes usades | : 4.325 | Codi lema: | 37.921 |

CONTEXT

«de llargada, proveïdes de surada i plom, uns petits *caps* de corda als extrems per a llegir unes xarxes a les altres. Aquestes peces són quatre; el»

Autor: Emerencià Roig i Raventós
Títol: La pesca a Catalunya
Obra: 530 N.1: N.2: Localització : 116,2,2

CONTEXT

«Quan les barques surten a pescar a les quatre de la tarda, hom diu que van de prima; quan surten a les dues de la matinada, van de matinada. Tant en un cas, com en l'altre, les embarcacions van a unes dues o tres milles mar endins, on calen. Les peces de sardinals són unes xarxes de cinc canes de llargada, proveïdes de surada i plom, uns petits *caps* de corda als extrems per a lligar unes xarxes a les altres. Aquestes peces són quatre; el llur conjunt, és anomenat calada de sardinals.»

FIGURA 8. Encara un tercer exemple de context corresponent a la mateixa forma, en aquest cas de l'obra *La pesca a Catalunya* d'Emerencià Roig i Raventós, publicada l'any 1927, i classificada dins el grup Ciències aplicades. Amb aquests tres casos posats com a mostra el lector pot fer-se una idea dels resultats que pot proporcionar una consulta sistemàtica del corpus.

BDTLC RECONSTRUCCIÓ DEL CONTEXT SEGONS L'ORIGINAL

Autor: Francesc Martínez i Martínez
Títol: Còses de la meua tèrra. (La Marina). Terça tanda i darrera
Obra: 441 N.1: N.2: Localització : 3,28,4

| | | | |
|---|---|---|---|
| Any d'edició: | 1947 | Forma. . | : cabet |
| Tipus public.: | Ciències Socials | C. morf. | : DS Aspectual singular |
| Total d'ocurrències | : 57.133 | Lema . . | : cap |
| Total de lemes usats | : 6.646 | C. Gram. | : M Nom masculí |
| Total de formes usades | : 10.735 | Codi lema: | 37.921 |

CONTEXT

«be de jòcs, ara per a dormir a aquéll; per açò coloquen el *cabet* del nin al muscle i alvançant un pèu queden les cames obertes»

Autor: Francesc Martínez i Martínez
Títol: Còses de la meua tèrra. (La Marina). Terça tanda i darrera
Obra: 441 N.1: N.2: Localització : 3,28,4

CONTEXT

«La passejadora es una gicona que no te més faena que la de aguantar el criançó, i llevant del rato que aquèst te que mamar, sempre están a la briva per els carrers juant i fent mal de cap al vehinat ab ses cantoríes be de jòcs, ara per a dormir a aquéll; per açò coloquen el *cabet* del nin al muscle i alvançant un pèu queden les cames obertes a estil de compás, i començen a engrunsarse carregant el còs en un i atre peu alternativament, cantant el consabut noni, noni..., pegantli al mateix temps suaus tronets en la esqueneta.»

FIGURA 9. Aquesta imatge ens mostra un exemple de forma amb codi D, és a dir, un derivat apreciatiu; en aquest cas la forma *cabet*, associada al mateix lema CAP. Correspon a l'obra *Còses de la meua tèrra*, de Francesc Martínez i Martínez, publicada l'any 1947 i classificada com a Ciències socials.

BDTLC RECONSTRUCCIÓ DEL CONTEXT SEGONS L'ORIGINAL

Autor: Manuel Sanchis Guarner
Títol: Gramàtica valenciana
Obra: 270 N.1: N.2: Localització : 113,32,2

| | | | |
|---|---|---|---|
| Any d'edició: | 1950 | Forma. . | : cap |
| Tipus public.: | Llengua i Liter. | C. morf. | : MET Ús metalingüístic |
| Total d'ocurrències | : 66.777 | Lema . . | : cap |
| Total de lemes usats | : 6.167 | C. Gram. | : M Nom masculí |
| Total de formes usades | : 12.648 | Codi lema | : 37.921 |

CONTEXT

«114. En posició final de paraula s'escriu generalment p; exemples: *cap*, cep, macip, glop, cup "recipient", camp, serp, colp, Calp, Asp, etc.»

FIGURA 10. Una darrera mostra, encara dins el procés de consulta de contextos, ens ofereix un exemple d'ús metalingüístic de la forma *cap*, corresponent al mateix lema CAP (substantiu masculí): es tracta de l'obra *Gramàtica catalana* de Manuel Sanchis Guarner, publicada l'any 1950, i classificada dins el grup Llengua i literatura.

## SITUACIÓ ACTUAL I OBJECTIUS IMMEDIATS

Com es desprèn del que ja he dit al llarg d'aquest article, els treballs de constitució del CTILC són molt avançats en aquest moment.[13] La part del corpus que es refereix als textos de caràcter no literari és totalment acabada. Els 29.234.544 mots de què consta han estat lematitzats i la informació que resulta d'aquesta operació ha estat incorporada a la base de dades, de tal manera que pot ésser consultada i se'n poden extreure els llistats que preveuen els programes d'explotació. El seu caràcter representatiu i equilibrat, tant pel que fa a la dimensió temporal, com pel que fa a la tipologia dels textos que configuren aquesta part del corpus, fa que puguem considerar-la un corpus acabat, que pot servir de punt de referència per a l'estudi de la llengua, en la seva vessant no literària.

La situació de la resta del corpus (la part corresponent a la llengua literària) en aquest moment és la que segueix: hi ha 18.110.247 mots introduïts a l'ordinador i revisats; d'aquests, però, només 4.245.262 han estat lematitzats. Per a acomplir les previsions del projecte manca introduir, doncs, poc més de 5.000.000 de mots i lematitzar-ne uns 19.000.000. Tanmateix, el conjunt de mots lematitzats de text literari forma part també de la base de dades i ofereix les mateixes possibilitats de consulta i explo-

(13) Per bé que les tasques preparatòries i l'inici del projecte es remunten a l'any 1985, el gruix de la tasca de constitució i el treball regular consolidat a gran escala es produí durant els anys 1989 a 1992; si s'hagués pogut continuar el treball al mateix ritme que aquests anys, el corpus s'hauria acabat a finals de 1994, però un replantejament en els fons de finançament produí una disminució primer i després una aturada del procés de constitució (introducció de nous textos i lematització). Durant aquests dos darrers anys s'ha treballat sobretot en la revisió de la base de dades, en el millorament dels programes d'explotació, a fi de poder-ne oferir la consulta externa, i en la realització de càlculs relatius a la freqüència dels elements lèxics i a la seva distribució.

tació que la resta. Sobre les perspectives d'acabament d'aquesta part del corpus, només puc dir que depenen de les contingències pressupostàries; en tot cas, però, la finalització del corpus és un objectiu prioritari.

Quant a les possibilitats d'explotació i consulta del corpus, en parlar del contingut de la BDTLC ja he fet referència a la naturalesa dels programes elaborats amb aquest objectiu i a la naturalesa de la informació que podem obtenir a través de la seva aplicació. En aquest moment, però, l'usuari del sistema de consulta té també accés a dades específiques sobre la distribució de la freqüència de cada lema en els diferents grups tipològics; aquests càlculs formen part del projecte de publicació d'uns índexs de lemes del corpus amb informació detallada sobre les freqüències i llur distribució. La informació sobre la distribució de la freqüència dels lemes en els diferents grups en què es pot dividir el corpus amb criteris de caràcter tipològic afegeix una informació qualitativament molt important a la simple freqüència entesa com el nombre absolut d'ocurrències en el corpus, o, fins i tot, a la freqüència relativa (percentatge que representa la freqüència d'un lema en relació amb l'extensió del corpus). És generalment admès que, entre dos mots de freqüència igual o similar, el que reparteix les seves ocurrències d'una manera més uniforme entre els diferents tipus de text representa millor el vocabulari fonamental d'una llengua que no pas aquell que ofereix una repartició més desequilibrada, el qual, contràriament, representa el lèxic usat amb preferència en un tipus o uns tipus de text determinats.[14] Actualment el programa de consulta ofereix, doncs, d'una manera provisional, per a cada un dels lemes del subcorpus no literari, la distribució de la seva freqüència entre els deu grups tipològics que distingim (això és el que anomenem la freqüència real); també ens ofereix altres dades, com el que anomenem freqüència previsible o teòrica, que representa el valor hipotètic de la freqüència d'un lema en cada grup tipològic en el supòsit que es repartís d'una manera completament uniforme (es tracta de la repartició de la freqüència total d'un lema d'una manera proporcional al nombre total d'ocurrències que correspon a cada grup tipològic);[15] la diferència entre la freqüència real i la teòrica és la desviació absoluta respecte de les previsions teòriques, i també apareix a la pantalla quan consultem les freqüències dels lemes del corpus (a la figura VI de l'apèndix podem observar les dades de repartició de la freqüència d'un lema i la desviació tal com apareixen en la presentació que s'ofereix actualment).

Els valors d'aquesta desviació són il·lustratius, però cal tenir en compte que no són directament comparables entre els diferents lemes, perquè es refereixen a valors absoluts de la freqüència, la qual és diferent per a cada lema (només serien comparables en lemes de freqüència igual); d'altra banda, tampoc són directament comparables els corresponents a diferents grups tipològics dins un mateix lema, perquè representen valors absoluts relacionats amb una frequència teòrica diferent per a cada grup. A fi de poder comparar les diferències en la repartició dels diferents elements lexicals podem recórrer a l'expressió de la freqüència real en termes proporcionals a

(14) Hi ha diferents propostes per a substituir la freqüència per algun altre paràmetre, o per a modificar el valor de la freqüència en funció de la distribució, a fi de ponderar millor el valor dels elements lexicals en la llengua; en aquests moments, dins els treballs del CTILC estem estudiant la manera més adequada d'avaluar aquests conceptes i de classificar la informació amb el criteri més ponderat possible.

(15) Aquest concepte substitueix el de freqüència mitjana, que utilitzaríem si el corpus tingués el mateix nombre d'ocurrències en cada grup tipològic.

la freqüència teòrica (raó entre la freqüència real i la freqüència previsible), que representa un valor proporcional referit a la unitat; tenint en compte, doncs, que, si un lema presentés una repartició uniforme entre els diferents grups, aquest valor seria 1, les variacions respecte de la unitat representen una desviació proporcional en més o en menys: els valors inferiors a 1 representen una freqüència defectiva i els superiors a 1 una freqüència sobreabundant. En l'apèndix d'aquest treball, a manera d'exemple que permet fer-se càrrec de les possibilitats d'aquestes informacions, podeu veure el valor de la freqüència proporcional d'alguns lemes del corpus per a cada grup tipològic (figura VII), i a continuació (figura VIII), una representació gràfica d'aquests valors, que permet observar d'una manera fàcilment perceptible i en termes comparables, les diferències de repartició de la freqüència dels lemes en els diferents tipus de text.

Deixant de banda les moltes i variades aplicacions de la informació de caràcter quantitatiu i les conseqüències que se'n poden derivar, faré, per a acabar, només un breu esment de les perspectives futures i de les necessitats immediates en la tasca d'anàlisi i d'explotació del corpus. Al començ d'aquest treball he mencionat l'objectiu primordialment lexicogràfic que motivà aquest projecte, i he dit aleshores que un dels principis metodològics que el justifiquen és la possibilitat d'establir les significacions dels mots a partir de l'ús, i de poder determinar les combinacions reals d'elements lèxics més freqüents en el text, i els tipus d'estructures sintàctiques en les quals cada un d'aquests es pot trobar usat. Hem vist que en la fase de lematització s'atribueix a cada una de les ocurrències que constitueixen el text unes categories (morfològiques i sintàctiques) que permeten classificar la informació des del punt de vista gramatical; d'una manera anàloga, i en una fase ja pròpiament prelexicogràfica s'ha de començar a classificar les diferents ocurrències de cada lema d'acord amb una tipologia d'usos que permeti arribar a treure'n conseqüències relatives a l'aspecte semàntic i a l'aspecte sintàctic, així com també a propòsit de les associacions recurrents d'unitats lexicals en el text; a partir d'aquí podem prefigurar l'estructura dels articles d'un futur diccionari basat realment en les dades empíriques que l'ús ens forneix. Des del punt de vista de l'organització del treball, això representa l'elaboració d'una sèrie d'eines informàtiques que permetin prosseguir l'anàlisi del corpus en les línies indicades, i emmagatzemar adequadament els resultats d'aquesta recerca per a utilitzar-los en la tasca lexicogràfica pròpiament dita, que ha de donar lloc, d'una banda, a un o a diversos diccionaris en format convencional i en forma de llibre, però també, d'una altra banda, a un gran diccionari electrònic susceptible d'ésser consultat interactivament a través d'una xarxa entrellaçada de possibilitats.

JOAQUIM RAFEL I FONTANALS
*Universitat de Barcelona*
*Institut d'Estudis Catalans*

# APÈNDIX

| | Lema | c.g. | proc. | freqüència | lema principal |
|---|---|---|---|---|---|
| | relliscant | AI | DCC | 1 | |
| | relliscar | VI | DFA | 147 | |
| | relliscós | A | DFA | 26 | |
| | relló | M | DCC | 2 | |
| | rellogament | M | DCC | 1 | |
| | rellogar | VTP | DFE | 3 | |
| | rellogat | A | DCC | 6 | |
| | rellogat | M | DFA | 8 | |
| * | rellonge | M | DCC | 14 | rellotge |
| * | rellonger | M | DCC | 1 | rellotger |
| * | rellongeria | F | DCC | 3 | rellotgeria |
| | rellotge | M | DFA | 969 | |
| | rellotger | A | DEC | 3 | |
| | rellotger | M | DFA | 53 | |
| | rellotgera | F | DFA | 1 | |
| | rellotgeria | F | DFA | 77 | |
| | rellotges | MP | DFA | 1 | |
| | relluent | AI | DFA | 5 | |
| | relluir | VI | DFA | 52 | |
| * | relonge | M | DCC | 1 | rellotge |
| * | relotge | M | DCC | 20 | rellotge |
| * | relotger | M | DCC | 2 | rellotger |
| | reluctància | F | DFA | 3 | |
| | relum | M | DCC | 1 | |
| | relumbró | M | DCC | 3 | |
| | rem | M | DFA | 336 | |
| * | rém | M | DCC | 10 | raïm |
| | remà | L | DFA | 1 | |
| | remador | M | DFA | 4 | |
| | remagencar | VT | DFA | 1 | |
| | remaire | M | DEC | 1 | |
| * | remament | M | DCC | 1 | romanent |
| | remanament | M | DCC | 1 | |
| * | remandre | V | DCC | 3 | romandre |
| | remanegar | V | DCC | 1 | |
| | remanència | F | DCC | 1 | |
| * | remanent | AI | DCC | 1 | romanent |
| * | remanent | M | DCC | 41 | romanent |
| * | remangar | V | DCC | 1 | arremangar |
| | remar | VI | DFA | 37 | |
| | remarca | F | DFA | 156 | |
| | remarcable | AI | DFA | 1.007 | |
| | remarcablement | AV | DFA | 51 | |
| | remarcadament | AV | DCC | 1 | |
| | remarcar | VT | DFA | 2.587 | |
| | remarcat | A | DCC | 34 | |
| | remat | M | DCC | 76 | |
| | rematada | F | DFA | 22 | |
| | rematadament | AV | DFA | 3 | |
| | rematador | A | DCC | 7 | |
| | rematament | M | DCC | 5 | |
| | rematant | M | DCC | 10 | |
| | rematar | VTP | DFA | 196 | |
| | rematat | A | DEC | 19 | |
| * | remate | M | DCC | 1 | remat |

FIGURA I.

*Índex general de lemes del subcorpus no literari. Ordenació alfabètica.* En aquesta mostra podem observar la informació que ens dóna aquest tipus de llistat. En la columna de l'esquerra apareixen els lemes classificats alfabèticament; els que porten un asterisc al davant corresponen a la categoria de secundaris, i cada un d'aquests té a l'extrem dret del llistat la indicació del lema principal a què està associat. A continuació del lema apareix la categoria gramatical, i al costat, el que anomenem codi de procedència (és a dir la informació sobre si el lema en qüestió és recollit en les fonts lexicogràfiques bàsiques tingudes en compte (DFA = *Diccionari Fabra*; DEC = *Diccionari Enciclopèdia Catalana*), o bé no hi és enregistrat; en aquest cas es tracta d'un lema que ha estat donat d'alta en el DBI com a conseqüència d'haver aparegut en els textos del corpus (té el codi DCC) [el codi DFE s'utilitza en els casos en què el mot es troba a tots dos diccionaris citats, però amb codificació gramatical no coincident]. La columna numèrica ens dóna la freqüència absoluta de cada lema, és a dir, el nombre total d'ocurrències que presenta. El nombre total de lemes que apareixen en aquests índexs generals del corpus no literari és de 119.000.

FIGURA II.1

| freqüència Absoluta | relativa | lema | c.g. | proc. |
|---|---|---|---|---|
| 3.069.217 | 10,764607 | el | AR | DFA |
| 2.023.094 | 7,095559 | de | PO | DFA |
| 918.865 | 3,222718 | i | C | DFA |
| 713.677 | 2,503066 | a | PO | DFA |
| 697.763 | 2,447251 | ell | P | DFA |
| 617.377 | 2,165314 | un | AR | DFA |
| 578.536 | 2,029088 | en | PO | DFA |
| 565.692 | 1,984041 | ésser | VI | DFA |
| 502.994 | 1,764141 | del | CT | DFA |
| 486.397 | 1,705931 | que | P | DFA |
| 410.940 | 1,441282 | per | PO | DFA |
| 364.405 | 1,278070 | que | C | DFA |
| 318.723 | 1,117851 | haver | VIA | DFA |
| 282.654 | 0,991347 | no | AV | DFA |
| 238.759 | 0,837394 | amb | PO | DFA |
| 222.826 | 0,781513 | al | CT | DFA |
| 202.156 | 0,709017 | aquest | A | DFA |
| 163.743 | 0,574292 | com | AV | DFA |
| 155.990 | 0,547100 | seu | A | DFA |
| 149.752 | 0,525222 | més | AV | DFA |
| 148.471 | 0,520729 | fer | VVP | DFA |
| 147.592 | 0,517646 | o | C | DFA |
| 127.452 | 0,447010 | tot | A | DFA |
| 126.302 | 0,442976 | poder | VT | DFA |
| 119.389 | 0.418730 | jo | P | DFA |
| 116.943 | 0,410152 | tenir | VTP | DFA |
| 107.359 | 0,376538 | hi | P | DFA |
| 99.632 | 0,349437 | altre | A | DFA |
| 84.314 | 0,295712 | dir | VVP | DFA |
| 78.855 | 0,276566 | si | C | DFA |
| 76.208 | 0,267282 | però | C | DFA |
| 66.269 | 0,232424 | anar | VA | DFA |
| 61.548 | 0,215866 | qual | AI | DFA |
| 60.404 | 0,211853 | nostre | A | DFA |
| 58.546 | 0,205337 | en | P | DFA |
| 58.360 | 0,204684 | mateix | A | DFA |
| 52.215 | 0,183132 | pel | CT | DFA |
| 51.862 | 0,181894 | any | M | DFA |
| 51.523 | 0,180705 | molt | AV | DFA |
| 51.419 | 0,180340 | donar | VTP | DFA |
| 50.911 | 0,178559 | estar | VVP | DFA |
| 50.324 | 0,176500 | gran | AI | DFA |
| 50.114 | 0,175763 | també | AV | DFA |
| 48.647 | 0,170618 | aquell | A | DFA |
| 46.640 | 0,163579 | dos | AN | DFA |
| 46.164 | 0,161910 | veure | VTP | DFA |
| 46.143 | 0,161836 | entre | PO | DFA |
| 44.939 | 0,157613 | quan | AV | DFA |
| 44.680 | 0,156705 | ja | AV | DFA |
| 43.499 | 0,152563 | primer | A | DFA |
| 41.871 | 0,146853 | això | P | DFA |
| 41.292 | 0,144822 | trobar | VVP | DFA |
| 40.718 | 0,142809 | perquè | C | DFA |
| 40.408 | 0,141722 | ho | P | DFA |

FIGURA II.

*Índex general de lemes del subcorpus no literari. Ordenació per freqüència (decreixent).* Oferim al lector les tres primeres pàgines d'aquest índex, en les quals pot observar els 162 lemes més freqüents en el subcorpus no literari; a la primera columna apareix la freqüència absoluta; a la segona columna, la freqüència relativa o percentatge que representa la freqüència absoluta en relació al total de mots del corpus (en aquest cas del subcorpus no literari). Les altres dades ja són conegudes (grafia del lema, codi gramatical i codi de procedència).

FIGURA II.2

| freqüència Absoluta | relativa | lema | c.g. | proc. |
|---|---|---|---|---|
| 40.095 | 0,140624 | anar | VI | DFA |
| 39.065 | 0,137011 | algun | A | DFA |
| 38.426 | 0,134770 | part | F | DFA |
| 38.409 | 0,134711 | son | A | DFA |
| 37.801 | 0,132578 | què | P | DFA |
| 37.165 | 0,130348 | home | M | DFA |
| 37.156 | 0,130316 | sobre | PO | DFA |
| 36.602 | 0,128373 | encara | AV | DFA |
| 36.480 | 0,127945 | dia | M | DFA |
| 36.548 | 0,121169 | cosa | F | DFA |
| 34.341 | 0,120443 | voler | VT | DFA |
| 33.767 | 0,118430 | ni | C | DFA |
| 33.549 | 0,117665 | sense | PO | DFA |
| 32.590 | 0,114302 | caldre | VI | DFA |
| 32.263 | 0,113155 | temps | M | DFA |
| 32.215 | 0,112987 | fins | PO | DFA |
| 31.034 | 0,108844 | manera | F | DFA |
| 30.966 | 0,108606 | tan | AV | DFA |
| 30.889 | 0,108336 | molt | A | DFA |
| 30.029 | 0,105320 | així | AV | DFA |
| 29.833 | 0,104632 | qui | P | DFA |
| 29.004 | 0,101725 | vida | F | DFA |
| 28.158 | 0,098758 | cada | AII | DFA |
| 28.023 | 0,098284 | tu | P | DFA |
| 27.933 | 0,097968 | bo | A | DFA |
| 27.669 | 0,097042 | nou | A | DFA |
| 27.307 | 0,095773 | on | AV | DFA |
| 27.255 | 0,095590 | sinó | C | DFA |
| 27.150 | 0,095222 | cas | M | DFA |
| 26.668 | 0,093532 | tant | AV | DFA |
| 26.573 | 0,093198 | sempre | AV | DFA |
| 26.543 | 0,093093 | saber | V | DFA |
| 26.197 | 0,091880 | lloc | M | DFA |
| 26.194 | 0,091869 | vegada | F | DFA |
| 25.713 | 0,090182 | forma | F | DFA |
| 25.684 | 0,090081 | posar | VVP | DFA |
| 25.439 | 0,089221 | llur | AI | DFA |
| 25.205 | 0,088401 | tot | P | DFA |
| 24.919 | 0,087397 | passar | VVP | DFA |
| 24.813 | 0,087026 | arribar | VIP | DFE |
| 24.064 | 0,084399 | déu | M | DFA |
| 23.865 | 0,083701 | des | PO | DFA |
| 23.757 | 0,083322 | obra | F | DFA |
| 23.576 | 0,082687 | poble | M | DFA |
| 23.445 | 0,082228 | català | A | DFA |
| 23.147 | 0,081183 | ben | AV | DFA |
| 22.806 | 0,079987 | deixar | VTP | DFA |
| 22.527 | 0,079008 | terra | F | DFA |
| 22.425 | 0,078650 | fet | M | DFA |
| 21.511 | 0,075445 | cap | AII | DFA |
| 20.971 | 0,073551 | ara | AV | DFA |
| 20.571 | 0,072148 | punt | M | DFA |
| 20.357 | 0,071397 | deure | VT | DFA |
| 20.158 | 0,070699 | segons | PO | DFA |

FIGURA II.3

| freqüència Absoluta | relativa | lema | c.g. | proc. |
|---|---|---|---|---|
| 19.992 | 0,070117 | general | AI | DFA |
| 19.823 | 0,069524 | estat | M | DFA |
| 19.758 | 0,069296 | aigua | F | DFA |
| 19.345 | 0,067848 | parlar | VVP | DFE |
| 19.341 | 0,067834 | hom | P | DFA |
| 19.156 | 0,067185 | casa | F | DFA |
| 18.764 | 0,065810 | senyor | M | DFA |
| 18.580 | 0,065165 | cert | A | DFA |
| 18.312 | 0,064225 | portar | VTP | DFA |
| 18.207 | 0,063857 | semblar | VI | DFA |
| 18.182 | 0,063769 | segle | M | DFA |
| 17.954 | 0,062969 | nom | M | DFA |
| 17.935 | 0,062903 | món | M | DFA |
| 17.866 | 0,062661 | tres | AN | DFA |
| 17.828 | 0,062527 | venir | VI | DFA |
| 17.817 | 0,062489 | treball | M | DFA |
| 17.679 | 0,062005 | doncs | C | DFA |
| 17.670 | 0,061973 | creure | VVP | DFA |
| 17.612 | 0,061770 | en | AR | DFA |
| 17.344 | 0,060830 | només | AV | DFA |
| 17.254 | 0,060514 | bé | C | DFA |
| 17.199 | 0,060321 | etcètera | L | DFA |
| 17.129 | 0,060076 | propi | A | DFA |
| 16.910 | 0,059308 | formar | VVP | DFE |
| 16.794 | 0,058901 | exemple | M | DFA |
| 16.746 | 0,058732 | país | M | DFA |
| 16.408 | 0,057547 | després | PO | DCC |
| 16.384 | 0,057463 | petit | A | DFA |
| 16.110 | 0,056502 | avui | AV | DFA |
| 16.021 | 0,056190 | prendre | VVP | DFA |
| 15.868 | 0,055653 | seguir | VTP | DFA |
| 15.851 | 0,055593 | durant | PO | DFA |
| 15.827 | 0,055509 | ciutat | F | DFA |
| 15.815 | 0,055467 | moment | M | DFA |
| 15.684 | 0,055008 | començar | V | DFA |
| 15.469 | 0,054254 | quedar | VIP | DFA |
| 15.320 | 0,053731 | conèixer | VVP | DFA |
| 14.901 | 0,052262 | presentar | VTP | DFA |
| 14.864 | 0,052132 | social | AI | DFA |
| 14.829 | 0,052009 | segon | A | DFA |
| 14.786 | 0,051858 | produir | VTP | DFA |
| 14.720 | 0,051627 | contra | PO | DFA |
| 14.459 | 0,050711 | sols | AV | DFA |
| 14.320 | 0,050224 | tal | AV | DFA |
| 14.310 | 0,050189 | paraula | F | DFA |
| 14.243 | 0,049954 | dins | PO | DFA |
| 14.122 | 0,049529 | diferent | AI | DFA |
| 13.967 | 0,048986 | sentit | M | DFA |
| 13.942 | 0,048898 | dret | M | DFA |
| 13.915 | 0,048803 | tractar | VVP | DFA |
| 13.887 | 0,048705 | quin | A | DFA |
| 13.859 | 0,048607 | relació | F | DFA |
| 13.779 | 0,048326 | acabar | VVP | DFE |
| 13.738 | 0,048183 | força | F | DFA |

| lema | c.g. | proc. | freqüència absoluta | freqüència relativa. |
|---|---|---|---|---|
| regularitzar | VT | DFA | 134 | 0,000469 |
| irregularitzar | V | DCC | 2 | 0,000007 |
| singularitzar | VTP | DFE | 76 | 0,000266 |
| popularitzar | VTP | DFE | 132 | 0,000462 |
| proletaritzar | VT | DEC | 6 | 0,000021 |
| desproletaritzar | V | DCC | 5 | 0,000017 |
| militaritzar | VT | DFA | 8 | 0,000028 |
| remilitaritzar | V | DCC | 1 | 0,000003 |
| sedentaritzar | V | DCC | 2 | 0,000007 |
| tartaritzar | VT | DFA | 2 | 0,000007 |
| iberitzar | V | DCC | 2 | 0,000007 |
| suberitzar | VTP | DFE | 1 | 0,000003 |
| canceritzar | V | DCC | 1 | 0,000003 |
| atmosferitzar | V | DCC | 1 | 0,000003 |
| estrangeritzar | VTP | DFA | 1 | 0,000003 |
| encoleritzar | V | DCC | 1 | 0,000003 |
| polimeritzar | VT | DFA | 1 | 0,000003 |
| caracteritzar | VTP | DFA | 2.262 | 0,007933 |
| descaracteritzar | V | DCC | 5 | 0,000017 |
| eteritzar | VT | DFA | 1 | 0,000003 |
| cateteritzar | V | DCC | 1 | 0,000003 |
| cauteritzar | VT | DFA | 11 | 0,000038 |
| parqueritzar | V | DCC | 1 | 0,000003 |
| pulveritzar | V | DCC | 49 | 0,000171 |
| vampiritzar | V | DCC | 1 | 0,000003 |
| satiritzar | VT | DFA | 25 | 0,000087 |
| martiritzar | VT | DFA | 61 | 0,000213 |
| arboritzar | V | DCC | 1 | 0,000003 |
| herboritzar | VI | DFA | 19 | 0,000066 |
| ruboritzar | VP | DFA | 10 | 0,000035 |
| teoritzar | VI | DFA | 67 | 0,000234 |
| meteoritzar | VTP | DFA | 8 | 0,000028 |
| metaforitzar | VI | DFA | 1 | 0,000003 |
| euforitzar | V | DCC | 2 | 0,000007 |
| categoritzar | V | DCC | 6 | 0,000021 |
| rigoritzar | V | DCC | 1 | 0,000003 |
| vigoritzar | VT | DFA | 58 | 0,000203 |
| revigoritzar | VT | DEC | 2 | 0,000007 |
| inferioritzar | V | DCC | 4 | 0,000014 |
| superioritzar | V | DCC | 1 | 0,000003 |
| interioritzar | V | DCC | 45 | 0,000157 |
| exterioritzar | VTP | DFE | 186 | 0,000652 |
| prioritzar | V | DCC | 9 | 0,000031 |
| majoritzar | V | DCC | 2 | 0,000007 |
| valoritzar | V | DCC | 62 | 0,000217 |
| avaloritzar | V | DCC | 1 | 0,000003 |
| revaloritzar | V | DCC | 53 | 0,000185 |
| supervaloritzar | V | DCC | 1 | 0,000003 |
| desvaloritzar | V | DCC | 19 | 0,000066 |
| cloritzar | V | DCC | 1 | 0,000003 |
| memoritzar | VT | DEC | 17 | 0,000059 |
| atemoritzar | V | DCC | 21 | 0,000073 |
| minoritzar | V | DCC | 1 | 0,000003 |
| sonoritzar | VT | DEC | 4 | 0,000014 |

FIGURA III.

*Índex general de lemes del subcorpus no literari. Ordenació alfabètica inversa.* Aquest llistat ens presenta els lemes classificats alfabèticament per la dreta. Aquest tipus de classificació, anomenat correntment *invers* o *a tergo*, té com a objectiu de poder localitzar els mots per la seva terminació: és útil per a diversos tipus d'estudi, sobretot de morfologia (flexiva o lèxica). Les altres dades que hi apareixen ja han estat comentades.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| fibrós | | A | DFA | | 1 | 0,000353 | |
| | fibrós | | MS | DFA | | 1 | 0,000353 |
| ficar | | VTP | DFE | | 32 | 0,011316 | |
| | ficar | | IF | DFA | | 12 | 0,004243 |
| | ficat | | RMS | DFA | | 5 | 0,001768 |
| | fico | | 1PI | DFA | | 1 | 0,000353 |
| | fica | | 3PI | DFA | | 1 | 0,000353 |
| | fiquen | | 6PI | DFA | | 1 | 0,000353 |
| | ficava | | 1II | DFA | | 2 | 0,000707 |
| | ficava | | 3II | DFA | | 5 | 0,001768 |
| | ficà | | 3PT | DFA | | 2 | 0,000707 |
| | ficaràs | | 2FU | DFA | | 1 | 0,000353 |
| | fiquin | | 6PS | DFA | | 1 | 0,000353 |
| | fica | | 2IM | DFA | | 1 | 0,000353 |
| ficat | | A | DCC | | 2 | 0,000707 | |
| | ficat | | MS | DCC | | 1 | 0,000353 |
| | ficats | | MP | DCC | | 1 | 0,000353 |
| ficció | | F | DFA | | 2 | 0,000707 | |
| | ficció | | S | DFA | | 1 | 0,000353 |
| | ficcions | | P | DFA | | 1 | 0,000353 |
| fictici | | A | DFA | | 2 | 0,000707 | |
| | fictici | | MS | DFA | | 1 | 0,000353 |
| | fictícia | | FS | DFA | | 1 | 0,000353 |
| fidel | | AI | DFA | | 10 | 0,003536 | |
| | fidel | | S | DFA | | 9 | 0,003182 |
| | fidels | | P | DFA | | 1 | 0,000353 |
| fidelitat | | F | DFA | | 3 | 0,001060 | |
| | fidelitat | | S | DFA | | 3 | 0,001060 |
| figa | | F | DFA | | 21 | 0,007426 | |
| | figa | | S | DFA | | 3 | 0,001060 |
| | figues | | P | DFA | | 18 | 0,006365 |
| figuera | | F | DFA | | 14 | 0,004950 | |
| | figuera | | S | DFA | | 9 | 0,003182 |
| | figueres | | P | DFA | | 5 | 0,001768 |
| figura | | F | DFA | | 33 | 0,011670 | |
| | figura | | S | DFA | | 26 | 0,009194 |
| | figures | | P | DFA | | 7 | 0,002475 |
| figurar | | VVP | DFA | | 3 | 0,001060 | |
| | figurar | | 3PI | DFA | | 1 | 0,000353 |
| | figurava | | 3II | DFA | | 2 | 0,000707 |
| fil | | M | DFA | | 13 | 0,004597 | |
| | fil | | S | DFA | | 8 | 0,002829 |
| | fils | | P | DFA | | 3 | 0,001060 |
| | filet | | DS | DCC | | 1 | 0,000353 |
| | filets | | DP | DCC | | 1 | 0,000353 |

FIGURA IV.

*Índex de lemes i formes d'una obra.* Aquest llistat ens mostra un índex de lemes amb les seves formes, però, en comptes de tractar-se d'un índex general de tot el corpus (que tindria un aspecte molt més complex), presenta la informació corresponent a una sola obra del corpus: *Camins de França* (1934), de Joan Puig i Ferreter. En el llistat apareixen els diferents lemes classificats alfabèticament (seguits del codi gramatical, el codi de procedència, la freqüència absoluta i la freqüència relativa); a continuació de cada lema hi ha les formes que li han estat associades, seguides del codi morfològic, el de procedència, la freqüència absoluta i la freqüència relativa. La classificació de les formes es fa d'acord amb una taula dels codis morfològics (el singular abans que el plural, el masculí abans que el femení, etc.).

| | | | No Lit. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | agonitzant | AI | 45 | | 18 | 8 | 5 | 2 | 1 | 7 | 2 | 2 | |
| | 52 | 0,000158 | 0,000158 | | 0,000616 | 0,000145 | 0,000146 | 0,000091 | 0,000022 | 0,000255 | 0,000091 | 0,000059 | |
| 2 | agonitzar | VI | 61 | 3 | 8 | 12 | 4 | | 1 | 6 | 18 | 9 | |
| | 80 | 0,000243 | 0,000214 | 0,000177 | 0,000274 | 0,000218 | 0,000117 | | 0,000022 | 0,000219 | 0,000825 | 0,000269 | |
| 3 | agora | F | 39 | 4 | | 3 | | | 2 | 6 | | 24 | |
| | 50 | 0,000152 | 0,000136 | 0,000237 | | 0,000054 | | | 0,000045 | 0,000219 | | 0,000718 | |
| 4 | agosaradament | AV | 11 | | 5 | 1 | 2 | | 1 | 1 | 1 | | |
| | 11 | 0,000033 | 0,000038 | | 0,000171 | 0,000018 | 0,000058 | | 0,000022 | 0,000036 | 0,000045 | | |
| 5 | agosarat | A | 118 | 5 | 5 | 27 | 19 | 6 | 10 | 9 | 22 | 15 | |
| | 170 | 0,000517 | 0,000414 | 0,000296 | 0,000171 | 0,000492 | 0,000556 | 0,000275 | 0,000226 | 0,000329 | 0,001008 | 0,000448 | |
| 6 | agost | N | 2474 | 9 | 94 | 414 | 466 | 211 | 395 | 211 | 40 | 614 | 20 |
| | 2.688 | 0,008176 | 0,008688 | 0,000533 | 0,003219 | 0,007546 | 0,013642 | 0,009704 | 0,008942 | 0,007714 | 0,001834 | 0,018375 | 0,017048 |
| 7 | agotament | M | 25 | | | 4 | 2 | | 16 | | 2 | 1 | |
| | 26 | 0,000079 | 0,000087 | | | 0,000072 | 0,000058 | | 0,000362 | | 0,000091 | 0,000029 | |
| 8 | agotar | V | 73 | 6 | 3 | 9 | 11 | 3 | 22 | 2 | 3 | 14 | |
| | 73 | 0,000222 | 0,000256 | 0,000355 | 0,000102 | 0,000164 | 0,000322 | 0,000137 | 0,000498 | 0,000073 | 0,000137 | 0,000418 | |
| 9 | agotat | A | 35 | 2 | 3 | 4 | 6 | 2 | 9 | 3 | 3 | 2 | 1 |
| | 36 | 0,000109 | 0,000122 | 0,000118 | 0,000102 | 0,000072 | 0,000175 | 0,000091 | 0,000203 | 0,000109 | 0,000137 | 0,000059 | 0,000852 |
| 10 | agraciat | A | 43 | | 10 | 5 | 9 | 7 | 3 | | 2 | 7 | |
| | 47 | 0,000142 | 0,000151 | | 0,000342 | 0,000091 | 0,000263 | 0,000321 | 0,000067 | | 0,000091 | 0,000209 | |
| 11 | agradable | AI | 1010 | 46 | 140 | 111 | 180 | 81 | 175 | 98 | 67 | 102 | 10 |
| | 1.400 | 0,004258 | 0,003547 | 0,002727 | 0,004795 | 0,002023 | 0,005269 | 0,003725 | 0,003961 | 0,003583 | 0,003072 | 0,003052 | 0,008524 |
| 12 | agradablement | AV | 39 | 2 | 2 | 2 | 8 | 2 | 11 | 3 | 2 | 6 | 1 |
| | 47 | 0,000142 | 0,000136 | 0,000118 | 0,000068 | 0,000036 | 0,000234 | 0,000091 | 0,000249 | 0,000109 | 0,000091 | 0,000179 | 0,000852 |
| 13 | agradar | VIP | 1789 | 77 | 180 | 187 | 650 | 90 | 124 | 205 | 109 | 126 | 41 |
| | 3.470 | 0,010555 | 0,006283 | 0,004565 | 0,006165 | 0,003408 | 0,019029 | 0,004139 | 0,002807 | 0,007495 | 0,004998 | 0,003770 | 0,034950 |
| 14 | agradat | A | 12 | | 2 | 4 | 1 | 1 | | 2 | 1 | 1 | |
| | 20 | 0,000060 | 0,000042 | | 0,000068 | 0,000072 | 0,000029 | 0,000045 | | 0,000073 | 0,000045 | 0,000029 | |
| 15 | agradívol | A | 35 | 2 | 4 | 5 | 1 | 12 | 2 | 2 | 3 | 4 | |
| | 59 | 0,000179 | 0,000122 | 0,000118 | 0,000137 | 0,000091 | 0,000029 | 0,000551 | 0,000045 | 0,000073 | 0,000137 | 0,000119 | |
| 16 | agrado | N | 61 | | 28 | 7 | 10 | 1 | 3 | 1 | 7 | 3 | 1 |
| | 61 | 0,000185 | 0,000214 | | 0,000959 | 0,000127 | 0,000292 | 0,000045 | 0,000067 | 0,000036 | 0,000321 | 0,000089 | 0,000852 |
| 17 | agradós | A | 222 | 5 | 16 | 61 | 19 | 18 | 16 | 36 | 22 | 25 | 4 |
| | 254 | 0,000772 | 0,000779 | 0,000296 | 0,000548 | 0,001111 | 0,000556 | 0,000827 | 0,000362 | 0,001316 | 0,001008 | 0,000748 | 0,003409 |
| 18 | agradosament | AV | 30 | 1 | 1 | 8 | 6 | 5 | | 2 | 2 | 5 | |
| | 34 | 0,000103 | 0,000105 | 0,000059 | 0,000034 | 0,000145 | 0,000175 | 0,000229 | | 0,000073 | 0,000091 | 0,000149 | |
| 19 | agraïment | N | 506 | 14 | 123 | 72 | 107 | 18 | 32 | 33 | 29 | 73 | 5 |
| | 610 | 0,001855 | 0,001777 | 0,000830 | 0,004213 | 0,001312 | 0,003132 | 0,000827 | 0,000724 | 0,001206 | 0,001329 | 0,002184 | 0,004262 |
| 20 | agrair | VI | 702 | 17 | 64 | 78 | 298 | 19 | 36 | 31 | 43 | 66 | 50 |
| | 920 | 0,002798 | 0,002465 | 0,001007 | 0,002192 | 0,001421 | 0,008724 | 0,000873 | 0,000815 | 0,001133 | 0,001971 | 0,001975 | 0,042622 |

FIGURA V.

*Index de lemes del subcorpus no literari. Ordenació alfabètica amb distribució tipològica de la freqüència.* Aquest tipus de llistat ens mostra la distribució de la freqüència de cada lema entre els diferents grups tipològics en termes de freqüència absoluta i de freqüència relativa; aquest darrer valor permet comparar la freqüència d'un lema en els diferents grups, o la de diferents lemes en un grup. Els criteris de selecció permeten limitar l'aplicació d'aquest format als lemes que apareixen en *n* grups com a mínim i/o tenen una freqüència igual o superior a *x*. La mostra que oferim aquí pertany a un llistat restringit en el qual només figuren els lemes que apareixen almenys en 5 grups tipològics i presenten una freqüència total superior a 9 (10 o més).

Lema: 14394 F cresta

| Grup tipològic | Freq. prev. | Freq. real | Desviació |
|---|---|---|---|
| 1 Filosofia | 28 | 1 | 27- |
| 2 Religió i teologia | 49 | 4 | 45- |
| 3 Ciències Socials | 93 | 22 | 71- |
| 4 Premsa | 59 | 12 | 47- |
| 5 Ciències Pures. Ciències Naturals | 37 | 99 | 62 |
| 6 Ciències Aplicades | 75 | 173 | 98 |
| 7 Belles Arts. Divertiments. Esports | 47 | 42 | 5- |
| 8 Llengua i Literatura | 37 | 10 | 27- |
| 9 Història i geografia. Biografia | 58 | 123 | 65 |
| 0 Correspondència | 2 | | 2- |
| | | Total desviació : | 1 |

Freqüència total: 562 Freqüència NO LITERARI: 486
Dispersió: 0,63426284 Ús: 308,25174024

Lema: 17671 F idea Acumulat de principal i secundaris

| Grup tipològic | Freq. prev. | Freq. real | Desviació |
|---|---|---|---|
| 1 Filosofia | 633 | 1.756 | 1.123 |
| 2 Religió i teologia | 1.102 | 1.058 | 44- |
| 3 Ciències Socials | 2.071 | 2.239 | 168 |
| 4 Premsa | 1.324 | 966 | 358- |
| 5 Ciències Pures. Ciències Naturals | 825 | 502 | 323- |
| 6 Ciències Aplicades | 1.660 | 1.076 | 584- |
| 7 Belles Arts. Divertiments. Esports | 1.042 | 1.181 | 139 |
| 8 Llengua i Literatura | 827 | 1.060 | 233 |
| 9 Història i geografia. Biografia | 1.298 | 935 | 363- |
| 0 Correspondència | 44 | 54 | 10 |
| | | Total desviació : | 1 |

Freqüència total: 12.441 Freqüència NO LITERARI: 10.827
Dispersió: 0,85991584 Ús: 9.310,30879968

FIGURA VI.

Una de les possibilitats de consulta del corpus es refereix a la informació sobre la freqüència, i, especialment, sobre la repartició de la freqüència d'un lema entre els diferents grups tipològics. Això és el que ens ofereix aquesta imatge per als lemes CRESTA i IDEA. La columna central és la que ens mostra la freqüència del lema en cada grup; la columna de l'esquerra ens mostra la freqüència teòrica o previsible (calculada repartint la freqüència total de manera proporcional al nombre total d'ocurrències que constitueix cada grup); i la columna de la dreta ens mostra la diferència entre la primera i la segona. Aquesta darrera xifra ens indica si en un grup determinat aquell lema és sobreabundant o deficitari, o bé si el nombre d'aparicions s'acosta molt al previsible en el cas que fos uniformement distribuït. Les xifres que apareixen més avall (dispersió i ús) formen part de les recerques que es fan en aquest moment per a expressar les diferències de repartició en forma de coeficient únic i per a modificar el valor de la freqüència en funció d'aquest coeficient.

| | 1. Filosofia | 2. Religió i teologia | 3. Ciències socials | 4. Premsa | 5. Ciències pures i naturals | 6. Ciències aplicades | 7. Belles arts, divertiments, esports | 8. Llengua i literatura | 9. Història, geografia, biografia | 0. Correspondència |
|---|---|---|---|---|---|---|---|---|---|---|
| **cresta** | 0.03 | 0.08 | 0.23 | 0.20 | 2.67 | 2.30 | 0.89 | 0.27 | 2.12 | 0 |
| **idea** | 2.77 | 0.96 | 1.08 | 0.72 | 0.60 | 0.64 | 1.13 | 1.28 | 0.72 | 1.22 |
| **invocar** | 1.29 | 3.32 | 1.36 | 0.46 | 0.12 | 0.28 | 0.62 | 0.61 | 0.87 | 0.66 |
| **rambla** | 0 | 0.19 | 0.84 | 3.90 | 0.22 | 0.35 | 1.13 | 0.18 | 1.19 | 0.20 |
| **escórrer** | 0.10 | 0.20 | 0.20 | 0.30 | 2.18 | 3.58 | 0.59 | 0.37 | 0.77 | 0 |
| **sanció** | 1.35 | 0.67 | 2.56 | 1.10 | 0.10 | 0.44 | 0.53 | 0.40 | 0.58 | 0 |
| **jovenesa** | 2.09 | 1.84 | 0.44 | 0.95 | 0.14 | 0.24 | 0.77 | 2.71 | 1.34 | 1 |
| **esposa** | 0.59 | 2.67 | 1.11 | 1.35 | 0.12 | 0.14 | 0.47 | 0.74 | 1.42 | 2.50 |
| **telèfon** | 0.05 | 0.02 | 0.12 | 7.47 | 0.08 | 0.19 | 0.05 | 0.08 | 0.06 | 0.08 |
| **fauna** | 0.08 | 0.05 | 0.13 | 0.32 | 9.70 | 0.20 | 0.57 | 0.13 | 0.76 | 0 |
| **rima** | 0.23 | 0.20 | 0.07 | 0.45 | 0.31 | 0 | 1.17 | 9.54 | 0.22 | 0 |
| **sarment** | 0 | 1.11 | 0.05 | 0.21 | 0.25 | 5.07 | 0.24 | 0.05 | 0.25 | 1 |

FIGURA VII.

En aquesta taula podem veure, per a una petita mostra de lemes del corpus, les diferències de repartició de la freqüència expressada de manera proporcional a la freqüència teòrica (és a dir, el valor que té respecte de la previsió si la repartició fos uniforme); aquesta forma d'expressar els valors de la desviació de la freqüència té l'avantatge de poder comparar-los entre ells, tant per grups, com per lemes. Si comparem els valors d'aquesta taula amb els que apareixen a la figura anterior (per al primer lema), observarem que realment la freqüència de CRESTA en el grup 1 és només el 0.03 de la previsió que podem fer si partim de la hipòtesi que el lema està uniformement repartit; en el grup 2, és el 0.08; en el 3, el 0.23; en el 4, el 0.2; en el grup 5, en canvi, és 2.67 vegades la previsió; en el 6, 2.3 vegades; en el 7 s'acosta a la previsió (el 0.89); en el 8, torna a ésser deficitari (el 0.27) i en el 9 és altra vegada sobreabundant (2.12 vegades la previsió). Si comparem les dades de la taula per columnes, podem observar, per exemple, que en el grup 1 (Filosofia) els mots *idea, jovenesa, sanció* i *invocar* són sobreabundants (en ordre decreixent) i la resta són deficitaris progressivament per aquest ordre: *esposa, rima, escórrer, fauna, telèfon, cresta*; *rambla* i *sarment* tenen en aquest grup una freqüència nul·la. I així podem anar comparant la resta de les dades; el judici, els comentaris i les conclusions els deixo per al lector.

FIGURA VIII.1

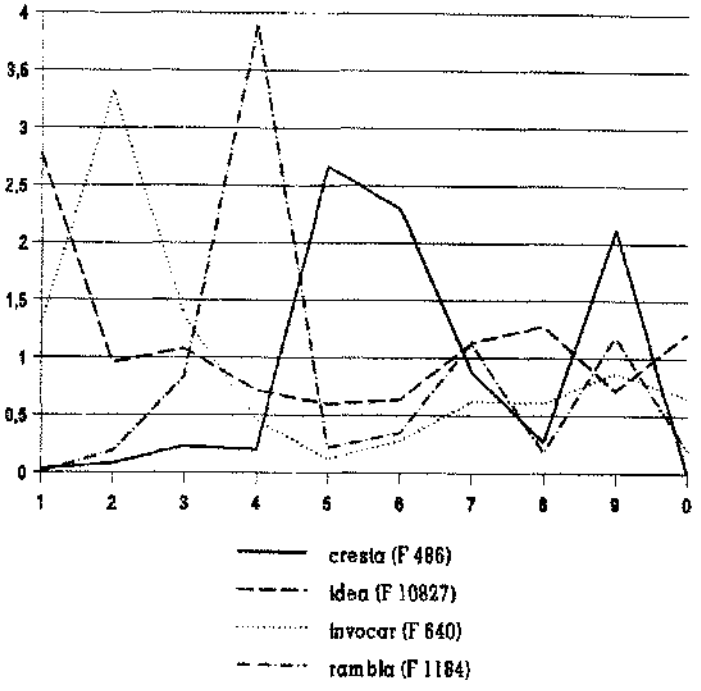



FIGURA VIII.2

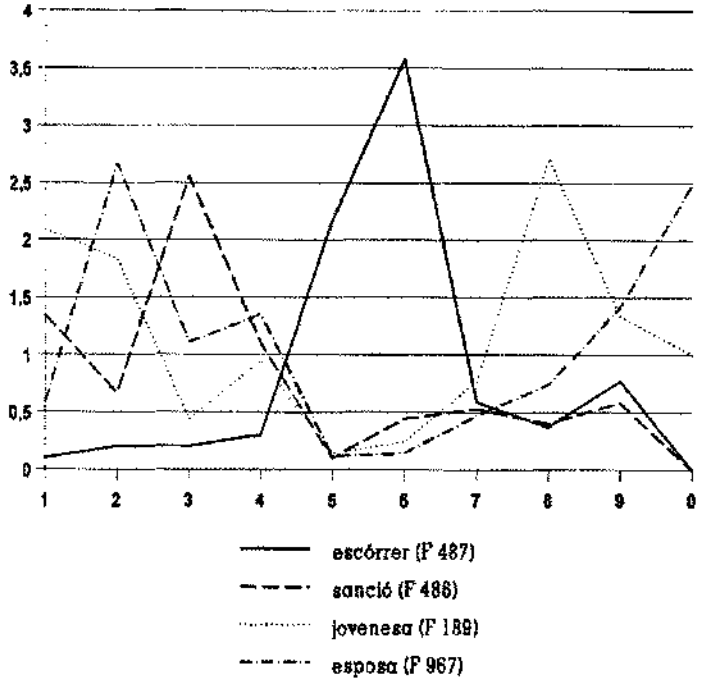
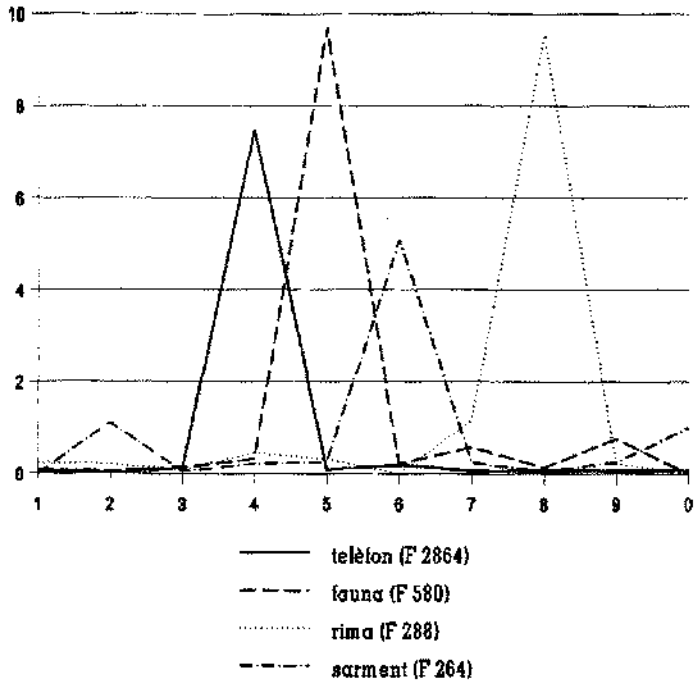



FIGURA VIII.

En aquestes imatges veiem la representació gràfica de les dades presentades en la figura anterior. Per a la interpretació d'aquests gràfics cal recordar que en el cas hipotètic que la freqüència d'un lema tingués una repartició completament uniforme entre els diferents grups, el valor proporcional de la freqüència seria 1, i, per tant, la seva representació gràfica seria una línia horitzontal situada en el valor 1 de l'escala de freqüències. Qualsevol desviació d'aquesta horitzontal hipotètica representa una variació, en més o en menys, respecte de la freqüència teòrica.

FIGURA VIII.3



CARLES DUARTE

# LA TERMINOLOGIA CATALANA ACTUAL

Dins de la bibliografia actual sobre la llengua catalana, constatem l'existència de nombroses publicacions terminològiques, que, en general, són el resultat de l'aplicació d'una metodologia moderna i que han estat elaborades amb un nivell digne de qualitat. N'hem inclòs una tria a la bibliografia que figura a l'acabament d'aquest article.

Si comparem aquesta situació amb la que presentava la terminologia catalana l'any 1980, ens adonarem tot seguit que s'hi ha produït una notable progressió.

Sens dubte ja aleshores s'havia dut a terme un treball important. D'una banda, amb algunes publicacions anteriors a la Guerra Civil, les quals mantenien encara, si més no en part, la seva vigència l'any 1980; per exemple les obres de J. Amades sobre la pesca o els pastors, el *Vocabulari jurídic* de R. Folch i Ll. Serrallonga o el *Diccionari de medicina* de M. Corachan. D'altra banda, amb els treballs publicats durant l'època franquista i els primers anys de règim democràtic, com l'*Assaig d'un vocabulari meteorològic català* d'E. Fontserè o el *Diccionari d'electrònica* de Ll. Marquet, i amb l'activitat de la Comissió Coordinadora Lexicogràfica de Ciències (1976-1982), que editava un «Full lexicogràfic». I finalment amb l'ingent esforç de sistematització i d'innovació terminològica que va representar l'elaboració i la publicació de la *Gran Enciclopèdia Catalana* (1969-1980).

La progressió a què m'he referit és el fruit sobretot d'una certa superació dels plantejaments terminològics tradicionals, basats més en el costum o en la intuïció que en l'aplicació de criteris rigorosos en aspectes com el procediment d'establiment de neologismes, la definició dels termes o la selecció i l'ordenació dels termes que han de formar part d'una obra terminològica.

*Caplletra 17* (Tardor 1994), pp. 251-266

Aquesta renovació metodològica s'ha produït com a conseqüència d'un esforç de reflexió i del manteniment de contactes amb els especialistes i les institucions que en altres països s'ocupen d'aquesta disciplina.

En aquest sentit, és representativa l'edició l'any 1984 del llibre *Metodologia de la recerca terminològica*, traducció i adaptació catalana, dirigida per M. Teresa Cabré, de l'obra *Méthodologie de la recherche terminologique* de Pierre Auger i Louis-Jean Rousseau publicada en la seva versió original per les edicions oficials de Quebec. Uns mesos abans, concretament el 28 de setembre de 1983, l'Escola d'Administració Pública de Catalunya havia organitzat una *Taula rodona sobre la lexicografia catalana i els llenguatges tècnics*, amb la participació de Joan Bastardas, M. Teresa Cabré, Xavier Lamuela, Isidor Marí i Lluís Marquet, en la qual es féu una anàlisi de la situació existent i s'esbossaren diversos projectes en matèria terminològica, que després s'han anat concretant. Dos anys i mig després, el 5 de maig de 1986, coincidint amb la celebració del II Congrés Internacional de la Llengua Catalana, la Fundació Joaquim Torres Ibern va organitzar també una *Taula rodona sobre l'ús del català científic i tècnic*, que va comptar amb la participació de M. Teresa Cabré, Enric Casassas, Joan Guitart, Joan Majó, Lluís Marquet i Carles Riba. I en el marc del II Congrés Internacional de la Llengua Catalana, l'àrea IV, de lingüística social, va dedicar força atenció als aspectes terminològics, amb una conferència de Pierre Auger, una ponència de M. Teresa Cabré i diverses comunicacions presentades per destacats experts. En aquest sentit, té un notable interès el volum corresponent a aquesta àrea de les actes del II Congrés, publicat, a cura d'Isidor Marí, per la Universitat de les Illes Balears (1992).

Així mateix, cal assenyalar la importància de l'organització de cursos de terminologia i l'esforç d'incorporació dels criteris internacionals sobre aquesta matèria. En aquesta matèria, ha estat decisiva la tasca de difusió que s'ha fet de les normes i recomanacions del Comitè Tècnic 37 de la International Standardization Organization (ISO), el qual, creat l'any 1952, ha impulsat, sota la inspiració inicial d'E. Wüster, una renovació profunda de la terminologia, que ha aconseguit una notable implantació en llengua catalana.

En el treball d'actualització i d'enfortiment de l'activitat terminològica en llengua catalana han complert una funció destacada les universitats i els diversos organismes encarregats de la normalització lingüística en els diferents territoris de llengua catalana. Ara bé, la responsabilitat central d'aquest procés ha pertocat al TERMCAT, centre de terminologia creat (1985) i sostingut pel Departament de Cultura de la Generalitat de Catalunya i l'Institut d'Estudis Catalans i vinculat a la xarxa internacional de terminologia TermNet. Recentment la Generalitat de Catalunya ha aprovat, per mitjà del Decret 47/1994, de 22 de febrer, la constitució del Consorci del Centre de Terminologia TERMACAT, integrat inicialment per la Generalitat de Catalunya, l'Institut d'Estudis Catalans i el Consorci per a la Normalització Lingüística. El TERMCAT actua com a punt de convergència de les activitats terminològiques en llengua catalana. Concre-

tament fomenta la recerca terminològica (per mitjà d'ajuts concedits per la Direcció General de Política Lingüística –l'any 1991 se'n van atorgar 17–); difon una metodologia normalitzada (destaquem aquí la publicació del llibre *Metodologia del treball terminològic* –1990– i dels documents de la col·lecció Criteris lingüístics per a la terminologia: *Manlleus i calcs lingüístics en els treballs de terminologia catalana* –1990–, *Formació de termes amb elements cultes* –1990–, *Guia d'establiment de criteris per a la denominació catalana d'animals* –1992– i *Formació del plural dels manlleus* –1992–); atén consultes externes (l'any 1992 el TERMCAT va atendre 11.795 consultes –8.289 per telèfon i 3.506 per correspondència; la xifra de consultes ateses entre 1987 i 1992 és de 45.064); fa un assessorament i un seguiment dels treballs terminològics en curs d'elaboració; ha establert un banc de dades de terminologia catalana (BTERM), concebut perquè s'hi connectin usuaris (centres universitaris,...) interessats (en dades procedents de la *Memòria del curs 1991/1992* del TERMCAT, el BTERM tenia, el dia 30 de setembre de 1992, 201.561 fitxes terminològiques); el TERMCAT té també un banc de dades documentals (BDOC) i té en projecte una base de dades de termes normalitzats (BNORM); organitza cursos de formació en terminologia; ha contribuït, per mitjà del seu Consell Supervisor, a la normalització de termes en català (des de 1986 a 1992 el Consell Supervisor del Termcat ha aprovat 4.300 neologismes); duu a terme ambiciosos programes d'elaboració i edició d'obres terminològiques en àrees temàtiques fins ara desateses o ateses només parcialment, com és el cas de la col·lecció de vint-i-nou diccionaris dels esports olímpics, que ha comportat la recopilació de 14.306 termes catalans amb les definicions i les equivalències en castellà, francès i anglès i la normalització de 1.303 termes; edita des de 1988 un «Full de difusió de neologismes», que ha esdevingut un instrument d'una gran eficàcia per a la implantació efectiva de la terminologia en llengua catalana; etc.

El Centre de Terminologia TERMCAT ha iniciat, a més, recentment (1993), en col·laboració amb l'Institut d'Estudis Universitaris Josep Trueta i la Fundació Barcelona, una col·lecció de diccionaris terminològics sobre diverses àrees del coneixement, com la biologia cel·lular, l'electromagnetisme, la sociologia o la lingüística, amb un ambiciós programa que preveu la publicació de més de cent diccionaris.

El TERMCAT va ser dirigit en una primera etapa per la professora M. Teresa Cabré i des de l'any 1988 n'és director Isidor Marí, subdirector general de Política Lingüística de la Generalitat de Catalunya. Tots dos han exercit una influència positiva en el funcionament del TERMCAT, organisme que avui constitueix una referència inexcusable en qualsevol treball sobre terminologia catalana, per la qualitat i el rigor de la seva tasca i per l'autoritat i la difusió que ha aconseguit.

El desenvolupament actual de la terminologia catalana permet, doncs, de disposar d'un mecanisme àgil de treball, que, per exemple, fa possible l'adopció de neologismes com *programari* (per *software*), *maquinari* (per *hardware*), o *ala de pendent* (per *parapente*) o *descens en bot* (per *rafting*).

Ara bé, la situació actual de la llengua catalana queda ben lluny de la normalitat a la qual aspirem. Això fa que hi hagi potser una actitud especialment amatent per part dels organismes responsables respecte a la importància de donar una solució ràpida a les necessitats terminològiques que sorgeixen de les innovacions que es difonen a partir d'altres llengües. Sorprèn, per exemple, que en altres llengües, amb una situació més consolidada, hi hagi més dificultats a definir una política terminològica reeixida. Però l'esmentada actitud amatent en el cas de la llengua catalana no és pas suficient, perquè són molts els problemes terminològics que arriben al català amb la interferència d'una altra llengua (el castellà a les terres de llengua catalana de l'Estat espanyol, el francès a la Catalunya Nord i l'italià a l'Alguer). I la convivència del català dins de la seva àrea lingüística amb una llengua de més difusió ha creat i continua creant encara avui alguns impediments a la utilització general del català en els àmbits tècnics i científics, en els quals la llengua oficial de l'Estat manté habitualment una posició preponderant.

En efecte, no n'hi ha prou que es prenguin mesures perquè existeixi un procediment àgil de resolució de problemes terminològics; cal també que funcionin amb eficàcia els mecanismes de difusió de la terminologia entre els seus usuaris naturals. Certament en català s'editen, com hem assenyalat, nombrosos diccionaris i vocabularis, i es publica l'esmentat «Full de difusió de neologismes», però és necessari que alhora, com ja s'ha començat a fer, s'estengui l'ús del català en els àmbits tècnics i científics (a les universitats –amb la publicació de manuals de les diverses disciplines en català i amb la selecció i la formació de professorat–, en el món professional,...). L'experiència d'aquests darrers anys ens demostra la viabilitat de la generalització de solucions terminològiques pròpies (per exemple en el domini del llenguatge administratiu i jurídic, amb força casos reeixits com *emplenar* 'omplir amb les dades necessàries els espais en blanc d'un imprès' o *provisió* 'tipus de resolució judicial'), però també posa en relleu els entrebancs que cal superar per aconseguir una implantació efectiva de termes catalans en àmbits professionals on el català té una presència escassa.

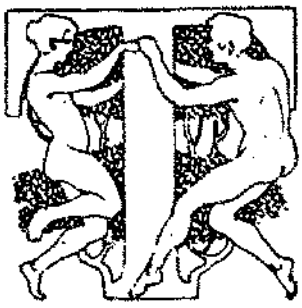
Amb aquest propòsit cal considerar un encert del TERMCAT l'impuls que ha donat a la constitució de comitès tècnics amb una representació adequada dels sectors professionals corresponents. El treball terminològic no pot ser concebut en cap cas com una activitat reservada en exclusiva a lingüistes, sinó que ha de ser el fruit de la col·laboració entre terminòlegs i experts en l'àrea temàtica objecte de treball terminològic.

Entre les publicacions terminològiques editades recentment en català, em sembla convenient de destacar dues obres monumentals per les seves dimensions i alhora modèliques des d'un punt de vista metodològic: la ja esmentada col·lecció de vint-i-nou diccionaris dels esports oficials i de demostració dels Jocs Olímpics de Barcelona (amb una extensió global de més de tres mil cent pàgines) duta a terme pel TERMCAT i editada per Enciclopèdia Catalana i el *Diccionari enciclopèdic de medicina* dirigit per Oriol Casassas, realitzat per l'Acadèmia de Ciències Mèdiques de

Catalunya i Balears, i editat també per Enciclopèdia Catalana, amb una extensió de més de mil set-centes pàgines. Totes dues obres són el fruit del treball de diversos anys d'equips integrats per terminòlegs i experts (en el primer cas, en les diferents disciplines esportives i, en el segon, metges, biòlegs, farmacèutics,...). Dins d'aquestes mateixes àrees s'han publicat recentment altres obres terminològiques o d'un gran interès terminològic. Així, per exemple, cal esmentar dins l'àmbit esportiu la col·lecció de reglaments oficials dels diferents esports promoguda per la Generalitat de Catalunya i editada per Enciclopèdia Catalana o altres diccionaris com el *Vocabulari del joc de pilota* publicat l'any 1991 per la Conselleria de Cultura, Educació i Ciència de la Generalitat Valenciana, el *Lèxic de l'esquí* editat l'any 1988 per la Conselleria d'Educació i Cultura del Govern d'Andorra o el *Vocabulari de submarinisme* editat pel Departament de Cultura de la Generalitat de Catalunya, i el *Vocabulari d'halterofília* i el *Lèxic català-anglès de futbol americà* elaborats en el marc del TERMCAT i editats també pel Departament de Cultura de la Generalitat de Catalunya. I dins l'àmbit de la medicina han aparegut altres publicacions terminològiques d'interès, com el *Vocabulari d'odontologia* (Ed. Doyma 1991) o un *Vocabulari mèdic* (Conselleria de Cultura, Educació i Ciència de la Generalitat Valenciana 1986) o el *Vocabulari de la prescripció* (Departament de Sanitat i Seguretat Social de la Generalitat de Catalunya 1991).

Dins del panorama terminològic actual en llengua catalana, cal destacar, a més, l'activitat fructífera que duen a terme les universitats. Així, la Universitat Politècnica de Catalunya ha publicat obres tan importants com el *Lèxic de la ciència del sòl* –1989–, el *Vocabulari de la indústria tèxtil* –1990– o el *Diccionari de robòtica industrial* –1991–; la Universitat de Barcelona edita una útil col·lecció de vocabularis, en què trobem, entre altres, un *Vocabulari de relacions laborals* –1991–, un *Vocabulari de la microinformàtica* –1991– o un *Vocabulari de l'ofimàtica* –1992–...

Així mateix, col·legis professionals (ja m'he referit al *Diccionari enciclopèdic de medicina*; afegim-hi altres exemples: *Terminologia dels assistents socials* –1989– publicació del Col·legi Oficial de Diplomats en Treball Social i Assistents Socials de Catalunya, el *Diccionari manual de la construcció* –1986– editat per l'Institut de Tecnologia de la Construcció de Catalunya, el *Diccionari d'assegurances* –1987– publicat pel Consell de Col·legis d'Agents i Corredors d'Assegurances de Catalunya o la participació del Col·legi d'Enginyers Industrials de Catalunya en el *Diccionari conceptual de l'electrònica* –1991– o el *Diccionari del taller mecànic* –1991–), organismes governamentals (esmentem a tall d'exemple el *Diccionari de carreteres* –1991– publicat pel Departament de Política Territorial i Obres Públiques, *Espècies pesqueres d'interès comercial. Nomenclatura oficial catalana* –1992; un excel·lent treball de sistematització d'una matèria complexa on la creativitat lingüística ha produït una notable variació dialectal– editat pel Departament de Cultura i dut a terme amb la col·laboració del Departament d'Agricultura, Ramaderia i Pesca de la Generalitat de Catalunya, i l'edició catalana del *Pla general de comptabilitat*, que inclou un útil glossari,



publicat pel Ministeri d'Economia i Hisenda i els departaments d'Economia i Finances i de Cultura de la Generalitat de Catalunya –1992–) i l'Institut d'Estudis Catalans (amb obres terminològiques com el *Vocabulari forestal* –1990– o el *Glossari de corrosió* –1991–) han realitzat una tasca admirable en el camp de la terminologia.

Una aportació destacada a la difusió de la terminologia catalana és la publicació de l'obra d'Hilari Arnau i altres *Diccionari terminològic català* (Edicions Vicens Vives 1993), concebuda com un instrument pedagògic al servei de l'ensenyament en català de les diverses àrees, amb vuit mil entrades.

A l'hora de fer un balanç de la situació actual de la terminologia catalana, cal que esmentem diverses obres que no tenen un caràcter terminològic però que tracten d'aspectes relatius al llenguatge tècnic i científic, com *Pesos, mides i mesures dels Països Catalans* de Claudi Alsina, Gaspar Feliu i Lluís Marquet –1990–, els tres volums del llibre de Lluís Marquet *Novetat i llenguatge* –1979-1985– i els excel·lents llibres *Manual de català científic* i *El llenguatge científic català* de Carles Riera –1992 i 1994–. Així mateix cal que dediquem un esment especial a la publicació l'any 1992 del llibre de M. Teresa Cabré *La terminologia. La teoria, els mètodes, les aplicacions.* Aquesta obra constitueix un clar i complet manual sobre terminologia, que inclou una síntesi brillant dels coneixements actuals en matèria terminològica, amb una detallada exposició dels aspectes més rellevants dels principis teòrics, dels procediments metodològics i de les aplicacions pràctiques de la terminologia. Aquesta obra és, a més, una aportació destacada a la bibliografia internacional sobre aquest camp. Finalment cal assenyalar la publicació recent (1993) del llibre *El llenguatge científic i tècnic* del prestigiós especialista Lluís Marquet i Ferigle, autor d'una bibliografia abundant sobre aquesta matèria. Aquest llibre de Lluís Marquet tracta amb detall i rigor d'aspectes de lèxic, de terminologia, de la neologia, de les obres lexicogràfiques, de la traducció científica i tècnica i dels antecedents i la situació actual del català científic i tècnic.

En resum, doncs, la terminologia catalana presenta a hores d'ara un balanç clarament positiu, gràcies a l'existència del TERMCAT, al nombre creixent d'obres teòriques i pràctiques sobre aquesta matèria, a la difusió d'un model de treball en aquest camp que s'ajusta als estàndards internacionals, a la formació de terminòlegs,...

Això no obstant, tot i els progressos assolits, són encara evidents les mancances que hi ha en algunes àrees temàtiques i des d'un punt de vista de la difusió i la implantació efectiva de la forma adequada dels neologismes.

CARLES DUARTE I MONTSERRAT
*Lingüista*

## REFERÈNCIES BIBLIOGRÀFIQUES

ABADAL ROCA, J. (1987) *Diccionari filatèlic*. Barcelona, Fausi.

ACADÈMIA DE CIÈNCIES MÈDIQUES DE CATALUNYA I BALEARS (1990) *Diccionari enciclopèdic de medicina*. Barcelona, Fundació Enciclopèdia Catalana.

ALEGRE, M.; LLEONART, J.; VENY, J. (1992) *Espècies pesqueres d'interès comercial: nomenclatura oficial catalana*. Barcelona, Generalitat de Catalunya.

ALEGRET, S (1977) *Diccionari de l'utillatge químic*. Barcelona, Institut d'Estudis Catalans.

ALEIX CAMP, J.; TREMOSA FITÉ, M. (1986) *Glossari de termes d'informàtica*. Andorra la Vella, Crèdit Andorrà.

ALSINA, C. (1977) *Vocabulari de matemàtica bàsica*. Barcelona, Rosa Sensat.

ALSINA, C.; FELIU, G.; MARQUET, L. (1990) *Pesos, mides i mesures dels Països Catalans*. Barcelona, Curial Edicions Catalanes.

ÁLVAREZ, J.; RIERA, C.; RAFECAS, M.; CODONY, R.; BOATELLA, J. (1993) *Diccionari dels aliments*, Barcelona, Generalitat de Catalunya.

ARAN MARTÍNEZ, J. i al. (1990) *Diccionari d'arquitectura*. Barcelona, Generalitat de Catalunya.

ARDID, R.; MATA, C.; ROMERO, J. (1988) (2a ed.) *Vocabulari de l'utillatge de cuina*. Barcelona, Generalitat de Catalunya.

ARNAU, H. i al. (1993) *Diccionari terminològic català*. Barcelona, Edicions Vicens Vives.

ARNAU REITG, R. (1989) *Lèxic terminològic de concentració de menes*. Barcelona, Universitat de Barcelona.

AUGER, P.; ROUSSEAU, L-J. (1978) *Méthodologie de la recherche terminologique*. Quebec, Éditeur Officiel du Quebec [Traducció en català (1987) *Metodologia de la recerca terminològica*, 2a. ed., Barcelona, Generalitat de Catalunya].

BADIA, A. (ed.) (1977) *Vocabulari terminològic de publicitat castellà-català*. Barcelona, Club de la Publicitat.

BALBASTRE FERRER, J. (1972) *Diccionari tècnic de l'automòbil català-castellà castellà-català*. Barcelona, Ed. Pòrtic.

BALLESTA, J.M. (1991) *Diccionari de gramàtica generativo-transformacional*. Barcelona, Universitat Autònoma de Barcelona.

BALLESTEROS, E.; GINER, M. (1989) *Vocabulari de submarinisme*. Barcelona, Generalitat de Catalunya.

BANC DE BILBAO (1978) *Terminologia bancària*. Barcelona, Banc de Bilbao.

BANCA MAS SARDÀ (1975) *Diccionari de banca i borsa català-castellà*. Barcelona, Banca Mas Sardà.

BAÑERES, E.; SEABRA, M. de; BONET, E. (1989) *Diccionari de l'esport català-castellà castellà-català*. Barcelona, Enciclopèdia Catalana.

BARDAGIL, M. (1990) *Iniciació al vocabulari de l'art*. Barcelona, Distrimapas-Telstar.
BASSEGODA MUSTÉ, B. (1972) *Diccionari de la construcció*. Barcelona, B. Bassegoda.
BATLLE GARGALLO, L. (1986) *Diccionari per a ferroviaris*. Barcelona, El Llamp.
BAULIES, J.; XURIGUERA, J.B. (1980) *Diccionari terminològic de l'administració municipal*. Barcelona, Ajuntament de Barcelona.
BESORA, J.; LLADÓ, F.; PALAU, M.; PLAZA, C.; VIVES, M. (1993) *Diccionari d'escalada esportiva*.
BLANCH BARBERÀ, E. i al. (1986) *Vocabulari telefònic*. Barcelona, CTNE.
BLAS, M. de; SERRASOLSES, J (1992) *Diccionari d'energia solar: català, castellà, francès i anglès*. Barcelona, Universitat Politècnica de Catalunya.
BLAS, M. de i al. (1991) *Diccionari de robòtica industrial: català-castellà-francès-anglès*. Barcelona, Universitat Politècnica de Catalunya.
BOZZO, M. (1990) *Vocabulari de la joieria*. Barcelona, Generalitat de Catalunya.
BROTO, P. de (1991) *Vocabulari bàsic per als jutjats*. Barcelona, Generalitat de Catalunya.
BRUCH, M. L.; PLAZA, J. (1988) *Vocabulari dels basters*. Reus, Carrutxa.
— (1986) *Vocabulari dels boters*. Reus, Carrutxa.
BRUGUERA, J. (1985) *Història del lèxic català*. Barcelona, Enciclopèdia Catalana.
BUCHON, C.; IBÁÑEZ MAS, M. (1986) *Vocabulari de ciències de la natura*. València, Generalitat Valenciana.
CABRÉ, M. T. (1992) *La terminologia. La teoria, els mètodes, les aplicacions*. Barcelona, Editorial Empúries.
— (dir.) (1984) *Vocabulari de merceria*. Barcelona, Generalitat de Catalunya.
CABRÉ, M. T.; LORENTE, M. (1991) *Els diccionaris catalans de 1940 a 1988. Barcelona*, Universitat de Barcelona.
CABRÉ, M. T.; RIGAU, G. (1986) *Lexicologia i semàntica*. Barcelona, Enciclopèdia Catalana.
CAIXA D'ESTALVIS DE CATALUNYA (1981) *Vocabulari terminològic castellà-català*. Barcelona, Caixa d'Estalvis de Catalunya.
CAMBRA OFICIAL DE COMERÇ, INDÚSTRIA I NAVEGACIÓ DE BARCELONA (1991) *Vocabulari de comerç*. Barcelona, Cambra Oficial de Comerç, Indústria i Navegació de Barcelona.
CAMPOS FRECHINA, D.; MARTÍNEZ LÓPEZ, E. (1986) *Vocabulari d'electricitat i electrònica*. Valencia, Generalitat Valenciana.
CAMPOS FRECHINA, D. (1985) *Imatge i so. Vocabulari*. València, Generalitat Valenciana.
CAMPOS PARDILLOS, M. A. (1992) *Vocabulari comercial anglés/català/castellà*. Alacant, Universitat d'Alacant.
CANDÉ, R. de (1982) *Diccionari de la música*. Barcelona, Ed. 62.
CANYAMERES, F. (1983) *Diccionari de la marina*. Barcelona, Ed. Pòrtic.

CARDONA, O. (1987) *Diccionari d'assegurances*. Barcelona, Consell de Col·legis d'Agents i Corredors d'Assegurances de Catalunya.

CARRERAS PONS, G. (1990) *Vocabulari matemàtic: català, castellà, anglès*. Vic, EUMO.

CASAS, X. i al. (1991) *Vocabulari d'odontologia: equivalències català, castellà, anglès i francès*. Barcelona, Doyma.

CASASSAS, O.; RAMIS, J. (1991) *Vocabulari de la prescripció: els mots i les locucions d'una recepta*. Barcelona, Generalitat de Catalunya.

CASTELLANOS LLORENÇ, C.; FERRÀNDIZ I CIVIL, E. (1986) (2a. ed.) *Diccionari d'informàtica*. Barcelona, Cambra Oficial de Comerç, Indústria i Navegació de Barcelona.

CASTELLANOS LLORENÇ, R. i al. (1991) *Diccionari del taller mecànic*. Barcelona, Col·legi d'Enginyers Industrials de Catalunya - La Llar del Llibre.

CERVERA I CAMINAL, A.; MERENCIANO I SALADRIGUES, J. M. (1994) *Diccionari de termes informàtics*. Barcelona, Universitat Politècnica de Catalunya.

CLIMENT COLOMA, J. J. (1985) *Vocabulari de les matemàtiques*. València, Generalitat Valenciana.

COL·LEGI D'ADVOCATS DE BARCELONA (1986) *Diccionari jurídic*. Barcelona, Enciclopèdia Catalana.

COMBALIA VILA, M. i al. (1992) *Diccionari de la indústria d'adobar pells*. Barcelona, Rafael Dalmau.

COMISSIÓ COORDINADORA LEXICOGRÀFICA DE CIÈNCIES (1977) *Vocabularis bàsics: matemàtiques, física, química*. Barcelona.

— (1977) *Vocabularis bàsics: biologia, geologia*. Barcelona.

— (1977) *Vocabularis bàsics: anatomia, bioquímica, física mèdica*. Barcelona.

— (1978-1982) *Full lexicogràfic*, 1-11, Barcelona.

COMPANY, F. J.; VILARMAU, J. M.; CERVANTES, A. (1992) *Eurodiccionari*. Barcelona, Generalitat de Catalunya.

CORACHAN, M (1936) *Diccionari de medicina*. Barcelona, Salvat.

COSTA, J. M.; CABOT, P. L. (ed.) (1991) *Glossari de corrosió: amb equivalències en anglès, alemany, francès i espanyol*. Barcelona, Institut d'Estudis Catalans.

DÍEZ QUIJANO, D. (1986) *Vocabulari d'electrònica i electricitat*. Barcelona, Arimany.

— (1985) *Vocabulari de delineació-construcció i de fusteria*. Barcelona, Arimany.

— (1985) *Vocabulari d'administració i relacions públiques*. Barcelona, Arimany.

DUARTE MONTSERRAT, C. (1993) *Llengua i administració*. Barcelona, Columna.

ESTÉVEZ, M. A.; RUBIRALTA, J. (1988) (2a ed.) *Vocabulari de perruqueria i bellesa*. Barcelona, Generalitat de Catalunya.

FABRA, P.; FONT I QUER, P.; GARGANTA, M. de (1991) *Un epistolari fonamental per a la lexicografia científica catalana (1928-1935)*. Barcelona, Institut d'Estudis Catalans.

FEDERACIÓ DE CATALUNYA DE BANCA I ESTALVI DE COMISSIONS OBRERES (1992) *Banca i estalvi: vocabulari*. Barcelona.

ESPUÑES, F.; MARQUET, L. (1982) *Vocabulari d'electrònica català-castellà-anglès*. Barcelona, Marcombo.

FERRER SANTALÓ, J. (1984) *Diccionari pràctic de comerç exterior català-anglès, anglès-català*. Girona, Cambra de Comerç i Indústria.

FLUVIÀ, A. de (1982) *Diccionari general d'heràldica*. Barcelona, EDHASA.

FOLCH, R.; SERRALLONGA, L. (1934) *Vocabulari jurídic català*. Barcelona, Col·legi d'Advocats de Barcelona.

FONTSERÈ, E. (1948) *Assaig d'un vocabulari metereològic català*. Barcelona, Institut d'Estudis Catalans.

FRANQUESA; E.; LLUCIÀ, I.; RIBA, C. (1990) (2a ed.) *Vocabulari dels electrodomèstics*. Barcelona, Generalitat de Catalunya.

FULLANA, M. (1984) (4a ed.) *Diccionari de l'art i dels oficis de la construcció*. Mallorca, Ed. Moll.

FULLAT, O. (dir.) (1984) *Diccionari de filosofia castellà-català català-castellà*. Barcelona, Ed. Vicens Vives.

FUNDACIÓ BARCELONA, INSTITUT D'ESTUDIS UNIVERSITARIS JOSEP TRUETA, TERMCAT (1992) *Diccionari de sociologia*. Barcelona, Fundació Barcelona-Termcat.

— (1992) *Diccionari de lingüística*. Barcelona, Fundació Barcelona-Termcat.

— (1992) *Diccionari d'electromagnetisme*. Barcelona, Fundació Barcelona-Termcat.

— (1992) *Diccionari de dibuix tècnic*. Barcelona, Fundació Barcelona-Termcat.

— (1993) *Diccionari d'anatomia*. Barcelona, Fundació Barcelona-Termcat.

— (1993) *Diccionari d'antropologia*. Barcelona, Fundació Barcelona-Termcat.

FUNDACIÓ JOAQUIM TORRENS IBERN (1992) *L'ús del català científic i tècnic*. Barcelona, Publicacions de l'Abadia de Montserrat.

GARCIA FRASQUET, G; LLOPIS I BAUSET, F. (1991) *Vocabulari del joc de pilota*. València, Generalitat Valenciana.

GARCIA SEVILLA, L.; GARAU I FLORIT, A. (1991) *Vocabulari de la psicologia del condicionament i de l'aprenentatge*. Barcelona, Institut d'Estudis Catalans.

GARCIA SOLER, J.; RIUS I ALCARAZ, L.; SOLER I CEREZO, E. (1991) *Diccionari de l'empresa elèctrica*. Barcelona, La Llar del Llibre.

GARCIA VALLS, D. i al. (1986) *Vocabulari d'automoció*. València, Generalitat Valenciana.

GARRETA TORNER, F. (1983) *Vocabulari gasista català-castellà*. Barcelona, Catalana de Gas i Electricitat.

GENERALITAT DE CATALUNYA (1991) *Vocabulari de bars i restaurants-bars*. Barcelona, Generalitat de Catalunya.

— (1991) *Vocabulari de restaurants*. Barcelona, Generalitat de Catalunya.

— (1993) (2a. ed.) *Menjars xinesos*. Barcelona, Generalitat de Catalunya.

— (1984) *Vocabulari de delineació per a batxillerat i formació professional*. Barcelona, Generalitat de Catalunya.

— (1982) *Vocabulari de fusteria*. Barcelona, Generalitat de Catalunya.

— (1990) *Documents dels tallers d'automòbils.* Barcelona, Generalitat de Catalunya.
— (1990) *Avaries de l'automòbil.* Barcelona, Generalitat de Catalunya.
— (1992) *Vacances a Catalunya. Vocabulari en imatges.* Barcelona, Generalitat de Catalunya.
GENERALITAT DE CATALUNYA-AENOR (1993) *Terminologia: transport per cable.* Barcelona, Generalitat de Catalunya.
GENERALITAT VALENCIANA (1986) *Vocabulari comercial.* València, Generalitat Valenciana.
— (1993) *Vocabulari taurí.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari de dret.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari d'informàtica.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari de matemàtiques.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari d'enginyeria.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari de medicina.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari d'economia.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari d'agrònoms.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari de física.* València, Generalitat Valenciana – Colomar editors.
— (1993) *Vocabulari d'arquitectura.* València, Generalitat Valenciana – Colomar editors.
GENOVARD ROSSELLÓ, C. (1980) *Vocabulario básico trilingüe de psicología científica.* Barcelona, Fontanella.
GINER ROURICH, M. (1990) *Vocabulari dels oficis d'art.* Barcelona, Generalitat de Catalunya.
GOVERN D'ANDORRA (1988) *Lèxic de l'esquí.* Andorra la Vella, Govern d'Andorra.
— (1989) *Lèxic de bars i cafeteries.* Andorra la Vella, Govern d'Andorra.
— (1989) *Lèxic de restaurants.* Andorra la Vella, Govern d'Andorra.
GUÀRDIA, M.; PUIG, J. (1977) *Vocabulari nàutic.* Barcelona, Banca Catalana.
GUMÍ, J.; LLUÍS, R. (1988) *Diccionari de tècniques pictòriques.* Barcelona, Ed. 62.
HERRANZ LLUÍS, J. i al. (1979) *Diccionari d'enginyeria elèctrica. Termes bàsics.* Barcelona, Universitat Politècnica de Catalunya.
IBORRA, R.; LLUCH, G. (1986) *Vocabulari d'informàtica.* València, Generalitat Valenciana.
ICTINEU (1979) *Diccionari de les ciències de la societat als Països Catalans (segles XVIII-XX).* Barcelona, Ed. 62.

INGLÉS, M.; ROSELL, L. (1981) *Vocabulari de geologia català-castellà castellà-català*. Barcelona, Institució Catalana d'Història Natural.

ITEC (1986) *Diccionari manual de la construcció*. Barcelona, Institut de Tecnologia de la Construcció de Catalunya.

LLINARES, À. (1986) *Vocabulari de física i química*. València, Generalitat Valenciana.

LLORET ORRIOLS, A. (1979) *Diccionari de la ciència i la tecnologia nuclears*. Barcelona, Ed. 62.

MARÍ, I. (ed.) (1992) *Segon Congrés Internacional de la Llengua Catalana IV. Àrea 3: Lingüística social*. Palma de Mallorca, Universitat de les Illes Balears.

MARQUET, L. (1971) *Diccionari d'electrònica*. Barcelona, Ed. Pòrtic.

— (1979) *Vocabulari de luminotècnia*. Barcelona, Institut d'Estudis Catalans.

— (1983) *Vocabulari de la metal·lúrgia*. Barcelona Universitat Politècnica de Catalunya.

— (1979-1985) *Novetat i llenguatge, 3 vols.*. Barcelona, Ed. Barcino.

— (1993) *El llenguatge científic i tècnic*. Barcelona, Associació Col·legi d'Enginyers Industrials de Catalunya.

MARTÍ RODRÍGUEZ, J.; GARCÍA DÍEZ, J. V. *Vocabulari mèdic*. València, Generalitat Valenciana.

MARTÍNEZ HIDALGO, J. M.; CARBONELL, L. (1984) *Vocabulari marítim català-castellà i castellano-catalán*. Barcelona, Diputació Provincial.

MARTÍNEZ ROMERO, T.; MICÓ I CATALÀ, I. (1991) *Vocabulari de mitologia*. València, Generalitat Valenciana.

MAS SOLENCH, J. M. (1981) *Diccionari breu de la sardana*. Santa Coloma de Farners, Estudis Colomencs.

MASCLANS, F. (1981) *Els noms de les plantes als Països Catalans*. Barcelona, Montblanc-Martín.

MASSUTÍ, M. (1991) *Glossari de biotecnologia, administració, ordenació i institucions pesqueres i de cultius marins*. Palma de Mallorca, Govern Balear.

MERENCIANO, F.; RICART, A. (1972) *Vocabulari automobilístic*. Barcelona, Ed. Claret.

MESTRES QUADRENY, J.; ARAMON STEIN, N. (1983) *Vocabulari català de la música*. Barcelona, Ed. Millà.

MILÀ I GALLART, R. (1993) *Terminologia ferroviària*. Barcelona, Generalitat de Catalunya.

MILLAN NESTARES, I. (1985) *Vocabulari de filosofia*. València, Generalitat Valenciana.

MOGA FERRÉS, A.; DOLCET I LLAVERIA, J. (1991) *Vocabulari de recursos didàctics i procediments educatius*. Barcelona, La Llar del Llibre.

MONÉS, J. (1987) *Diccionari abreujat de l'educació*. Barcelona, Graó.

MUMBRÚ LAPORTA, J. i al. (1990) *Vocabulari de la indústria tèxtil català-castellà*. Barcelona, Universitat Politècnica de Catalunya.

ÒMNIUM CULTURAL (1982) *Breu vocabulari tèxtil*. Mataró, Òmnium Cultural.

— (1985) *Vocabulari comercial*. Barcelona, Generalitat de Catalunya.

PALOMERO, J. (1985) *Vocabulari de llengua i literatura*. València, Generalitat Valenciana.

PARAIRA, M. (1984) *Diccionari de química*. Barcelona, Universitat Politècnica de Catalunya.

PARÉS ESPANYOL, E. (1990) *Vocabulari forestal*. Barcelona, Generalitat de Catalunya.

PLANAS, J. (1982) *Vocabulari de ràdio i televisió*. Barcelona, Instituto Oficial de Radio Televisión.

PORTA, J. i al. (1989) *Lèxic de la ciència del sòl: català-castellà-francès-anglès*. Barcelona, Universitat Politècnica de Catalunya.

RAS, E. (1990-1991) *Diccionari conceptual d'electrònica, 2 vols.*. Barcelona, Marcombo-Associació-Col·legi d'Enginyers Industrials de Catalunya.

RIERA, C. (1982) *Lèxic de l'educació en l'esplai*. Barcelona, Generalitat de Catalunya.

— (1992) *Manual de català científic*. Barcelona, Ed. Claret.

— (1994) *El llenguatge científic català*. Barcelona, Ed. Barcanova.

SALVADÓ ROVIRA, G. (1984) *Vocabulari bàsic d'educació especial*. Barcelona, Generalitat de Catalunya.

SÁNCHEZ, J. J.; RIVERA, J. J. (1992) *Diccionari de meteorologia: català, castellà, francès, anglès*. Barcelona, Universitat Politècnica de Catalunya.

SANSALVADÓ TRIBÓ, M. (1987) *Diccionari d'economia*. Barcelona, Universitat Politècnica de Catalunya.

SANTACANA, A. (1990) *Lèxic bàsic de la visita mèdica*. Barcelona, Associació Professional d'Informadors Tècnics Sanitaris.

SELFA FORT, E. (1986) *Vocabulari de fusteria, mobiliari i decoració*. València, Generalitat Valenciana.

SIGALÉS, J. M. (1984) *Diccionari nàutic*. Barcelona, Juventud.

SITJÀ, M. (1988) *Terminologia dels assistents socials*. Barcelona, Col·legi Oficial de Diplomats en Treball Social i Assistents Socials de Catalunya.

SÒRIA RÀFOLS, R. (dir.) (1989) *Diccionari Barcanova d'història de Catalunya*. Barcelona, Barcanova.

TERMCAT (a partir de 1988) *Full de difusió de neologismes*. Barcelona, Generalitat de Catalunya.

— (1990) *Bibliografia dels treballs terminològics*. Barcelona, Generalitat de Catalunya.

— (1992) *Bibliografia dels treballs terminològics: actualització*. Barcelona, Generalitat de Catalunya.

— (1990) *Manlleus i calcs lingüístics en els treballs de terminologia catalana*. Barcelona, Termcat.

— (1990) *Formació de termes amb elements cultes*. Barcelona, Termcat.

— (1992) *Guia d'establiment de criteris per a la denominació catalana d'animals*. Barcelona, Termcat.

— (1992) *Formació del plural dels manlleus*. Barcelona, Termcat.
— (1990) *Metodologia del treball terminològic*. Barcelona, Generalitat de Catalunya.
— (1992) *Cicle de conferències. La terminologia en el seu entorn*. Barcelona, octubre de 1990, Barcelona, Generalitat de Catalunya.
— (1992) *Diccionari d'atletisme*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari d'esgrima*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari d'halterofília*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari d'handbol*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari d'hípica*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari d'hoquei*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari d'hoquei sobre patins*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari de bàdminton*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de basquetbol*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de beisbol*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari de boxa*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de ciclisme*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari de futbol*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de gimnàstica*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de judo*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de lluita*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de natació*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari de pentatló modern*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari de pilota*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de piragüisme*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de rem*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari de taekwondo*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de tennis*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de tennis de taula*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de tir amb arc*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de tir olímpic*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de vela*. Barcelona, Enciclopèdia Catalana.
— (1991) *Diccionari de voleivol*. Barcelona, Enciclopèdia Catalana.
— (1992) *Diccionari general dels esports olímpics*. Barcelona, Enciclopèdia Catalana.
— (1993) *Lèxic català-anglès de futbol americà*. Barcelona, Generalitat de Catalunya.
— (1993) *Lèxic d'arts gràfiques i edició*. Barcelona, Generalitat de Catalunya.
— (1993) *Lèxic de productes alimentaris*. Barcelona, Generalitat de Catalunya.
— (1993) *Lèxic de begudes*. Barcelona, Generalitat de Catalunya.
— (1993) *Lèxic de tèxtil i confecció*. Barcelona, Generalitat de Catalunya.
— (1993) *Lèxic d'electrodomèstics*. Barcelona, Generalitat de Catalunya.

TORRA, O. (1989) *Diccionari de les professions*. Barcelona, Col·legi d'Advocats de Barcelona-Generalitat de Catalunya.

UNIVERSITAT DE BARCELONA (1993) (2a ed.) *Vocabulari de la microinformàtica*. Barcelona, Universitat de Barcelona.

— (1993) (2a ed.) *Vocabulari de les relacions laborals*. Barcelona, Universitat de Barcelona.

— (1991) *Vocabulari de la terminologia*. Barcelona, Universitat de Barcelona.

— (1993) (2a ed.) *Vocabulari de l'ofimàtica*. Barcelona, Universitat de Barcelona.

— (1993) (2a ed.) *Vocabulari del material d'oficina*. Barcelona, Universitat de Barcelona.

— (1993) *Vocabulari de reprografia*. Barcelona, Universitat de Barcelona.

— (1993) *Vocabulari de previsió social*. Barcelona, Universitat de Barcelona.

— (1993) *Vocabulari de matriculació*. Barcelona, Universitat de Barcelona.

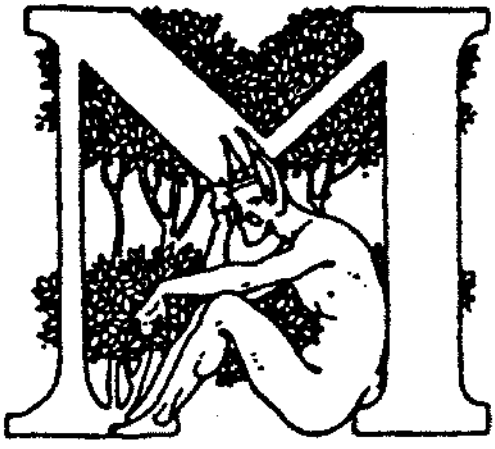
— (1993) *Vocabulari de gestió documental*. Barcelona, Universitat de Barcelona.

— (1994) *Vocabulari de formació*. Barcelona.

— (1994) *Vocabulari de comptabilitat*. Barcelona, Universitat de Barcelona.

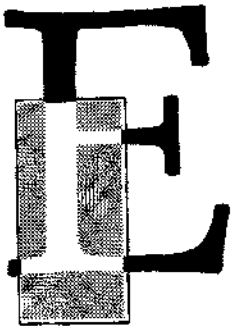
URGELL, M. D.; DURAN, C. (1988) *Vocabulari econòmico-administratiu*. Barcelona, Generalitat de Catalunya.

VILARÓ, F. (Cord.) (1991) *Diccionari de carreteres*. Barcelona, Generalitat de Catalunya.



JORDI BRUGUERA

# LA LEXICOGRAFIA BILINGÜE



Evidentment, el que hom espera de qualsevol diccionari és que li digui què vol dir una paraula determinada que no coneix. En altres mots, que li defineixi les paraules que no sap. És a dir, que li doni la definició d'alguna paraula, talment que aquesta darrera pugui eventualment ser commutada per aquella en un context donat, sense que el seu sentit global, no sols no en resulti alterat, sinó, al contrari, més ben precisat i explicat. Aquesta és, naturalment, la descripció d'un diccionari monolingüe o de definicions.

En un diccionari bilingüe, el correlat de la definició d'un diccionari monolingüe és l'equivalència de cada mot o, més àmpliament, de cada unitat lingüística de qualsevol nivell, en una altra llengua. Se suposa que cada unitat lingüística d'una llengua de partença ha de tenir un equivalent satisfactori en una altra llengua terminal. I viceversa. En certa manera, i no pas abusivament, l'equivalent de qualsevol llengua terminal o de destí ha de dir què és qualsevol significant, simple o complex, d'una llengua de partença: ha de definir què és aquella unitat. Fins a un cert punt, pot fer-ho per una mena de definició sinonímica, com les que sovint fan els diccionaris monolingües, i ben lícita també, equiparable a la més pròpia feta per gènere pròxim i diferència específica o per anàlisi de semes. Podríem parlar –ara sí, potser abusivament– d'una definició per sinònims, però amb fonètica diferent.

Però tot seguit sorgeix la gran pregunta: hi ha o hi pot haver equivalències totals i satisfactòries entre llengües? Hi ha equivalències entre tota unitat lingüística, lèxica i morfosintàctica, perífrasis i modismes? Té aquesta qüestió res a veure amb el concepte d'universal lingüístic?

*Caplletra 17* (Tardor 1994), pp. 267-272

Un tema, aquest dels universals lingüístics, que durant molt temps va desinteressar els lingüistes, perquè hi sentien un regust de massa filosofia medieval, més recentment ja és objecte de més consideració. Si tots els homes habitem un mateix planeta, bé podem esperar que hi hagi un cert paral·lelisme entre els idiomes, observa A. Martinet, un dels millors adeptes de la idea d'universals del llenguatge.[1] Malgrat que la correspondència pugui tenir alguns curts circuits en certs casos, hem d'acceptar l'existència, si més no, d'universals cosmològics, biològics i antropològics.

D'altres autors, com A. Meillet, creuen que, si hi ha algunes actituds molt generals, comunes a totes les llengües del món, es deuen al tipus mental de l'espècie humana i cal cercar-ne l'explicació en la psicologia.[2] Aquesta forma part de les dades comunes, que arreu són més o menys les mateixes pel que fa a la seva acció sobre les llengües. De fet, el llenguatge, amb varietats formals segons les diverses llengües, recorre en tots els homes a un mateix tipus de procediments, i en aquest sentit, el llenguatge n'és un. Però, més enllà d'aquests caràcters més nocionals, comuns a totes les llengües, que podem considerar com a elementals, hi ha veritables universals lingüístics en morfologia, en sintaxi, en semàntica?

Malgrat la repetida discussió sobre la correlació entre categories lògiques aristotèliques i categories gramaticals, hi ha un cert acord general a admetre que almenys es donen dos universals lingüístics en els idiomes: el nom i el verb, la designació de les coses, concretes i abstractes, i l'expressió dels estats i esdeveniments d'aquestes coses. Per bé que aquesta categoria verbal pugui expressar-se també formalment mitjançant una expressió nominal.[3] Afinant més, trobaríem també en la majoria de llengües altres elements comuns, com els determinants i els relacionants.

Molts autors, preocupats pel que creuen diferències irreductibles entre les llengües, accepten que el contingut latent de tot llenguatge és el mateix, i és el coneixement intuïtiu engendrat per l'experiència.[4] Les categories semàntiques, molt sovint, són considerades com a universals.[5] Bloomfield admetia també expressament l'existència d'universals semàntics, perquè, segons ell, la qüestió pràctica de saber quines coses poden ser dites en llengües distintes, es confon sovint amb la de les significacions de les paraules i de les categories. Una llengua usarà una frase on una altra usa una paraula, o bé on una tercera usa una forma composta. Una significació que disposa d'una categoria lingüística per a expressar-se en una llengua (per exemple, la pluralitat dels objectes en anglès) pot no aparèixer només sota estímuls pràctics limitats en una altra llengua. Però en el que afecta la denotació, qualsevol cosa que es pugui dir en una llengua, sens dubte es pot dir en una altra.[6]

Aquesta mena de predisposició de les llengües a deixar-se transportar en d'altres s'accentua potser entre llengües que històricament han anat compartint sempre més i més les pròpies riqueses culturals, fins arribar a una certa comunitat de cultura, per exemple europea. Les paraules no poden ser compreses correctament, separades dels fenòmens culturals que expressen. La majoria de les traduccions amb què solem es-

(1) A. Martinet (1950), «Réflexions sur le problème de l'opposition verbo-nominale», a *Journal de Psychologie normale et pathologique*, Paris, 1, pàg. 104.

(2) A. Meillet (1938), *Linguistique historique et linguistique générale*, II, Paris, Klincksieck, pàg. 16.

(3) Benveniste (1958), «De la subjectivité dans le langage», a *Journal de Psychologie normale et pathologique*, 3, pàg. 261-262.

(4) E. Sapir (1953), *Le langage*, Paris, Payot, pàg. 20-21.

(5) L. Hjelmslev (1949), «Existe-il des catégories qui soient communes à l'universalité des langues humaines?» a *Actes du 6e C.I.L.*, Paris, Klincksieck, pàg. 421.

(6) L. Bloomfield (1955), *Language*, 2 ed., London, Henderson and Spalding, pàg. 139-140.

tar familiaritzats han estat realitzades dins la família lingüística indoeuropea, i, en la seva major part, la cultura d'aquest domini lingüístic és relativament homogènia.[7]

Per tant, tot i admetent plenament totes les diferències entre els sistemes lingüístics, hem d'admetre que hi ha una quantitat tan important i extensa de trets universals comuns a totes les llengües, que el pas de qualsevol llengua a qualsevol altra és possible, si més no, en l'àmbit dels universals. És aquí on almenys és possible de trobar l'equivalència lingüística entre dues llengües.

## MACROESTRUCTURA

Un diccionari monolingüe qualsevol sol proposar un extret normalitzat del conjunt del lèxic real. Difícilment un diccionari monolingüe pot pretendre d'oferir un recull exhaustiu de tot el lèxic d'una llengua. El lexicògraf, fent una tria en el lèxic total, transforma l'inventari obert d'una llengua en un conjunt clos, el dels mots que figuren en el diccionari, per més ampli que aquest pugui ser i per més que aquest es pugui anar eixamplant segons les exigències dels usuaris.

Com que aquestes poden i solen ser variades, segons les diverses necessitats, el lexicògraf es troba en la dificultat de la tria del material que ha de constituir el diccionari: un recull màximament complet per a un conjunt el més ampli possible de consultants o un aplec reduït, selectiu, destinat a un cercle de lectors menys exigent o menys necessitat, per exemple, un diccionari escolar, on hi haurà només un vocabulari bàsic amb els seus corresponents modismes també bàsics. Més delicada la tria en aquest cas.

Un diccionari no és una obra que depengui directament de la lingüística com a ciència del funcionament del llenguatge, ja que tot diccionari és, per naturalesa, metalingüístic. Tracta dels mots fora de la parla, del discurs. És un instrument que entra en el joc de les relacions entre llengua i parla. Hom pot jutjar del grau de conformitat d'un diccionari amb una teoria lingüística: el diccionari depèn de la lingüística aplicada. Més enllà de la permanència aparent de l'estructura didàctica tradicional del diccionari, nombroses innovacions poden traslluir, d'una manera més o menys sistemàtica, l'evolució de la teoria lingüística. En funció d'aquesta, en l'anàlisi semàntica i la classificació dels sentits de les paraules, hom ha pres en consideració les relacions de la unitat en l'estructura de l'enunciat, en lloc d'analitzar-la només com a unitat de llengua abstracta, tot distingint els usos diferents segons les relacions sintagmàtiques, sobretot pel que fa als verbs.

En l'elaboració de la definició, s'ha aplicat l'anàlisi de l'estructura semàntica en unitats jerarquitzades, cosa que ha implicat una distinció i una classificació més sistemàtiques i més coherents dels sentits en funció de la presència o l'absència d'un o més semes en relació a un ús precedent. S'han reconegut més netament els diferents

(7) E. A. Nida (1945), «Linguistics and ethnology in translation problems». *Word*, 2, pàgs. 207 i 194.

sememes per a un mateix significant. Aquesta nova perspectiva en l'anàlisi semàntica, combinada amb els criteris funcionals saussurians sobre l'oposició entre diacronia i sincronia, ha tingut per corol·lari la distinció més rigorosa entre els usos envellits i els usos vigents i habituals en el lèxic contemporani, ja sigui per l'eliminació dels primers en un diccionari estrictament sincrònic, ja sigui per la marca d'un ús no contemporani. Els nivells del discurs solen ser netament marcats, talment que el lèxic descrit, en lloc de ser una suma d'unitats d'un estatut lingüístic homogeni en l'abstracció de la llengua, esdevé un mosaic de termes trets de diferents sociolectes (popular, familiar, literari, dialectal, argot, tècnic, científic, etc.).

El punt de vista funcional fa reconèixer als mots un estatut diferenciat sistemàticament en funció de les oposicions de categories gramaticals (verb/adjectiu, verb transitiu/intransitiu, masculí/femení, singular/plural). La definició de la unitat lexical segons l'anàlisi lingüística contemporània ha introduït en la nomenclatura unitats sintagmàtiques, si no en forma d'entrades principals, almenys en forma d'entrades secundàries. Es pot, doncs, afirmar que la descripció lexicogràfica contemporània ha estat àmpliament penetrada pels conceptes de la lingüística moderna, sota formes més o menys sistematitzades, segons la concepció de cada diccionari. Revela l'influx de la lingüística teòrica tot allò que afecta el metallenguatge i, per tant, la teoria posada en pràctica en l'anàlisi de les dades lexicològiques de llurs relacions sintagmàtiques o paradigmàtiques, la delimitació dels sentits que en deriva, la notació de les condicions gramaticals, semàntiques i sovint també estilístiques d'ús.

Totes aquestes consideracions, referides a un diccionari monolingüe, de definicions, poden ser aplicables proporcionalment a un diccionari bilingüe, particularment pel que fa a la tria i la jerarquització dels lemes amb les diverses accepcions.

Tenint present un diccionari bilingüe no pas restringit, el nombre d'entrades, més o menys extens, sol dependre en bona part, i en primer lloc, de la major o menor oferta lexical de la llengua d'origen a causa de l'interès que pot suscitar en la transposició a la llengua terminal.

Un diccionari bilingüe pot beneficiar-se més o menys de les trajectòries lexicogràfiques de cada llengua confrontada i acollir amb més o menys generositat dades que pròpiament pertanyen al diccionari de definicions, com són les diverses accepcions possibles i l'ampli camp de la fraseologia. Així evita el perill de tot diccionari bilingüe: el recurs fàcil i còmode –a vegades també per a l'usuari– a l'equivalència única, amb el greu risc de caure en errors greus de traducció.

Però també ací se suscita una altra qüestió interessant. Un diccionari bilingüe, que pretén de ser generós, sobretot a reflectir la llengua viva, ha de procurar incorporar el major nombre de mots del cabal lèxic de la llengua d'origen indiscriminadament? El major nombre de mots de la llengua «viva», ni que es tracti també de barbarismes o vulgaritats? Com cal resoldre la dialèctica entre terme «viu», real en la parla, i la seva acceptació per la normativa o el (difícil!) bon gust?[8] Són conciliables l'exigència

(8) Sempre són il·luminadores –i per això dignes de ser rellegides de tant en tant– les paraules de Carles Riba en el seu prefaci a la segona edició del diccionari Fabra, sobre el difícil problema de la selecció de mots en un diccionari normatiu, i concretament les que dedica al bon gust (pàg. XV).

de reflectir la realitat d'una llengua viva a un seu possible usuari estranger, que hi pot topar pel carrer, i la bona intenció d'oferir a aquest estranger el més genuí de la llengua amb què s'enfronta, que també és realitat autèntica? En altres mots, tot el que és un fet de parla ha de figurar sempre en un diccionari de llengua, monolingüe o bilingüe?[9]

## MICROESTRUCTURA

L'assumpció d'aquesta mena de dilema (esquemàticament: realitat/genuïnitat) amb què s'ha d'enfrontar un diccionari bilingüe afecta també per derivació el tractament de la seva microestructura. La gran realitat d'una llengua determinada, més encara quan es presenta i s'ofereix als usuaris d'una altra, consisteix sobretot –més que en el fet de mostrar-se trufada amb tota mena d'expressions incontrolades que semblen fer-la més autèntica– en l'extraordinària riquesa de construccions, locucions i modismes en què els mots solen revestir la seva variada força expressiva. D'aquí que se solgui esperar d'un bon diccionari bilingüe un bon cabal del lèxic més usual i genuí en els diversos registres, amb un bon esmicolament semàntic de cada mot, desplegat en el major nombre de sentits i matisos, incorporats en la fraseologia pròpia de cada llengua. Només així un usuari d'un bon diccionari bilingüe hi podrà veure satisfetes, segons cada moment, les seves necessitats semasiològiques i onomasiològiques.

És molt important l'afinament de la informació gramatical sobre cada lema, sobretot en els casos ben freqüents en què les categoritzacions gramaticals o els règims verbals no es corresponen exactament en cada llengua.

Els qui no obliden que la funció primordial d'una llengua és de ser parlada abans que de ser escrita agraeixen de trobar en un diccionari bilingüe la pronunciació figurada de la llengua d'origen. I de tots els sistemes de transcripció fonètica, ara com ara, i admetent tot el que calgui admetre, el millor i més fàcilment practicable és el de l'Associació Fonètica Internacional. Sempre acaba essent més econòmic de recórrer-hi que no pas d'inventar-ne de nous, intencionadament acomodats a un cert públic restringit.

Com en tants àmbits de les manifestacions de l'esperit humà, és impossible d'aconseguir una satisfacció i un perfeccionament plens en l'elaboració d'un diccionari, tan monolingüe com bilingüe. I la raó segurament es troba en les mateixes característiques del lèxic de tota llengua: que sigui, per principi, d'un inventari il·limitat i d'una mutabilitat constant. I aquesta realitat, en un diccionari bilingüe, es dóna per igual en cada llengua, però amb modalitats diferents segons les seves circumstàncies. Alhora, però, cal tenir en compte l'influx cada vegada més creixent entre llengües, sobretot entre algunes, tant en el camp de l'expressió com del contingut.

Aquest estat de coses, implica un esforç constant de revisió dels diccionaris bi-

(9) Sobre tot aquest tema m'he servit també en general –i les recomano– de totes les obres següents: G. Haensch, L. Wolf, S. Ettinger, R. Werner (1982), *La lexicografía. De la lingüística teórica a la lexicografía práctica*, Madrid, Gredos, particularment els capítols 6.4 i 9.4; R. Werner, C. Chuchuy (1992), «¿Qué son los equivalentes en el diccionario bilingüe?», a *Estudios de lexicología y metalexicografía del español actual*, Compilación y edición: Gerd Wotjak, Tübingen, Max Niemeyer Verlag, pàg. 99-107 (amb abundosa bibliografia); G. Mounin (1971), *Los problemas teóricos de la traducción*, Madrid, Gredos; M. Seco (1987), *Estudios de lexicografía española*, Madrid, Paraninfo.

lingües per a mantenir-los aptes per a facilitar la traducció d'una llengua a una altra, en una doble direcció. En definitiva, aconseguir d'entendre bé què vol dir una expressió en una llengua determinada i mirar de dir-ho en una altra en la mateixa proporció de significant a significat. Comesa delicada i difícil, però no impossible, quan més es va penetrant en la idiosincràsia de cada sistema lingüístic en qüestió. La comunicació entre parlants de llengües diferents és possible, fins i tot, potser, molt satisfactòria, i no pas menys, a vegades, que la que es dóna entre parlants d'una mateixa llengua, que tampoc no sempre s'entenen o s'interpreten ben bé del tot. Fatalitat, en tot cas, de la no perfecta adequació entre signes i continguts de significació entre els humans. No ens és possible el contacte directe dels pensaments, la intuïció perfecta. No podem eludir la mediatització. I, ben mirat, potser sortosament.

JORDI BRUGUERA

JOAN SOLÀ

# LA LEXICOGRAFIA GENERAL I LA NORMATIVA

## 0. OBJECTE D'AQUEST ESCRIT*

Les dificultats i les preocupacions actuals dels usuaris corrents de la llengua i dels especialistes que s'hi dediquen de manera aplicada (serveis lingüístics) són pràcticament il·limitades i afecten els aspectes més pròpiament gramaticals i lexicogràfics, però també altres aspectes heterogenis (tipogràfics, estilístics, etc.): entre els últims hi ha sobretot els que afecten les relacions del català amb altres llengües (§§ 1.1-2). Hauria pogut intentar de seleccionar per a aquest article les qüestions estrictament relacionades amb la lexicografia, però se m'ha fet difícil i, d'altra banda, he preferit de no excloure totalment els altres terrenys, perquè es tracta d'una realitat social de gran interès.

Intentaré, doncs, de fer veure quina part d'aquesta múltiple matèria preocupant podria ser acollida, i com podria ser-ho, en un «diccionari general» i quina part en quedaria exclosa. No insisteixo gaire, però ho dic clarament des d'ara, que un diccionari general no hauria de pressuposar cap intuïció en els lectors: és a dir, que hauria de servir exactament igual a un catalanoparlant que a un foraster que aprèn la llengua.

El títol que m'han proposat per a aquest treball inclou els conceptes de «lexicografia general» i «normativa»: he de suposar que se'm demana que els relacioni. Ho faré intentant de precisar quina part de la matèria que faré servir seria (més) pròpia d'un codi normatiu i quina en seria aliena (§ 2.3).

* Aquest treball forma part del projecte «Variació en el llenguatge» de la Universitat de Barcelona, finançat per la CIRIT (CS93-1017) i per la DGICYT (P90-0505).

Com que durant els últims decennis hi ha hagut un concepte de «normativa» excessivament lax i sovint equivocat i desorientador, aprofitaré marginalment l'ocasió per dir com s'hauria d'entendre aquest concepte (§ 1.3). Però diré també dues paraules sobre un aspecte de la qüestió que és totalment verge: quina matèria d'una llengua hauria de ser considerada «normativa» (§ 2.3*c*).

## 1. NORMATIVA I NECESSITATS REALS DELS USUARIS

### 1.1. El dubte lingüístic sistemàtic, característica de la nostra societat

És una característica palesa de la nostra societat el dubte, la vacil·lació i la inhibició sistemàtics en matèria de llengua. La causa originària d'aquesta situació foren les circumstàncies polítiques i socials; i els efectes d'aquella primera causa, algun dels quals ha esdevingut nova causa al seu torn, no són fàcils de reduir a esquema, però s'hi han de comptar almenys els següents: inseguretat general en l'ús de la llengua (cada vegada més greu, a causa de la barreja lingüística ambiental i d'una pressió de la llengua castellana infinitament superior a la que havia exercit mai), por generalitzada a decidir res personalment sobre aquesta llengua, manca d'actuació de l'autoritat lingüística i algun cop actuació excessiva d'aquesta autoritat.

Pot estranyar que parli d'inhibició sistemàtica, després d'uns anys en què hi ha hagut considerables batusses lingüístiques. Hem de tenir en compte que aquestes batusses han estat protagonitzades per mitja dotzena mal comptada de persones, i jo ara em refereixo a l'àmplia xarxa de serveis lingüístics, formada per centenars de professionals; a part que l'esmentat enrenou no ha estat de cap manera un diàleg.

El fenomen a què em refereixo consisteix en el següent. Avui els professionals dels serveis lingüístics pateixen de dubte i d'inseguretat personals sistemàtics sobre qualsevol aspecte de la llengua i no es tranquil·litzen fins que troben la resposta a cada dubte en algun indret fora d'ells mateixos: principalment els llibres d'estil i els telèfons lingüístics (de la Generalitat de Catalunya, del Termcat i de molts altres llocs). Inversament, es torna de manera clara a una actitud que anys enrere va produir qui-sap-los estralls: es desestima (fins i tot amb angoixa) qualsevol cosa que no es trobi en els indrets referits. Finalment, les «respostes» que es troben en aquests llocs, o en documents de l'Institut, no van acompanyades de cap actitud crítica constructiva: ningú no discuteix res, moltíssimes persones ni tan sols reflexionen un moment sobre una novetat descoberta allà on sigui, però tothom considera aquesta novetat com a cosa bona, segura i millor que la possible alternativa, i en conseqüència tothom corre a aplicar aquesta novetat, no fos cas que algú els descobrís el llautó. Si es veu al metro de Barcelona, per exemple, un avís que diu «No es permet de fumar a les instal·lacions *de metro*», la primera reacció és de pensar que «deu ser» *de metro* i no pas *del metro*,

i tot seguit es corre amb angoixa a remenar diccionaris, llibres d'estil i gramàtiques per localitzar-hi el que no és altra cosa que un lapsus evident.

En l'estat actual de la llengua i de la societat, opino que seria molt convenient per a tots que les novetats, sobretot les de l'autoritat, anessin acompanyades de raonaments adequats (els quals, evidentment, podrien figurar en lloc diferent d'aquell en què hi ha la novetat). Aquesta mesura, d'una banda, augmentaria la seguretat personal en els professionals, que posseirien raons a més a més de respostes concretes i s'acostumarien una mica a veure que moltes coses no són ni absolutes ni segures; i, d'una altra banda, permetria que tal o tal proposta pogués ser sospesada i assajada abans de rebre la sanció definitiva.

Qui hagi conegut per dins els actuals serveis lingüístics acceptarà que és real en gran part el panorama descrit. No sé quina solució global efectiva podria tenir aquesta situació, però la mesura que apunto podria contribuir-hi. Si dels professionals dels serveis lingüístics passem als simples usuaris, el panorama és molt més alarmant. El gran perill és que la nostra societat estigui formada per una immensa majoria d'usuaris que no tenen (o no es reconeixen o no se'ls reconeix) competència lingüística i per uns centenars de professionals que actuen a cop de diccionari, de llibre d'estil i de telèfon lingüístic. La llengua pot acabar encaixonada, mil·limetrada, absolutament uniforme en aquests llocs comuns de saviesa. Qui no s'hi hagi fixat, ha de saber que en aquests llocs de consulta hi ha qualsevol inimaginable dubte possible: des del plural de paraules com ara *tanca* (que deu ser, doncs, d'una «irregularitat» difícil de retenir) fins a tòpics seculars com l'esmena d'*abogat* en *advocat*, passant per la qüestió de saber si després de dos punts «és obligatori» o no de començar amb majúscula o si és una falta greu d'escriure «carrer Balmes» en lloc de «carrer de Balmes».

### 1.2. Les necessitats actuals dels serveis lingüístics

El panorama de vacil·lació universal es reflecteix, doncs, en els llibres d'estil i serveis telefònics de tota mena, els quals acaben recollint una heterogènia quantitat de materials que van des d'aspectes estrictament lèxics (neologismes, tecnicismes, acceptabilitat d'una paraula determinada) o gramaticals (si és o no correcta la construcció *no obstant* tota sola o com és el femení de *jutge*), fins a aspectes diversos d'ortografia (quan és que l'article *la* no s'apostrofa), de tipografia (com s'abreugen els dies de la setmana, si cal apostrofar o no davant una cursiva), d'onomàstica (si les comarques duen o no duen un article inicial i si aquest article s'escriu en majúscula o en minúscula), de relacions humanes (què s'ha de fer per evitar que un escrit sembli adreçat exclusivament a homes) o d'administració (quina forma física i lingüística té un saluda, quin tractament té un alcalde i si *alcalde* va en majúscula o en minúscula).

Entre totes aquestes matèries, la que representa el noranta per cent de les necessitats

dels usuaris d'aquests materials són els dubtes lèxics, fraseològics i de certs detalls sintàctics. I la immensa majoria de les vegades es tracta de saber si es pot dir o com s'ha de dir en català una determinada forma (lèxica, fraseològica o sintàctica) castellana. M'interessa arribar aquí: resulta que la immensa majoria de problemes pràctics dels actuals serveis lingüístics o dels usuaris corrents *no* poden ser resolts per un «diccionari general», perquè són detalls de comparació amb una altra llengua.

### 1.3. Concepte restrictiu de «normativa»

De manera que la realitat sobre els llibres de consulta ha canviat radicalment. I aquesta realitat pot fer que persisteixi la interpretació excessiva del concepte de normativa a què tot seguit em referiré: al cap i a la fi, en la pura pràctica, «és normativa» tot allò que amb una certa autoritat se'ns ha dit que és així o això. La primera autoritat és el professor d'una classe de llengua, d'un seminari, d'un curset de preparació per a uns exàmens o uns exercicis d'oposició; la segona, el manual que es fa servir (o, pitjor encara, els apunts) i que per a molts és l'única cosa que queda quan el professor desapareix. En definitiva, doncs, som una societat de manuals pràctics. Aclarim que també altres cultures tenen i han tingut muntanyes d'aquests manuals, però hi ha una diferència fonamental: només en la nostra hi ha la sensació (¡i la realitat!) profunda i generalitzada d'ignorància i/o d'incompetència a què m'he referit.

Vista així la societat, no és fàcil de saber quin paper hi ha de tenir el «diccionari normatiu», i per tant, quina relació hi ha d'haver entre lexicografia i normativa. Dit d'una altra manera: de tot aquest món bigarrat de dubtes, què és el que pot assumir un diccionari general. Aquesta qüestió és la que ens ocuparà en el § 2. Precisem ara el concepte de normativa.

Almenys des dels anys quaranta s'ha donat, dins el nostre domini lingüístic, al concepte de *normativa* uns valors molt amplis i ambigus, fins i tot sovint contradictoris: s'hi ha inclòs, de manera vaga i genèrica, tota mena de materials de consulta escolar *si eren ortodoxos*, això és, si reduïen cada cop més la llengua a les categories de bo i dolent de manera fulminant. Aquests materials ben sovint restringien, allargaven, interpretaven, matisaven o silenciaven la doctrina pròpiament normativa, probablement perquè aquesta s'havia anat quedant insuficient i, alhora, era en part inassequible. Però ningú no s'adonava de tals maniobres. Certs altres materials que a poc a poc anaven apareixent al mercat eren curosament relegats per la «saviesa popular» a la categoria d'heterodoxos; però, també a poc a poc, han anat penetrant en aquest teixit cada cop més complex i avui la confusió és considerable. La confusió sobre el concepte de normativa.

L'única manera de no caure en aquest parany d'interpretació vaga del concepte és reduir-lo al màxim. Opino que el concepte restrictiu no pot ser altre que el següent:

la normativa gramatical i lèxica del català està constituïda *exclusivament* per la gramàtica de Fabra de 1918 (però en la seva versió final, de 1933), el diccionari de Fabra de 1932 i les aportacions o modificacions que l'Institut d'Estudis Catalans ha fet a les dues obres anteriors.

## 2. DICCIONARI GENERAL I DICCIONARI NORMATIU: POSSIBILITATS I LÍMITS

Acabem de veure les necessitats reals de la nostra societat. La intuïció ens diu que moltes d'aquestes necessitats no es poden atendre en un diccionari. Però resulta que el concepte de diccionari no és un concepte unívoc. Hi ha lexicògrafs que diuen que un diccionari ha de contenir tot allò que els usuaris hi poden buscar (i de la manera que als usuaris els sigui més còmoda). Per això immediatament es diu que hi ha diccionaris «de llengua» i diccionaris «enciclopèdics», etc. I per això hi ha llibres que es titulen «diccionaris» i que ens informen de les capitals dels estats, dels sistemes de pesos i mesures, de les sigles «més importants», dels «principals» barbarismes de la llengua, dels signes de puntuació, etc.: matèria que hem vist que figura o pot figurar als manuals o altres mitjans de consulta dels nostres serveis lingüístics.

I de fet, amb el refinament de les tècniques lexicogràfiques actuals i per mica que eixamplem el diccionari general en el sentit enciclopèdic, força dels aspectes problemàtics esmentats més amunt hi poden ser inclosos. Però també n'hi ha uns altres que no hi cabrien sense desvirtuar el concepte que avui tenim de les obres linxicogràfiques. Més en concret: admetent ja l'eixamplament enciclopèdic del diccionari i limitant-nos a la matèria que avui preocupa els usuaris, sembla que hi ha uns aspectes que s'hi poden recollir directament a través dels articles més pròpiament lexicogràfics (§ 2.1*a*), uns altres que requeririen articles temàtics específics (§ 2.1*b*) i uns altres que clarament són aliens a una obra lexicogràfica i que en tot cas haurien d'anar en apèndixs (§ 2.2). Finalment, una part d'aquesta matèria seria (més) pròpiament adequada per a un diccionari de tipus normatiu (§ 2.3).

### 2.1. Matèria que podria assumir un diccionari general

a) *Dins els articles lexicogràfics ordinaris*

*Qüestions lèxiques i fraseològiques.* Aquesta és la matèria més primària d'una obra lexicogràfica.

*Qüestions morfològiques.* Els dubtes morfològics també han estat tradicionalment resolts dins les entrades lèxiques normals: forma de certs singulars (si és *complex* o *complexe*), forma de certs lexemes (si és *poguer, vúlguer* o *poder, voler*) i flexions

nominals i verbals (un cas ben concret i avui interessant: els femenins o masculins de certes paraules que indiquen professions i oficis). Algun altre residu («*es* miren» vs. «*se* sap») tampoc no presenta dificultats per ser acollit.

*Qüestions semàntiques*. També és possibilitat primària d'un diccionari d'indicar relacions semàntiques de sinonímia, antonímia, etc.

*Qüestions ortogràfiques i altres de relacionades*. A part l'accentuació, la dièresi, el guionet i aspectes relacionats amb la composició de paraules (escriptura d'una erra o dues, etc.), un diccionari fàcilment pot informar de les regles d'apostrofació (perquè afecten molt pocs elements), de la partició de paraules i d'altres aspectes no pròpiament ortogràfics com ara el desxiframent d'abreviatures (però la reducció de paraules i conceptes a abreviatures ja no és matèria d'un diccionari general).

*Qüestions sintàctiques*. Dins articles lèxics concrets un diccionari pot donar moltes menes d'informació sintàctica o sintàctico-semàntica: valor i funcionament de les partícules, règims i altres relacions prepositives de verbs, noms, adjectius i adverbis, usos de l'article determinat, de l'article personal, de cada pronom (per exemple, dels relatius), etc. I pot informar del fet que els noms siguin «comptables» o «no comptables» (v. § 3.2*b*).

*Qüestions de combinatòria* (vegeu aquest concepte més avall, § 3.3). Un diccionari general pot també, sense gaires dificultats, donar compte de les relacions d'aquest tipus.

b) *Dins articles temàtics especials*

Hi ha una altra sèrie d'aspectes que un diccionari no pot abordar dins els articles lèxics, almenys d'una manera raonablement econòmica (sense excessives repeticions) i pedagògica (sense deslligar fenòmens relacionats), però que poden ser acollits clarament dins d'articles temàtics gramaticals. Aquests articles tenen l'inconvenient que no faciliten a l'usuari una solució ràpida dels seus dubtes, però cal acceptar que es tracta més pròpiament de seccions d'una gramàtica que no pas d'un diccionari. En teoria, doncs, certs aspectes cabrien dins articles lèxics normals, però seria amb els inconvenients esmentats. Aquí em refereixo a tots dos tipus conjuntament. La majoria de temes d'aquest apartat són sintàctics, però n'hi ha algun d'ortogràfic.

*Qüestions ortogràfiques*. Es podrien abordar en articles temàtics detalls com l'ús de majúscules i minúscules i la divisió de paraules en síl·labes (que així es podria exposar amb visió global). Altres qüestions ja potser són més llunyanes dels interessos lexicogràfics: la puntuació, l'ús de tipus de lletra (cursiva, negreta, versaleta).

*Qüestions sintàctiques*. Em limito a enumerar-ne unes quantes. Diversos aspectes de l'infinitiu (com ara la substantivació: com que no es manifesta únicament a través de l'article *el*, no pot anar dins l'article dedicat a la paraula *el*; valors que pot tenir; ús de la partícula *de* al seu davant: si va posposat i si va anteposat al verb principal; etc.), usos del gerundi, les maneres d'expressar els conceptes de «neutre» i d'«intensificació» (per tradició i comoditat, acollides a les gramàtiques en el capítol de l'article),

el canvi i la caiguda de preposicions (encara que el fenomen afecti pocs elements, té uns aspectes generals imprescindibles en una bona exposició), usos dels possessius i dels pronoms personals forts (se n'ha de dir el mateix que del tema anterior), la construcció «set *de* llargs i sis *de* gairebé invisibles» (que podria anar a l'article *de*, però és millor independitzar-la'n, perquè afecta també d'altres categories: els quantificadors, la paraula «(d')altre(s)», etc.), la col·locació dels adverbis, les possibilitats de col·locació dels pronoms febles en conjunts verbals («*la* va haver de deixar mullar», «va haver-*la* de deixar mullar», etc.), usos dels temps i dels modes verbals, ordre d'elements (qüestió extremament complexa que, a més a més, afecta diverses categories), relacions transformacionals diverses (per exemple, les que es produeixen amb verbs factitius i amb altres de pròxims: «*La Roser* llegeix (un conte)» / «Deixa-*la* llegir en pau» / «Deixa-*li* llegir el conte en pau»), certs aspectes de la coordinació (si es pot o no repetir elements comuns entre membres coordinats, si convé o no de marcar amb coma certes el·lipsis, etc.).

### 2.2. Matèria que no podria assumir un diccionari general

Tota matèria no continguda dins els articles especificats al § 2.1 s'ha de considerar aliena a una obra estrictament lexicogràfica. La qual cosa no vol dir que el comerç hagi de canviar d'hàbits en la producció de les obres anomenades «diccionaris»: són dues qüestions conceptualment diferents. Seria aliena a la lexicografia, i per tant no tindria més remei que anar en apèndixs (que podrien disfressar-se també d'articles o quadres temàtics) matèria com ara la següent.

*Qüestions de morfologia.* Presència o absència de l'article en topònims (majors o menors).

*Qüestions de sintaxi.* A primera vista, no sembla que hi hagi gaires qüestions de sintaxi que haguessin de quedar relegades a aquesta secció, però caldria mirar-ho amb més compte. Potser hi quedaria alguna cosa sobre la coordinació (qüestions de tipus pragmàtic o més pròpiament contextual), a part d'aspectes de teoria general (parts de l'oració, anàlisi sintàctica). Potser també hi haurien de quedar qüestions que semblen prou vagues com ara el fet que hi hagi verbs transitius que poden funcionar normalment sense complement (*beure, llegir, menjar, obrir, escriure, comprar, pagar*) i que n'hi hagi d'altres que no tenen aquesta possibilitat (*estripar, llençar, posar, portar*; *llançar* i *treure*, llevat en certs esports; *tocar* en general).

*Qüestions de combinatòria* (vegeu § 3.3). Com han fet veure diversos lexicòlegs, la combinatòria té uns límits, fins i tot si el concepte s'eixampla considerablement: cap diccionari no pot fer constar amb quins altres conceptes poden relacionar-se, per exemple, els conceptes físics com ara *paper*, perquè són il·limitats (un paper, podem comprar-lo, vendre'l, estripar-lo, cremar-lo, veure'l, pesar-lo, endur-nos-el, regalar-

lo; ens pot, agradar, etc.) i depenen de restriccions d'un tipus molt general (del fet de ser abstractes o no, de ser comptables o no, de ser animats o no, de ser humans o no, etc.; *paper* no pot, per exemple, plorar, penedir-se, invitar; no podem meditar-lo, invitar-lo; *fang*, a més a més, no admet un numeral: tres). Certament, si un diccionari pogués marcar adequadament els trets d'aquesta mena (primer caldria tenir-ne una llista totalment controlada i orientada als fins de la lexicografia pràctica), per exemple a la manera de Mel'čuk, aleshores aquesta dificultat es resoldria dins els articles estrictament lexicogràfics. Mentre no pugui fer-ho, unes hipotètiques orientacions en aquest terreny seran alienes a un «diccionari general».

*Qüestions tipogràfiques.* Qui volgués incloure en un diccionari general un tractat d'abreviatures, no tindria més remei que obrir un apèndix independent, sobretot si, a més a més de la descodificació d'aquests signes, volia donar també l'operació inversa, la codificació.

*Qüestions estilístiques.* Avui preocupa positivament entre els professionals de la llengua l'ús ampli (anomenat abús) que es fa de certs termes i de certes expressions. La veritat és que aquesta preocupació ha existit sempre en ambients del ram i fins ha provocat més d'una polèmica. Vegeu-ne uns quants exemples al § 3.1. Potser sí que un diccionari convencional podria tractar aquesta matèria dins els articles lexicogràfics corresponents a cada cas, però aleshores entraria en un terreny difícil de delimitar (difícilment un diccionari pot precisar amb tota claredat on s'acaba l'ús adequat del verb *existir*, de l'adverbi *majoritàriament* o del nom *tema*, i forçosament hauria de recórrer a explicitar els casos abusius, pràctica impròpia d'un diccionari convencional). Tampoc no veig manera d'encabir en un diccionari certes orientacions, de tipus estilístic, destinades a evitar la preponderància del gènere masculí en els escrits administratius o de qualsevol altra mena: un apèndix solucionaria el cas in extremis.

*Relacions amb altres llengües.* Com he indicat més amunt, aquí es concentren la immensa majoria de dubtes i vacil.lacions actuals de la nostra llengua, cosa ben natural. Hi ha sobretot els termes i les expressions castellans, tant de camps específics (administració, economia, esports, etc.) com de la llengua general; però hi ha també la interferència d'altres llengües, específicament de l'anglès, interferència que avui ja crea llistes considerablement llargues. Aquesta matèria és pròpia dels diccionaris bilingües, i per tant un diccionari general, si volia acollir-la, la relegaria forçosament a un apèndix. I seria un altre cas evident d'aquesta possibilitat extrema, d'aquest límit clar.

### 2.3. Límits d'un diccionari normatiu

L'exposició que he fet de la matèria que avui és objecte de preocupació pràctica sembla que ens permetria de diferenciar entre dos tipus de límits: els que hi hauria entre un diccionari i una gramàtica (*a*) i els que hi hauria entre un diccionari comer-

cial i un de normatiu o acadèmic (*b*). Però, a més a més, ens hem de preguntar quina és la matèria d'una llengua que s'ha de considerar normativa (*c*).

a) *Límit pràctic entre un diccionari i una gramàtica*

No veig que sigui fàcil d'establir límits teòrics entre aquestes dues menes de codis. Els límits serien, si de cas, pràctics: els constituirien les possibilitats tècniques d'un diccionari de donar compte de tal o tal fenomen. De fet, dins el nostre domini lingüístic aquest límit ha estat sempre tan borrós que es pot considerar inexistent: tota persona que ha tingut al davant un text per corregir ha tret la informació d'on ha pogut, i majoritàriament l'ha treta precisament del diccionari (que sempre o per a molts és més fàcil de consultar).

Per exemplificar aquests límits n'hi ha prou de fer un cop d'ull al que he dit al § 2.2.

b) *Límit entre un diccionari normatiu i un de comercial*

Des d'un punt de vista lexicogràfic, no veig que sigui fàcil de traçar límits entre un «diccionari general» i un «diccionari normatiu». Partint de la matèria que ens ha servit al llarg d'aquest treball, crec que es pot afirmar que no n'hi ha: en un diccionari hi ha d'haver (almenys idealment) tot allò que un usuari necessita consultar-hi, tant si es tracta d'un diccionari «general» com si es tracta d'un de «normatiu». Alguna acadèmia d'alguna llengua veïna a la catalana sembla que no ho entén totalment així, i sembla que el diccionari acadèmic o normatiu per a ells no tindria altra finalitat que establir quines paraules són admeses i quines no. Discrepo totalment d'aquest punt de vista extravagant. Tots els casos que donaré al § 3 afecten de ple el diccionari, normatiu o no.

El límit es podria veure (més) clar si parlem de diccionaris «comercials» i de diccionaris no comercials. Ja hem vist que els primers acullen, de la manera i en el lloc que poden, com més coses millor de les que se suposa que han de ser «útils» a l'usuari i que, per tant, faran vendre el producte. Als §§ 2.1 i 2.2 m'he referit a aquests aspectes, i hi he aclarit quina part de la matèria afectada és assumible i quina no per un diccionari que sigui només una obra lexicogràfica en el sentit tècnic del concepte.

c) *Quina matèria s'ha de considerar «normativa»*

Hi ha un altre concepte que en català (ni segurament en la majoria de llengües) no ha estat mai sotmès a clarificació ni a discussió: quina és o ha de ser exactament la matèria «normativa» d'una llengua. Concretament, ¿és o ha de ser matèria normativa la teoria i la pràctica de les abreviatures, la teoria i la pràctica dels signes de puntuació (inclosos els signes d'interrogació i d'admiració, els claudàtors, els guions, etc.), la teoria i la pràctica de les majúscules, la teoria i la pràctica dels tipus de lletra (cur-

siva, versaleta, negreta)? No entraré en aquesta qüestió, que exigiria més temps i espai dels que ara tinc, però és sens dubte una qüestió ben interessant.

Jo crec que en principi no es pot marcar una frontera en aquest terreny. El que deu haver passat en la història de les llengües cultes deu haver estat el següent: que hi ha uns aspectes de les llengües que s'han vist (almenys en un cert moment històric) com a més essencials o més urgents (establiment de la llista de lexemes i de paraules, establiment de la grafia d'aquestes paraules, establiment de la morfologia i de quatre fets de sintaxi), i n'hi ha uns altres que s'han anat descobrint més tard, de manera més marginal o com a fets (més) particulars dels tipògrafs (aquí hi hauria els altres detalls esmentats). Però a mesura que les llengües es tecnifiquen i entren en relacions mútues i internacionals, és més necessari o urgent de «normalitzar» també els altres detalls (en el sentit tècnic que avui té el concepte de normalitzar o estandarditzar), els quals, doncs, acaben confonent-se inextricablement amb els primers.

El català s'ha trobat que fins fa molt pocs anys només s'havia prestat atenció a aquelles primeres qüestions més essencials o clàssiques, però avui ja preocupen positivament moltes altres, com el lector ha vist. De manera que probablement encara passarem uns quants anys abans que el concepte de «normatiu» quedi prou ben clarificat.

## 3. EXEMPLES CONCRETS

Els exemples que segueixen no tenen cap pretensió d'exhaustivitat, ni de molt lluny: pretenen simplement d'il·lustrar amb una mica de detall el panorama a què m'he referit al llarg d'aquest article. Com que és matèria més aviat vulgar, divulgadíssima, esquivo la pedanteria de fer constar d'on els trec, llevat d'algun cas en què la dada pot ser especialment útil.

### 3.1. D'estilística

Els llibres d'estil i similars es preocupen cada vegada més de recomanar als lectors que no eixamplin excessivament els valors semàntics de paraules o expressions com les següents, o que no n'abusin en escrits més formals. Entre parèntesis dono alguna de les «solucions» recomanades.

> a nivell de (en, a, a tot, com a, en l'àmbit de ), accedir (anar, entrar, arribar, aconseguir, atènyer), amb anterioritat/posterioritat (abans, després, anteriorment), amb l'objectiu de (a fi de), contemplar (preveure, incloure, considerar), cosa (punt, assumpte, afer, detall), descartar (excloure, rebutjar), efectuar (programar, fer, treballar, actuar), emblemàtic (representatiu, important, destacat, propi), en base a (a partir de, sobre la base de, tenint en compte que, segons),

escollir (elegir, triar), existir (haver-hi, tenir), fonamental(ment) (principal, essencial, sobretot, més que res, bàsic, especial), incidir i incidència (influir, repercutir), impartir (fer, donar), majoritàriament (generalment, principalment, majorment), present (en textos com «el present document», etc.: aquest, l'esmentat), realitzar (fer, preparar, organitzar, intervenir), suposar (implicar, comportar), tema (dificultat, reunió, matèria...), una sèrie de (diversos, uns quants)

### 3.2. De sintaxi

a) *Construccions amb verbs*

Pel que fa als verbs, hi ha dos aspectes que avui no estan atesos de manera suficient a la nostra lexicografia normativa (ni a la no normativa) o que han entrat en crisi si el diccionari normatiu en teoria ja els tenia tipificats d'una manera o altra. En primer lloc, la consideració de transitius o intransitius; en segon lloc, les preposicions que un verb pot dur per introduir els seus complements estrictes (règims) o bé els complements més laxos (circumstancials). La nostra lexicografia s'acontenta d'indicar si un verb és «tr.» o «intr.», que és una informació molt poc útil en la pràctica (vegeu, per exemple, els dos usos ben diferents i tots dos «intr.» del verb *reeixir*, a sota, (ii)). En les llistes següents no em preocupo de mirar què diuen el diccionari normatiu o d'altres sobre cada cas. Els exemples adduïts són trets de la realitat, però prescindeixo en general d'indicar-ne els autors o el lloc d'aparició.

(i) Casos de problemes pel que fa a la categoria de transitius o intransitius:

acompanyar ('adjuntar' un document, etc.), atendre, aprofundir, augmentar («Els criats, *els havien augmentats* de vint pessetes», DGLC, *augmentar*), aprofundir, argumentar, complaure, contestar, cridar (en el sentit d''escridassar, renyar'), delegar («comissió formada pel rector o persona que aquest delegui / o persona en qui aquest delegui»), dictaminar, dimitir, encantar («*li/l'* encanten els moixernons»), interessar, matricular / matricular-se de, obeir, obviar, operar («el demostratiu és un determinant que opera la selecció de les entitats denotades pel conjunt»), pagar, pegar, plaure, pregar, recórrer, renyar, respondre, telefonar, trucar

*Presència o absència del verb «fer» com a principal*: avortar, botar, bullir, cessar, extraviar, onejar («Les mateixes fonts insisteixen que la sanció no respon al fet que el pres onegés la bandera catalana, sinó...»: dels diaris), ressaltar

(ii) Tipus de verbs. Aquí anoto casos que van més enllà del problema de la transitivitat (*deixar*) o de l'especificació del complement (*interferir*). Però en general són

casos de vacil·lació en el tipus de complement o de preposició. Dono entre parèntesis les preposicions o les funcions que es troben en la pràctica.

> abaixar(-se) / apujar(-se) / augmentar / avantatjar / baixar / créixer / disminuir / dominar / guanyar / minvar / perdre / pujar / rebaixar / ... (el preu ha baixat el 7,3% / un 7,3% / en un 7,3% / d'un 7,3% / (en) 300 ptes., l'equip X guanya o perd per dos gols / de dos gols / (per) dos a zero / per dos / de dos)
>
> *Anar*: ¿anar de compres / a comprar, de passeig / a passejar, de viatge / a viatjar?, ¿anar de vacances, d'excursió, de putes, de copes? (cf.: anar a dinar / *de dinada, d'excursió / a fer una excursió o excursions, de copes / a prendre una copa / ?a prendre copes, de viatge / ?a viatjar; es va endur el noi de vacances / ?a fer vacances, de copes / ?a prendre copes, de putes / ??)
>
> agafar (de/per: «caminaven agafats del/pel braç»), caracteritzar-se / distingir-se (per + infinitiu / pel fet de), contagiar (Ø / amb / de: «Aquesta actitud que tu ens has contagiat»), costar (quant costa, a quant costa, què costa una cosa), dedicar (dedicar X a/per Y: «no he dedicat gaires esforços extraordinaris per desenterrar cartes inèdites»), deixar (Ø / + infinitiu: «¿Continuaràs lluitant? —Si em deixen, sí»), dir («a aquest satèl·lit, li diuen Hispasat» / «d'aquest s., en diuen H.»), distingir-se (v. caracteritzar-se), erigir- se («Bush i Clinton voten amb la seguretat d'erigir-se *amb* la victòria»), escapar (de/a), esperar (Ø / a / fins que), estirar (Ø / de / per: estirar (d')un carro, estirar una persona de/per la mànigа), incidir (en/sobre), interessar-se (en/per), interferir(-se) («els factors extralingüístics que sempre han interferit en la nostra llengua»), lligar ('relacionar' a/amb), preferir (que / a + infinitiu: «La solidaritat àrab ha preferit tenir màrtirs que no pas solucionar problemes»), pudir (?a: «Aquella casa pudia a indigència»), reeixir («Tot li reïx», «Ell ha reeixit en la seva empresa»: DGLC i altres), relegar («el català havia quedat relegat de la vida pública»), reemplaçar («En aquest cas [...] *un* és reemplaçat *u*»: Fabra, gramàt. pòstuma, § 39, idèntic en totes les eds.), sobreviure (Ø/a), sol·licitar (a/d'una persona una cosa), substituir(-se) (substituir X per Y, substituir Y a X; la 2ª possibilitat és habitual en els escrits de Fabra: «es comprèn que la preposició que denota proximitat (*ad*) pugui fàcilment substituir-se a la que denota interioritat (*in*) [...]. D'aquí construccions com [...] *Anar a París*», conversa 236 de l'ed. Pey), substreure's (a/de), tallar (pa en/a llesques), titllar (de/com/com a: «En un apartat de les instruccions, el Minivolgrama és titllat com la perfecta síntesi de les qualitats buscades»), valer (quant val, a quant val, què val tal cosa)

(iii) Possibilitat de passivitzar-se. Un bon diccionari hauria de poder advertir si un verb admet o no la passiva, almenys en general (car aquest fenomen és gradual en certa mesura i en molts casos resulta impossible d'advertir-ho categòricament); en

dono tres o quatre exemples i em fixo més particularment en un altre detall: si són possibles certs participis passius o passats.

?«el català fou *supeditat* a les altres llengües romàniques»; «La seva obra ha estat bastant *parasitada*, ¿no?» (*El Temps*, 17.II.1992, p. 34; sc. 'hom n'ha extret matèria sense advertir- ho'); «Sens dubte aquest ciutadà havia estat ben *impactat* per la vostra tasca» (*El Temps*, 17.II.1992, p. 39)

*Caure, pujar*: diem «el noi ha pujat a l'arbre», però no «he vist el noi pujat a l'arbre» (sinó *enfilat*); diem «amb aquestes nevades ha caigut molts arbres», però no «he vist molts arbres caiguts» (en tot cas, *tombats, abatuts, a terra*)

Altres casos de participis: «Vint homes *adreçats* d'espases» (DGLC, *adreçar* 'proveir'); «una empresa *participada*» (sc. que una altra empresa hi té participació: d'*El barco fantasma*, p. 245; «Tabacalera Española, participada al 50% per la companyia suïssa Fabriques de Tabac Reunis», *El Temps*, 23.XI.92, p. 80; «L'empresa Tabasa està participada en gairebé un XX% per la Generalitat», Catalunya Ràdio, setembre 1992; «Aquesta última [societat] està participada en un 50 per cent», *Avui*, 31.X.1992, p. 36; cf.: «La empresa Materiales Hidráulicos Griffi, participada en un 86 % por Valenciana de Cementos», *La Vanguardia*, 11.VI.1992, p. 80); «Nom de la gran ploguda i inundació *haguda* en temps de Noè» (DGLC, *diluvi*); «*Deixat* portar per la fòbia política, la justesa de criteri de l'advocat i l'bjectivitat de l'historiador desapareixen» (Romà Planes, *El País*, Quadern, 16.I.1992, p. 6); «l'únic objectiu [...] és el de desprestigiar com sigui el *finat* president» (Id.); «va creure *arribada* la seva hora d'escalar a la més alta magistratura» (Id.); «El *dimitit* García Valverde...» (Catalunya Ràdio); «Els dos *fugats* són...» (*Avui*, 17.I.1992, p. 17); «L'entrenador *dimitit* no ha volgut fer declaracions»; «En la conferència de premsa *tinguda* al Palau de Pedralbes» (*El Temps*, 23.XI.92, p. 18) «un cop *florits* els jacints / *arribats* els vaixells / **agradats* els colors», «les persones *acabades* d'arribar / *d'agradar» (G. Rigau, *Caplletra* 8, p. 9 i 10). Cf.: «La tormenta *descargada* en Mallorca» (*La Vanguardia*, 8.X.1992)

De Joan Fuster, *Estudis d'història cultural*, Publicacions de la Universitat Jaume I, Castelló de la Plana 1992: «i en no pocs documents [...] trobem molt *insistida* una observació: l'escassa eficiència [...]» (p. 131), «Vull deixar-ho insistit, això» (135), «degué haver-hi una certa resistència a admetre-la [aquesta proposició], perquè la trobem insistida» (150), «la disputa no solament queda reflectida en [...], sinó que també ens arriba *adverada* per conductes laterals» (152); «En el lèxic poètic [...] apareixen uns quants mots insistits» (*Obres completes* 2, B. 1969, p. 191), «Els nostres pares inicials [...] aspiraven a ser com déus, aspiració absolutament *infatuada*» (ibid., p. 396)

b) *Construccions amb altres categories*

(i) Amb noms

contribució (fer contribució/ons sobre/en), dret (a/de), esment (fer — de/a), falta (fer — un jugador a/contra/sobre un altre), màster (organitzar o estudiar un — en/de tal matèria), mida («30 pàgines (de) mida foli»), motiu (de/per), opció (tenir — a/de/sobre/per a: «Les primes... costaran... 100 dòlars, 80 pagats per l'empresa i 20 pel treballador, que tindrà l'opció per a un servei més generós», d'un diari), por (de/a: avui completament habitual *a*, sense que desaparegui del tot *de*), raó (de/per), temps (tenir — de/per)

A part dels possibles complements, als noms s'hauria d'indicar si són comptables (o tendeixen a ser-ho) o no comptables. Cada dia són més freqüents els usos de certes paraules en plural: *orientacions, informacions, gents, calçats*, etc.

(ii) Amb altres categories. En altres llocs m'he referit a la vacil·lació constatable en l'ús d'altres peces lèxiques, com ara *altri* o *d'ençà*. En dono uns quants casos més:

correlatiu (a/de), distint/diferent (a/de), equivalent (a/de), pastat ('molt semblant': «és pastat al/del/amb el seu pare»), vell («les Bases de Manresa, velles de mig segle...»: *Revista de Catalunya* 61, 1992, p. 3)

malgrat: «malgrat tothom ho sap, ningú no ho recorda» (construcció estranya cada cop més freqüent); «malgrat (d')haver-ho vist...»

des de (des de... (fins) a; de... (fins) a)

hom («¿Com pot hom entendre l'existència de *capellans* casats i amb fills en algunes pel·lícules?», d'un diari; «les característiques principals del «Japanese-Style management», són les següents: 1. La col·locació vitalícia: és a dir, hom molt rarament canvia d'empresa. [...] 4. El sistema de sou per antiguitat: com més anys porta hom a l'empresa, més guanya», d'un llibre publicat a B. pel setembre de 1993; considero tots tres casos problemàtics)

### 3.3. De combinatòria

Hi ha un tipus d'informació que els lexicòlegs anomenen *collocation* i/o *combination* (en anglès i en francès), intensament debatuda durant els últims anys i, malgrat tot, encara no aclarida de manera unívoca o satisfactòria. Aquí l'anomenaré provisionalment *combinatòria*. Es tracta d'informació que es troba entre la sintaxi (estructures sistemàtiques previsibles), el lèxic (fraseologia) i la semàntica (compatibilitat semàntica entre lexemes). Uns quants exemples aclariran pràcticament de què es tracta. Al *Tirant* apareixen frases com les següents, que avui són de la manera que dono entre

parèntesis: *fer un somni* (*tenir* un s.), *fer dubte que* (*provocar* dubte/sospita que), *tancar els ulls* (*aclucar*/?*tancar* els u.). I a les criatures se'ls sent coses com ara *ficar un gol* (*fer* un g.), *fer una multa* (*posar* una m.). No es tracta exactament de sintaxi (*fer* i *tenir*, etc. admeten complements nominals), ni de semàntica (*aclucar* i *tancar* s'avenen exactament igual amb els ulls, etc.) ni de fraseologia (*tenir un somni* difícilment es podria qualificar d'unitat fixada). Així mateix diem *la verda paüra, la verda esperança* i no *la blava paüra/esperança*, no pas per cap relació (o incompatibilitat) semàntica entre aquests noms i aquests adjectius.

Les friccions que la combinatòria provoca en matèria de normativa potser no són gaire nombroses, perquè es tracta d'una matèria fonamentalment intuïda pels parlants, més que no pas regulada pels codis. Però com que la intuïció dels catalanoparlants es troba totalment desballestada sota la intensíssima influència del castellà, també el conflicte afecta aquest terreny ara i adés. I com que un diccionari ha de ser per a tothom (inclosos els no natius), és evident que aquesta matèria hi és fonamental. Sense anar més lluny, diversos casos dels que dono a continuació es troben resolts en general de manera diferent pels catalans del nord i del sud del Pirineu (per exemple, «*fer* o *formar* part d'una organització»). Dono tot seguit una llista de casos en què, independentment del fet que hi hagi o no hi hagi conflicte real i habitual, un diccionari podria orientar el lector o hauria de fer-ho. Entre parèntesis dono les paraules entre les quals s'hauria de seleccionar la o les adequades en la combinació. En general la informació dels parèntesis és totalment neutra pel que fa a la meva opinió, però algun cop hi afegeixo comentaris personals (sovint reduïts a un interrogant o un asterisc). En general suprimeixo els autors dels exemples. Em serveixo de notes inèdites d'Alba Fraser.

acció (emprendre, ?prendre: «No sabem quina acció podrà prendre el jutge durant els pròxims dies», de la ràdio)

una aixeta / una ampolla / una capsa / una porta / el gas / el foc / els llums (tapar, apagar, tancar)

atenció (posar, parar, fer, tenir, haver-hi, prestar, donar, atribuir, dedicar...)

autoritat (?fer)

baixa (?causar — una persona d'una empresa; el Termcat ha admès *causar estat* 'esdevenir definitiva [una resolució]')

bo (donar — de fer una cosa; v. *gust*)

calor/fred (fer-hi, haver-hi, agafar, ?entrar-li a un) [v. *olor*]

citació (fer un jutge una — a una persona: «citació que fa un jutge a una persona perquè comparegui davant seu el dia i a l'hora fixats», possible definició de l'expressió *citació a dia cert*)

càrrega (bèstia de; ?animal de; *brut de [sc. «brut» és algun cop sinònim d'«animal»)

emprar/usar/fer servir (?) (un criteri, un verb, sis pàgines, una forma mètrica, una citació, una determinada perspectiva en l'estudi d'una qüestió, un mètode, les seves influències per obtenir quelcom, un procediment: «sorprèn que Denk *empri* després sis planes per reimprimir l'inici de la traducció versificada del *Llibre dels set savis* [...]. Dedica força espai a la Gaia Sciència [...], exemplificant les diverses formes mètriques i poètiques *usades*», «[A qui escrivís aquesta història] li hauríem de donar plena llibertat [...] en l'*ús* de citacions, traduccions i il·lustracions», de la *Revista de Catalunya* 69, 1992, p. 125 i 129; «...materials de treball que ha calgut ordenar. Hem fet *servir*, per fer-ho [sc. ordenar-los], bona part del procediment que va *emprar* el mateix Sebastià Farnés», sdins Sebastià Farnés, *Paremiologia cat. comparada*, B. 1992, p. 13b)

difusió (fer, ?donar: «els impresos que arribin a l'entitat X per tal de donar-ne la difusió adequada...», d'un document del Consorci per a la Normalització Lingüística de Catalunya)

entrant/pròxim/vinent («la setmana / *el mes / *l'any / *el juliol / *l'estiu / *la nit / *el dia *entrant*»; «l'any / el mes / ?la setmana / *el segle / *el matí / *la nit / ?el juliol / ?l'estiu / ?la Pasqua *vinent*»; «el/la *pròxim(a)* estiu / setmana / juliol / ?tarda / ?any / ?segle / ?mes»)

estat/ordre (en perfecte —)

estiu/primavera/tardor/hivern (al pic de; cf. francès «au coeur de», castellà «en pleno/a»)

excel·lència (?alta: «Tindrem uns vins del 93 d'una alta excel·lència», *Avui*, 21.IX.1993, p. 13)

expedient (sofrir, posar, ?fer...)

fracàs (obtenir, tenir, sofrir, experimentar...)

gust (?donar — de fer una cosa; v. *bo*)

hivern (v. *estiu*)

importància (treure, llevar)

injecció (posar, introduir, clavar, injectar...)

meló/tomàquet/pa (una tallada, un tall, una llesca, una berla, un tros...)

membre (viril, ?masculí)

mercès (?«mercès a X» 'gràcies a X')

objectiu (atènyer, assolir, obtenir...)

olor / calor / fred / graus (de temperatura) / temperatura / pudor / soroll / silenci / remor / pols / fang / aigua / aiguat / tolls / núvols / ... (haver-hi, fer-hi, agafar, tenir, agafar-li a un, ?entrar-li a un)

pa (v. *meló*)

part (fer, formar: «fer part de», a *Diari de Barcelona*, 11.X.1991, p. IV)

pau (restablir, restaurar, ?refer) [v. *salut*]

pena (valdre la, ?pagar la; v. francès «*valoir* o *être* la peine o le coup», castellà «*valer, merecer* la pena»)

plats (fregar, ??rentar, ?netejar [sc. a mà]; *fregar, ?rentar, ?netejar [sc. a màquina])

premsa (roda, conferència de)

primavera (v. *estiu*)

pròxim (v. *entrant*)

salut (?restablir, *restaurar, refer) [v. *pau*]

tanda (demanar, donar, guardar (la))

tarannà (?«Totes aquestes persones m'admiren pel *tarannà heroic* amb què ho fan», sentit per ràdio)

tardor (v. *estiu*)

taula (parar, desparar, *posar, ?preparar)

temperatura/graus/... (haver-hi, fer(-hi), tenir) [v. *olor*]

tenda «de campanya» (parar, plantar, ?posar, ?muntar; desparar...)

tomàquet (v. *meló*)

tret (?fer, disparar, clavar...)

vicissituds (passar, haver-hi, tenir moltes...)

vinent (v. *entrant*)

volar/anar (quelcom — (molt) alt)

JOAN SOLÀ
*Universitat de Barcelona*

MERCÈ LORENTE / CARLES TEBÉ

# LEXICOGRAFIA I CORRECCIÓ: LA FUNCIÓ DELS DICCIONARIS EN LA REVISIÓ DE TEXTOS

## 1. INTRODUCCIÓ

Els diccionaris han estat considerats, des de sempre, eines de consulta bàsiques en l'aprenentatge d'una llengua, en la traducció i en la correcció de textos.[1] En aquest article volem introduir-nos en la xarxa de relacions establertes entre la lexicografia i la revisió de textos, i encetar una visió particular dels diccionaris i de les seves funcions.

Òbviament, la perspectiva d'un especialista en revisió de textos davant dels fenòmens lingüístics és ben diferent de la d'un aprenent de llengua, d'un ensenyant o d'un traductor. Intentarem establir des d'aquesta perspectiva la connexió entre les tasques del corrector i la consulta de diccionaris.

En primer lloc cal considerar que la consulta de diccionaris per part d'un corrector està mediatitzada pels aspectes propis de la revisió textual:

(1) *El text revisat té autor*

L'autor del text ha pres determinades decisions sobre la comunicabilitat del text que, malgrat que no hi siguin explícites, han de ser respectades en el procés de revisió. Cal que els canvis que imposa un corrector es limitin als aspectes més objectius i documentables (la correcció ha de ser una activitat controlada) i en tot cas justificables.

(2) *La revisió s'ocupa de detectar, en el text, elements contradictoris respecte d'una sèrie de normes lingüístiques*

(1) La denominació de la professió presenta una certa diversitat: correcció, revisió, supervisió, etc. Alguns professionals prefereixen les formes *revisió* i *supervisió* perquè fan referència al conjunt d'intervencions sobre el text, que superen la simple correcció lingüística i que es regeixen per criteris sistemàtics. En aquest article utilitzem indistintament les formes *revisió* i *correcció de textos* com a sinònimes.

*Caplletra 17* (Tardor 1994), pp. 291-310

Hem de tenir en compte que la norma lingüística amb què treballa un corrector no és única ni uniforme. Al marge de la norma oficial de la llengua estàndard, comuna a tot el col·lectiu lingüístic, hi ha un conjunt de convencions discursives, pragmàtiques, comunicatives i estilístiques que esdevenen normes lingüístiques tan fonamentals com la primera. Un text ha d'adequar-se a una finalitat comunicativa, ha de presentar continguts formalitzats de manera uniforme i s'ha d'integrar en una determinada variant lingüística: el corrector necessita tenir codificat un corpus de llengua extens i divers.

(3) *El procés de correcció ha de mantenir inalterables els continguts conceptuals, el grau de formalització del discurs i la funció comunicativa del text que pretenia l'autor*

Conseqüentment amb els punts (1) i (2), el corrector de textos ha de limitar-se a canviar aquells elements divergents amb el conjunt de normes que regulen l'adequació del text a una llengua particular i a un tipus de discurs determinat. Les modificacions, per tant, han d'abastar exclusivament aspectes lingüístics i paralingüístics, i no han d'incorporar tries estilístiques particulars ni el coneixement del món del corrector: les intervencions del revisor han de ser tan neutres com sigui possible.[2]

(4) *La revisió lingüística d'un text ha de garantir-ne la coherència dins de la variació lingüística*

La variació lingüística s'articula en dos eixos fonamentals: el dialectal, que integra les varietats temporals, geogràfiques i socials dels parlants, i l'eix funcional, associat a les varietats contextuals dels textos o registres (Payrató, 1988: 42-46). Tota producció lingüística és el resultat de la combinació de trets de l'eix dialectal i de trets de l'eix funcional.

Encara que el corrector ha de poder enfrontar-se amb tota mena de textos, de fet l'activitat professional li ofereix normalment un conjunt de textos limitat des de l'eix funcional, per bé que més o menys obert des del punt de vista dialectal: el corrector treballa fonamentalment amb textos de to formal.

I és en relació amb aquest relatiu grau de formalització on la revisió de textos connecta amb la varietat estàndard de la llengua, variant heterogènia que es caracteritza, al seu torn, per una certa variació i que s'identifica amb l'objecte de la codificació.

> *En la mesura que és la modalitat més neutra, menys marcada des dels punts de vista social, geogràfic i temporal, resulta força apropiada en qualsevol context. D'aquí neix l'ambivalència de l'estàndard com a varietat i registre. D'una banda és arrenglerable amb la resta de varietats dialectals (i es manifesta en registres diferents: oral o escrit, espontani o preparat...), però d'una altra apareix associada només amb determinats escenaris comunicatius (preferentment els formals).* (Payrató, 1988: 59).

(2) En alguns casos pot resultar lícit –i fins i tot convenient– que el corrector proposi esmenes encara que no les pugui documentar (en la revisió d'aspectes d'adequació, de coherència i cohesió del text, i en la correcció d'estil de textos literaris). En qualsevol cas, l'autor sempre les haurà de validar o rebutjar (vegeu la descripció de tipus de textos i correcció a l'apartat 3).

L'estandardització del català codifica variants fonètiques, morfològiques i lèxiques en dos nivells diatòpics: un d'àmbit general i un altre d'àmbit restrictiu per a algun dialecte. Per tant, la correcció d'un text català produït dins d'una varietat dialectal determinada ha de vetllar per mantenir coherentment els elements textuals en un d'aquests dos nivells, si fos el cas.

De manera paral·lela, la revisió ha de garantir la coherència del text en relació amb el canal, el camp temàtic, el tenor funcional i el grau de formalitat.[3]

(5) *Els elements susceptibles de revisió poden ser de tipus lingüístic i paralingüístic*

Els elements lingüístics que poden ser revisats es corresponen amb els nivells de representació fonètica, morfològica, sintàctica i semàntica.[4] Els elements paralingüístics revisables, en sentit ampli, comprenen el conjunt de convencions tipogràfiques, com l'ortografia, la puntuació o l'emfasització de textos; els aspectes d'adequació, de coherència i de cohesió; i la revisió d'estil.

D'acord amb totes aquestes premisses, ens podem plantejar les següents qüestions:

- Què busca un corrector en un diccionari?
- Quina mena de diccionaris li fan servei?
- Com ha de navegar pel diccionari a la recerca de la solució dels seus problemes?

Per la complexitat de la tasca de correcció de textos i per la diversitat de diccionaris que hi ha a l'abast d'un corrector, la consulta lexicogràfica pot presentar més entrebancs que no podríem imaginar.

En aquest treball introduirem una selecció de diccionaris apta per a la tasca de correcció de textos escrits d'entre la tipologia de diccionaris existents (2). A continuació, ens referirem als tipus de textos que són l'objecte de treball dels correctors professionals (3). Després observarem si, des de la perspectiva del diccionari com a inventari del lèxic d'una llengua, les classes de diccionaris seleccionades serveixen per revisar tota mena de textos (4). Des de la perspectiva del diccionari com a recull d'informacions relacionades amb les unitats lèxiques d'una llengua, intentarem esbrinar si els articles de diccionaris recullen tots els elements lingüístics i paralingüístics susceptibles de ser corregits (5). Finalment, i a tall de conclusió, farem una valoració del paper dels diccionaris com a eines per a la revisió de textos (6).

## 2. TIPOLOGIA DE DICCIONARIS

La història recent de la lexicografia està plena d'intents de classificació de diccionaris,[5] que han utilitzat criteris ben diversos. En aquest article partim de la tipologia establerta a Cabré i Lorente (1991), que resulta de l'observació de dades de la ma-

(3) Seguim, com Payrató (1988), la proposta de Halliday, McIntosh i Stevens (1964), reformulada per Gregory i Carroll (1978), que permet classificar els registres a partir de quatre factors: camp, mode, tenor funcional i to interpersonal.

(4) Qualsevol d'aquests aspectes poden referir-se a estructures lèxiques, oracionals o a la totalitat del text.

(5) Les classificacions de diccionaris s'han regit per diversos criteris lingüístics i extralingüístics. A tall d'exemple, podem referir-nos a la classificació històrica de B. Quemada (1968), a la genètica d'A. Rey (1970-1977) o a la tripartida de G. Haensch (1982).

croestructura i de la microestructura dels diccionaris existents i potencials. Aquesta classificació, que reproduïm en el quadre final, distingeix entre característiques generals de tota mena de diccionaris (A) i decisions selectives, en la macroestructura (B) o en la microestructura (C), que determinen diccionaris específics.

Aquí, però, per la caracterització que hem fet de la feina de revisió de textos, en presentem una versió reduïda, més acostada a les necessitats de l'usuari. El corrector consulta els diccionaris de manera diferent a la del traductor o de l'aprenent, a la recerca d'una alternativa adequada per a un element textual divergent amb la normativa o amb el tipus de discurs. Generalment el corpus lingüístic que necessita tenir codificat difereix dels interessos d'altres consultors de diccionaris, perquè treballa amb textos més variats per la temàtica, la funció i l'estructura.

Així, el corrector professional necessita corpus extensos, amb molta informació temàtica per als textos científico-tècnics, i per això hi incorporem les enciclopèdies.[6] Pel que fa al nombre de llengües, només hi recollim els diccionaris monolingües i bilingües, per si de cas li cal comprovar l'adequació de manlleus. Deixem de banda els diccionaris específics massa restrictius (d'autor i obra, de sociolectes,[7] o d'un sol dialecte geogràfic).

Només tenim en compte els diccionaris amb forma d'entrada lexemàtica o sintagmàtica (hi refusem els visuals, els de locucions, etc.). I, com que centrarem la nostra atenció en la revisió de textos escrits, no hi recollim els diccionaris fonètics.

De l'aplicació dels anteriors criteris, en resulta la següent tipologia reduïda:

1. Enciclopèdies i diccionaris enciclopèdics
2. Diccionaris d'ampli abast: etimològics i dialectals
3. Diccionaris generals monolingües
4. Diccionaris bilingües
5. Terminologies
6. Diccionaris associatius: de sinònims i ideològics
7. Diccionaris correctius: de barbarismes i ortogràfics.

## 3. TIPOLOGIA DE TEXTOS I CORRECCIÓ

En el procés de revisió d'un text escrit qualsevol,[8] el corrector, en identificar un element contradictori amb qualsevol de les normes que interpreta que afecten aquell text, segueix habitualment una sèrie de passes de manera seqüencial:

1- decidir què busca (*identificar el problema*)
2- decidir on ho busca (*identificar l'eina que li cal*)
3- un cop identificada l'eina, decidir com es localitza dins el diccionari (*manejar l'eina*)

(6) Les enciclopèdies no apareixien en la tipologia de Cabré i Lorente (1991), per la seva orientació onomasiològica i l'abast supralèxic de la seva nomenclatura.

(7) Els diccionaris específics per la selecció del lèxic d'un sociolecte són, en el cas de català, escassos i d'abast molt restrictiu i limitats temporalment: *Des del vapor de la O. Diccionari de veus populars i marineres, Vocabulari de l'argot de la delinqüència.*

(8) Quan parlem de *revisió* partim del supòsit que un autor s'encarrega de la redacció del text original, i que, un cop produït, una persona diferent duu a terme la revisió. Altrament, el procés tindria a veure amb les fases i estratègies de producció del text per part d'un mateix autor (vegeu Cassany, 1987).

4- si la recerca no ha donat cap fruit, decidir a quines altres obres que no siguin diccionaris es pot recórrer (*buscar altres eines*).

Abans d'abordar el primer punt –quins problemes poden plantejar-se a un corrector de textos escrits i quines eines pot fer servir per resoldre'ls– elaborarem una tipologia simplificada de textos a fi de classificar les contradiccions amb la norma que hi puguin aparèixer, ja que els diferents tipus de text plantegen problemes també diferents.

En l'establiment d'aquesta tipologia de textos per a la revisió, tindrem en compte dos criteris de classificació: *a*) el tenor funcional, i *b*) les llengües que intervenen en la producció del text.[9]

Segons el primer criteri del tenor funcional, dividirem els textos en dos grans grups:

- *textos referencials*: qualsevol text on predomini bàsicament la funció referencial, ja sigui de tipus descriptiu, narratiu, directiu o argumentatiu[10]
- *textos literaris*: qualsevol text produït amb una finalitat bàsicament literària o creativa (obres de ficció)

I, pel que fa al segon criteri de la llengua/les llengües de producció del text, distingirem entre:

- *textos creats en la mateixa llengua en què es revisen*
- *textos traduïts, produïts originalment en una altra llengua.*

El creuament d'aquests dos criteris de classificació ens permet d'obtenir quatre tipus textuals que, a grans trets, responen bastant fidelment a la varietat de textos escrits que habitualment són objecte de revisió per part de correctors professionals.

Intentarem a continuació de respondre a la primera pregunta que formulàvem més amunt (què es busca?), que ens permetrà de classificar els tipus de problemes més usuals en la revisió de textos. La segona pregunta (on es busca?) la contestarem en l'apartat següent (4); la tercera (com es busca dins el diccionari?) és tractada a (5) i la darrera pregunta (en quins casos la recerca és infructuosa) es respon a les conclusions.

En la revisió d'un text se'ns poden plantejar diversos dubtes, que varien considerablement en funció del tipus de text, de la seva finalitat, de l'emissor i el destinatari, del canal utilitzat, etc.[11]

1.1 Si ens enfrontem primerament amb un *text de tipus referencial, que s'ha de revisar en la mateixa llengua que s'ha produït*, els problemes més usuals amb què es pot trobar un corrector poden ser:

i) *referits a la interpretació general de la norma lèxica*:

- la paraula és desconeguda

(9) Simplifiquem molt la tipologia per raons pràctiques. Els criteris de classificació del text com a unitat de comunicació s'han basat en criteris diversos i no sempre coincidents, a partir de la formulació clàssica de Jakobson (1962) en la seva descripció de les funcions de llenguatge: emotiva, conativa, referencial, metalingüística, poètica i fàtica, i és una qüestió que resta encara oberta. Recordem que en aquest treball només considerem el codi escrit, i partim de la base que la gran majoria dels textos escrits que s'han de revisar són de caràcter formal. Per a una acurada descripció de les tipologies textuals, vegeu Fernández-Villanueva (1991).

(10) Dins els textos referencials encara podríem distingir, a propòsit del grau d'especialització del text, entre textos de caràcter 'general' i textos de tipus 'tècnic'. Per a alguns autors, molts textos 'generals' són susceptibles de ser tractats com a textos tècnics, en virtut de les circumstàncies específiques (pragmàtiques i comunicatives) que regeixen la producció de qualsevol text. Així, un article de la secció d'esports d'un diari, un text didàctic sobre història natural o un document administratiu o comptable poden rebre la consideració de 'textos tècnics', per bé que el grau d'especialització temàtica que palesin sigui relativament baix. Per a més aclariments sobre el concepte de discurs tècnic o especialitzat i la seva frontera amb el llenguatge comú, vegeu Cabré (1992).

(11) En aquest apartat contemplem tots els tipus d'esmenes possibles que duen a terme els correctors *'lingüístics'*–és a dir, aquells correctors que en la revisió d'un text intervenen en qualsevol aspecte relacionat genèricament amb la llengua–. La revisió professional de textos sovint se centra en un aspecte prèviament determinat (gramatical, estilístic, tipogràfic) però el corrector professional sol detectar aspectes no previstos per l'autor o per la ins-

- la paraula presenta algun dubte pel que fa a la seva forma ortogràfica (o flexiva)
- l'abast semàntic de la paraula presenta algun problema d'interpretació, perquè
    - és una extensió (o restricció, o modificació) de sentit no recollida als diccionaris
    - sembla utilitzada en un àmbit d'ús que no li és propi
    - sembla referir-se als trets semàntics propis d'una altra paraula
- hi ha dues formes (que semblen) concurrents,
    - i es vol comprovar si són realment equivalents
    - i es vol comprovar si totes dues són correctes
    - i se'n vol prioritzar una per davant de l'altra
- la paraula és formalment correcta però sembla un calc
- la paraula no està adaptada al sistema de la llengua (o en contradiu alguna regla de formació)
- la paraula és identificada com a problemàtica perquè:
    - els diccionaris de referència en proposen solucions diferents
    - no és acceptada per un diccionari de referència (però per d'altres sí)
    - les obres de referència no recullen la paraula perquè és un terme massa especialitzat
- altres.

ii) *referits a la interpretació general de la norma gramatical*

- *ortogràfics*: problemes o dubtes d'apostrofació (esses líquides, vocals aspirades) o de l'ús del guionet, referits a noms propis o comuns.
- *morfològics*: problemes de flexió nominal i verbal, de l'ús dels determinants i dels quantificadors
- *sintàctics*: problemes referits a vacil·lacions en l'ús dels locatius (a/en), en el règim verbal, en el canvi i caiguda d'algunes preposicions (anar amb cotxe / anar en cotxe), en l'ús del gerundi, pronoms febles, concordances, etc.

iii) *referits a l'adequació, la coherència i la cohesió*

- *adequació*: ¿la varietat funcional i el registre utilitzats són els més adequats? Hi ha algun tipus d'incoherències en el grau de formalitat o en el tenor del text?

tància que encarrega la revisió, i que afecten altres nivells de la producció del text.

- *estructura i coherència*: ¿la informació presentada en el text està seleccionada i organitzada en una estructura comunicativa prou adequada? L'ordre lògic i temporal del text és el més escaient?
- *cohesió*: ¿hi ha els elements de cohesió exigibles al llarg del text: referències endofòriques, exofòriques, enllaços, connectors, concordances, etc.?

iv) *referits a aspectes paralingüístics*

- *puntuació*: el text, està ben puntuat? Hi ha recursos no utilitzats? N'hi ha d'altres utilitzats abusivament?
- *aspectes tipogràfics i altres convencions*: hi ha una correcta representació de majúscules i minúscules, de les abreviacions i sigles?
- *disposició:* la presentació formal del text, és l'adequada? Hi ha tots els elements necessaris? Hi ha una adequada jerarquització de la informació (capítols, subcapítols)?

v) *referits a l'estil en general*

- ¿el text és prou *precís* i presenta un grau suficient de *varietat* lèxica o discursiva en general al llarg del text? ¿Hi ha mots considerats superflus o inútils, redundàncies, aspectes culturalment poc escaients, es poden evitar frases o expressions ambigües?

1.2 Si el *text*, a més del seu caràcter *referencial*, s'ha *produït en una llengua diferent de la qual es revisa*, el corrector, a més de la majoria d'aspectes esmentats, haurà de controlar específicament qüestions:

i) *referides a la fidelitat del text traduït respecte de l'original*, en tots els seus aspectes, i concretament:

- al correcte transvasament de les unitats d'un text a l'altre: el mot, la fraseologia, l'oració
- a l'adequació de les varietats i registres del text de partida i el text d'arribada

ii) *referides a les mancances que presenta el text en la llengua d'arribada respecte del text original*, singularment:

- paraules que apareixen en la forma de la llengua original, sense traduir o adaptar
- deficiències o errors en la solució donada a fenòmens idiosincràtics de la llengua de partida (xenismes, noms propis significatius, enquadrament sociolingüístic, etc.)[12]

(12) Coseriu (1977) parla de dos tipus de problemes en la traducció: *a*) els aspectes relatius a la 'transposició', és a dir, els problemes d'establiment de correspondències entre les diferents unitats de traducció: i *b*) les qüestions relatives a la 'traducció' pròpiament dita, que poden implicar «adopció, adaptació, explicació analítica de significats de la llengua de partida i comentari o aclariment». Els tipus *a* i *b* de Coseriu es corresponen bastant amb els problemes exposats als tipus *i)* i *ii)* d'aquest apartat.

2.1 Quan es troba davant un *text* de caràcter *literari, si està escrit en la mateixa llengua en la qual es revisa*, el corrector pot revisar teòricament aspectes pertanyents a qualsevol dels àmbits descrits a 1.1, és a dir:

i) *referits a la interpretació de la norma lèxica o gramatical*
ii) *referits a l'adequació, a la coherència i a la cohesió*
iii) *referits a aspectes paralingüístics*
iv) *referits a qüestions d'estil.*

Però la diferència amb la correcció del text referencial és que en aquest cas les contradiccions amb les diferents normes exposades més amunt poden ser fruit de la voluntat deliberada de l'autor –és a dir, violacions intencionades de la norma–. La revisió, per tant, esdevé més complexa, en el sentit que la lectura crítica del corrector ha d'estar guiada per una informació fiable sobre el registre funcional i la varietat dialectal on se situa voluntàriament l'autor, i a més li caldrà conèixer les convencions pròpies del gènere en què escriu (que poden afectar, entre altres, els aspectes paralingüístics). El procés de revisió esdevé per tant menys objectiu, menys documentable, més delicat. Per aquesta raó, les correccions de textos literaris (correntment anomenades 'correccions d'estil') se sotmeten, per regla general, a l'aprovació de l'autor del text.

2.2 Si el *text*, a més del seu caràcter *literari, està escrit en una llengua diferent de la qual es revisa*, el corrector haurà de controlar, a més, els aspectes descrits a 1.2, és a dir, en termes de Coseriu (1977):

i) aspectes relatius a la 'transposició' del text
ii) aspectes relatius a la 'traducció' del text.

3. En conclusió, la revisió d'un text de tipus referencial presenta una tipologia de possibles correccions que es poden classificar bastant bé en els grups establerts genèricament a 1.1: nivells lèxic i gramatical, adequació, estructura, coherència i cohesió, aspectes paralingüístics i deficiències d'estil. En termes generals, una revisió global d'un text considerat molt deficient implicaria corregir aspectes relatius a tots els apartats descrits.

Per contra, la revisió d'una obra literària en la pròpia llengua sol afectar un nombre de correccions molt més petit que la revisió d'un text de tipus referencial, ja que la competència de l'autor en el codi sol ser –per la pròpia naturalesa creativa del text– bastant elevada. Pel que fa al tipus de correccions, val a dir que poden afectar qualsevol àmbit dels esmentats a 2.1, però solen centrar-se més en aspectes d'inadequació a la varietat i al registre triats per l'autor (incloent-hi els problemes inclosos en el concepte d'estil) i en aspectes menors de tipus paralingüístic. Com que en un text literari

les diferents contradiccions amb la norma poden ser una opció deliberada de l'autor, aquest sol tenir la potestat d'acceptar o de rebutjar en darrer terme les esmenes que li proposa el corrector.

En les dues classes de textos caracteritzades per la procedència d'un codi lingüístic diferent del de la llengua de revisió, aquesta esdevé molt més complexa i els tipus de correccions possibles es multipliquen: poden afectar qualsevol dels aspectes esmentats a 1.2 i 2.2, a més dels descrits a 1.1 i 2.1, segons la qualitat de la traducció, les dificultats intrínseques que presenta el text i els coneixements del corrector de les dues llengües en joc.

## 4. LA CORRECCIÓ DE TEXTOS I ELS DICCIONARIS (I): LA MACROESTRUCTURA

En aquest apartat reprendrem els tipus de diccionaris seleccionats per a la correcció que havíem establert a (2) i comprovarem en quina mesura la nomenclatura de cada diccionari permet de resoldre els problemes de correcció que plantejàvem en l'apartat anterior.

### Tipus de diccionaris[13]

1. *Diccionaris enciclopèdics*: Les enciclopèdies i els diccionaris enciclopèdics solen contenir informacions tant de caràcter lingüístic com extralingüístic; dit d'una altra manera, a més de dades sobre la llengua, aquests diccionaris presenten informacions sobre les coses, o sobre el món.

En la correcció de textos, els diccionaris enciclopèdics serveixen: *a*) com a obres temàtiques de referència, ja que, cobreixen parcialment una gran quantitat de lèxic comú especialitzat temàticament que difícilment trobaríem en qualsevol diccionari de llengua, per complet que fos; *b*) com a diccionaris de llengua ampliats, ja que contenen gran quantitat d'informació sobre el sentit dels mots que pot ser necessària per desfer ambigüitats semàntiques; *c*) com a guies de referència per a la correcció d'aspectes paralingüístics referits als noms propis.

Globalment, els diccionaris enciclopèdics són útils per corregir determinats textos de caràcter referencial, on complementen l'ús del diccionari general.

2. *Diccionaris d'ampli abast: etimològics i dialectals*:

2.1 Els *diccionaris etimològics* contenen informació relativa a l'origen i –a vegades– a la història de les paraules (evolucions semàntiques, canvis formals, etc.) i per tant funcionen en la pràctica com a diccionaris de llengua complementaris des d'un vessant diacrònic.

(13) Per a la caracterització de l'abast dels diccionaris hem consultat Haensch (1982) i Landau (1984).

En la correcció de textos, els diccionaris etimològics serveixen: *a*) per localitzar usos antics o obsolets que han estat bandejats del diccionari de llengua general; *b*) per verificar usos literaris o poètics que tampoc són al diccionari general però que es basen en creacions del fons històric de la llengua.

Globalment, els diccionaris etimològics són útils per corregir textos antics, textos redactats en el present però que fan referència a temps passats, o bé textos poètics o literaris en general que continguin lèxic d'ús molt poc freqüent.

2.2 Els *diccionaris dialectals* presenten informacions relatives als usos lèxics i gramaticals propis d'una o de diferents varietats geogràfiques d'una llengua.

En la correcció de textos, els diccionaris dialectals serveixen per identificar el lèxic marcat geogràficament que no sol aparèixer en un diccionari general.

En general, els diccionaris dialectals s'usen en la revisió de textos referencials que tracten qualsevol aspecte relatiu a un àmbit geogràfic concret del domini d'una llengua, o en textos literaris que reprodueixen una variant dialectal del codi.

3. *Diccionaris generals monolingües*: Són els diccionaris que contenen informacions relatives a la 'llengua general', és a dir la variant estàndard d'una llengua. L'abast i les informacions que contenen aquests diccionaris varien molt segons els casos, però solen incloure àmplia informació semàntica, gramatical i d'ús (exemples, àmbit temàtic, registre sociolingüístic, etc.).

En la correcció de textos, els diccionaris generals serveixen: *a*) per consultar qualsevol problema d'interpretació de la norma lèxica; *b*) per consultar problemes referits a la norma gramatical (ja siguin ortogràfics, morfològics o sintàctics); *c*) per consultar aspectes relatius a l'adequació (en el cas que continguin informacions sobre registres i marques d'ús).

En general, els diccionaris generals són útils com a obres de referència bàsica en la revisió de qualsevol tipus de text, tant si són de caràcter referencial com literari.

4. *Diccionaris bilingües*: Els diccionaris bilingües són obres de referència que contenen equivalències entre les unitats lèxiques de dues llengües, a partir d'una llengua de partida (que pot tenir el seu corresponent invers).

En la correcció de textos, els diccionaris bilingües serveixen: *a*) per verificar la correcta equivalència entre les unitats de traducció de la llengua de partida i la llengua d'arribada; *b*) per trobar un recanvi per a un equivalent dubtós o problemàtic en la llengua d'arribada.

En general, els diccionaris bilingües són obres de referència obligada per a les revisions de textos traduïts, tant si són de tipus referencial com literari.[14]

5. *Terminologies*: De caràcter normalment multilingüe, les terminologies contenen subconjunts del lèxic d'àmbits temàticament especialitzats en sentit ampli, ja siguin

(14) En el cas del català, els diccionaris bilingües castellà-català han acomplert habitualment una funció correctora que no els era pròpia, per raons de tipus sociolingüístic. Aquí no abordarem aquest aspecte de la seva consulta.

àrees del coneixement o de l'activitat humanes (ciència, tecnologia, humanitats), sectors d'activitat professional o lúdica (comerç, esports, oficis) o qualsevol altre àmbit susceptible de tenir un lèxic propi.

En la correcció de textos, les terminologies serveixen: *a*) per a la identificació de lèxic especialitzat en la llengua pròpia, en qualsevol dels àmbits esmentats; *b*) per a l'establiment o la verificació de correspondències de termes entre llengües, en virtut del seu caràcter bilingüe o plurilingüe.

En termes generals, les terminologies són útils per corregir textos especialitzats temàticament, tant si són redactats en la llengua original com si són fruit d'una traducció.

6. *Diccionaris associatius*: Els diccionaris dits associatius (que comprenen generalment els diccionaris de sinònims i els ideològics) contenen informacions que relacionen els continguts semàntics de les paraules, i poden presentar una ordenació alfabètica (diccionaris de sinònims) o sistemàtica per conceptes (ideològics).

En la correcció de textos, els diccionaris associatius serveixen: *a*) per verificar el grau d'equivalència semàntica de les unitats del text: *b*) per buscar equivalents semàntics a les unitats del text.

En termes generals, els diccionaris associatius són útils en les revisions d'aspectes de coherència i cohesió lèxica dels textos referencials, i en les correccions d'estil (aspectes de varietat/precisió lèxica) que se solen practicar en textos de caràcter literari.

7. *Diccionaris correctius*: Per diccionaris correctius entenem fonamentalment els diccionaris ortogràfics, que contenen informacions sistemàtiques sobre la recta grafia dels mots i altres informacions gramaticals (però no de sentit: no solen incloure definicions).

En la revisió, els diccionaris correctius serveixen: *a*) per verificar aspectes dubtosos o problemàtics referits a la grafia dels mots; *b*) per verificar-ne altres aspectes gramaticals (plurals irregulars, forma femenina de noms i adjectius, etc.); *c*) per consultar-ne aspectes paralingüístics (particions síl·làbiques, pronúncia figurada, etc.).

En general, els diccionaris ortogràfics i correctius en general són eines útils per a consultes puntuals relacionades amb la grafia dels mots i amb problemes de tipus flexiu.

## 5. LA CORRECCIÓ DE TEXTOS I ELS DICCIONARIS (II): LA MICROESTRUCTURA

En l'apartat anterior hem vist com el corrector de textos tria els diccionaris que li són útils segons la mena de textos que ha de revisar. Ara ens fixarem en el conjunt d'informacions lingüístiques i paralingüístiques que formen part de la miscroestructura[15]

(15) La microestructura del diccionari és el patró d'informacions sistemàtiques i ordenades que apareixen lligades a un lema. De l'aplicació de la microestructura a cada entrada en resulten els articles del diccionari.

dels diccionaris de (2), per intentar respondre les següents preguntes:

- Quins problemes pot resoldre el corrector en la consulta de diccionaris i quins no?
- Quines mancances presenten els articles dels diccionaris consultats per a la revisió de textos?
- I d'altra banda, al corrector, li són útils totes les informacions de l'article de diccionari?
- Quina seria la microestructura ideal d'un diccionari general per a la revisió de textos?

El prototipus d'article lexicogràfic que resulta de la intersecció d'informacions dels articles de diccionaris seleccionats s'estructura com segueix:

*entrada* categoria gramatical
informació morfològica
etiquetes sociolingüístiques
etiquetes d'àrea temàtica o associativa
definició[16]
equivalents a d'altres llengües
remissions: sinònims, variants dialectals, alternatives de barbarismes
etimologia

1. Ortografia

El lema o entrada, a més de funcionar com a paraula clau en la recerca, permet la verificació d'aspectes ortogràfics. La resta d'elements paralingüístics revisables (puntuació, emfasització convencional, format de presentació) no són revisables mitjançant diccionaris, sinó mitjançant altra mena d'obres (tractats de puntuació, llibres d'estil).

2. Sintaxi

La categoria gramatical hauria de donar pistes sobre el comportament morfosintàctic d'una peça lèxica, cosa que no sempre aconsegueix ser exitosa. Les deficiències en aquest punt s'han de cercar en els codis de representació poc explícits i en els continguts poc aclaridors de la resta de l'article. Per exemple, un verb codificat com a intransitiu (v. intr.) pot portar en català complements obligatoris no referenciats en l'article del diccionari consultat. Només la intuïció del corrector o la consulta d'altra mena d'obres podrà aclarir-li qüestions com aquestes.

La informació gramatical del lèxic es presenta en el diccionari de manera dispersa, a l'etiqueta categorial, la definició i els exemples i citacions. Però aquesta dispersió, més que completar la informació, presenta mancances evidents. L'etiqueta de categoria

(16) Les definicions d'enciclopèdies, terminologies i diccionaris generals es diferencien pels continguts, l'organització de les dades i l'orientació.

gramatical, de transmissió tradicional, és insuficient i massa general. Les definicions d'alguns diccionaris generals, si són d'elaboració formal, aporten algunes dades,[17] però no són suficients. Els exemples i les citacions poden il·lustrar el comportament sintàctic d'una peça lèxica, però no cobreixen ni tots els contextos ni tota la variació sintàctica possibles.

Aquesta dispersió i asistematització alenteixen la consulta dels correctors professionals, que preferirien disposar de diccionaris de règim exhaustius.[18] Aquests diccionaris no són capaços de donar compte de la sintaxi oracional o d'aspectes de cohesió textual, però sí que haurien de reflectir les restriccions gramaticals (sintàctico-semàntiques) del lèxic.

3. Morfologia

La informació morfològica no disposa d'un espai fix en la microstructura dels diccionaris.[19] La majoria només recull de manera sistemàtica la flexió de gènere dels adjectius, representada per una doble entrada. La morfologia verbal només apareix reflectida en alguns diccionaris generals quan es fa referència al model de conjugació d'alguns verbs, mitjançant un verb patró, o quan s'introdueix la forma de participi si és irregular.

La presència de la informació morfològica en l'article a més d'irregular també és dispersa: de vegades s'afegeix entre parèntesis abans o després de la categoria gramatical, de vegades la metallengua de la definició ho fa explícit, i d'altres es reflecteix en l'entrada.

La morfologia lèxica també presenta mancances perquè les nomenclatures de diccionaris no recullen tots les formes derivades reals. En el cas de formes derivades possibles pels seus formants però no recollides en un diccionari general, la consulta de diccionaris de formants pot resoldre, de manera ràpida, la incorporació d'un neologisme en un text.

La informació morfològica sistemàtica només es reflectiria en diccionaris lexemàtics, en què cada article s'encapçalaria per un lexema i incorporaria totes les formes flexives i derivades relacionades.[20]

4. Indicacions de context

Les etiquetes de caire sociolingüístic es relacionen amb els aspectes tractats a l'apartat (4), perquè restringeixen l'abast de l'article o de part de l'article a un registre lingüístic determinat. De tota manera, les etiquetes usuals que fan referència a aquestes dades presenten uns límits molts difusos: Quina diferència hi ha entre qualificacions com *familiar* i *col·loquial*? Les unitats etiquetades com a *vulgars* poden considerar-se integrants de la llengua estàndard? Però no ens estendrem en aquest punt, perquè d'altres autors ja se n'han ocupat.[21]

(17) En el cas del DGLC (1932), la formalització de les definicions permet identificar el complement directe, en el cas dels verbs transitius, i copsar-ne la selecció semàntica. Exemple: **estuprar** *v. tr.* Violar (una verge).

(18) En català no disposem, ara per ara, de cap diccionari de règim (Cabré i Lorente, 1991).

(19) L'únic diccionari català que presenta informació morfològica de manera sistemàtica és el *Diccionari català invers amb informació morfològica* (1990) de J. Mascaró i J. Rafel. Però, per la seva ordenació les unitats lèxiques es relacionen per la flexió o per la sufixació, mai pel lexema.

(20) En català no disposem tampoc de diccionaris lexemàtics ni de diccionaris metòdics en què les unitats lèxiques es relacionin per la forma dins de l'article.

(21) Vegeu l'article de Ll. Payrató en aquest número de *Caplletra*.

La utilitat que tenen les marques de registre per al corrector de textos és molt limitada, perquè l'acotació general de canal escrit i to formal redueix força l'aparició d'unitats lèxiques marcades amb etiquetes com *familiar, col·loquial* o *vulgarisme*.

Els articles encapçalats amb marques temporals (*arcaisme, en desús, antigament,* etc.) serien interessants per al corrector, si li fornissin alhora equivalents sincrònics d'aquelles unitats.

La referència a àrees temàtiques o àmbits associatius es vehicula mitjançant etiquetes en diccionaris generals o per la restricció de la nomenclatura en terminologies o diccionaris ideològics. Fonamentalment, la seva funció és relacionar una unitat lèxica amb el discurs especialitzat.

Per tant, en la revisió de textos les etiquetes temàtiques serveixen per verificar la pertinença d'una unitat lèxica en un àrea determinada o per identificar-ne el referent semàntic en relació amb d'altres unitats contextualitzades. Per exemple, l'adscripció d'una unitat determinada a una especialitat i a un descriptor nocional permetrà al corrector combinar sinònims en la redacció, de manera més exacta.

De tota manera, si el text s'inscriu en una especialitat ben delimitada és preferible per al corrector utilitzar diccionaris terminològics. En aquest sentit, encara queda molta feina per fer, per tal que el panorama de les especialitats científiques, tècniques i professionals estigui cobert amb diccionaris específics. Aquesta mancança queda resolta provisionalment per les informacions etiquetades dels diccionaris generals i per les informacions extres de les enciclopèdies.

## 5. La definició

La definició de diccionari reflecteix informacions de nivells diversos. Hem vist que en alguns casos proporciona informació sintàctica formalitzada, però la informació fonamental que forneix la definició es relaciona amb el component semàntic.

La definició relaciona una denominació amb un referent, pot reflectir la jerarquia del camp semàntic en què s'inscriu aquest referent i delimita l'abast significatiu del referent en relació amb d'altres del mateix camp semàntic.

Però, quins problemes de la revisió de textos pot resoldre la consulta de definicions? El corrector utilitza la definició de diccionaris per solucionar aspectes relacionats amb la cohesió i la precisió lèxica. Per exemple, si l'ús d'un mot en un context determinat no és l'adequat, la consulta de la definició permet verificar-ne la inadequació; però, en canvi, no resol el recanvi lèxic.

Com observem, l'usuari-corrector no concedeix a la definició la primacia en l'article del diccionari que la tradició li ha atorgat. En aquest sentit, participa de la idea d'alguns autors, com Hausmann (1988), que consideren la definició prou feble com per esdevenir inútil.[22]

De tota manera, i en relació amb els aspectes semàntics, la definició pot incloure sinònims, perífrasis o hiperònims que permeten corregir aspectes de variació lèxica en un text.

(22) Hausmann (1988) argumenta, en contra de la definició, la circularitat, la asistematicitat, la ilimitació i la incomprensió en alguns casos. I és partidari de la seva substitució per fraseologia contextuada.

6. Remissions

Els equivalents en d'altres llengües no són gairebé mai pertinents per a la revisió de textos, si no és per encetar una nova via d'informació sobre aspectes semàntics o enciclopèdics d'una denominació concreta o per verificar equivalències en el cas de textos traduïts.

Les remissions monolingües varien en cadascun dels tipus de diccionaris. La remissió més usual, la sinonímia, permet incrementar la variació lèxica. En canvi, el corrector ha de ser molt prudent amb les remissions de variants dialectals, fornides per diccionaris d'ampli abast. Recordem que els textos revisats, generalment de to formal, s'inscriuen en la variant estàndard i els dialectalismes inclosos no han de ser massa connotats.

Els diccionaris de barbarismes són un cas molt peculiar.[23] Només seran utilitzats en la revisió de textos de forma auxiliar, quan el corrector detecti en el text un barbarisme i no en recordi possibles alternatives.[24]

Finalment, la etimologia és una informació de la lingüística històrica que no té utilitat en la revisió de textos, si no és en un cas molt complicat en què l'etimologia faciliti la delimitació semàntica d'una unitat lèxica.

7. La microestructura d'un diccionari general per a la revisió de textos

Tal com hem vist, el corrector professional necessita diccionaris específics caracteritzats per un tipus d'informació exhaustiva, que li permetin resoldre ràpidament problemes lingüístics i ortogràfics del lèxic de textos escrits formals.

Els diccionaris ortografics i els de règim li proporcionaran informació en un primer nivell de revisió. Els diccionaris terminològics i les enciclopèdies li resoldran, en un segon nivell, problemes relacionats amb l'especificitat temàtica del text. Finalment, els diccionaris de formants li oferiran informació pertinent per acceptar la inclusió d'unitats lèxiques no codificades ni intercanviables (neologismes).

Però, al marge d'aquests diccionaris especialitzats, la lexicografia actual sembla que tendeix a dissenyar diccionaris generals per a grups d'usuaris específics: diccionaris escolars, per a nivells educatius distints, diccionaris per a l'aprenentatge d'una segona llengua, diccionaris per a traductors, etc. Seguint la mateixa línia, un diccionari general que contemplés les necessitats d'un corrector de textos, atenent a la caracterització que hem fet de la seva feina, tindria la següent microestructura:

| | |
|---|---|
| *entrada* | informació sintàctica |
| | restriccions semàntiques |
| | contextos |
| | sinònims |
| | altres remissions lexemàtiques i semàntiques |

(23) Encara que presenten una estructura similar a la dels diccionaris (ordenació i equivalents), a M. Lorente (1991), no es consideren obres lexicogràfiques, sinó inventaris auxiliars, pel seu contingut i la seva orientació.

(24) L'aprenent d'una llengua, natiu o estranger, o el parlant en situacions de conflicte lingüístic no s'adonen generalment de la producció de barbarismes, fet que manlleva als diccionaris de barbarismes tota utilitat de consulta.

La microestructura resultant representaria un diccionari-tipus òptim per al corrector, equivalent a un diccionari general amb informació exhaustiva sobre subcategorització, selecció semàntica dels arguments, remissions sistemàtiques de tota mena i informació contextual.

## 6. CONCLUSIONS

Podem resumir aquesta aproximació a la funció dels diccionaris en la revisió de textos concloent que no totes les qüestions que es planteja un corrector professional troben resposta en els diccionaris que habitualment consulta, i que hem descrit a (2).

En efecte, *hi ha aspectes de la revisió de textos que es poden corregir satisfactòriament amb l'ajuda de diccionaris*: són els aspectes referits principalment a la interpretació de la norma lèxica, i secundàriament a la interpretació de la norma gramatical. Les obres bàsiques de referència en aquests casos són els diccionaris generals (i bilingües generals en el cas de traduccions), que necessàriament s'han de complementar amb altres diccionaris en funció del tipus de text que es revisa (vegeu-ne la tipologia a (4)).

En canvi, *hi ha aspectes de la revisió de textos que no es poden corregir satisfactòriament amb diccionaris però sí que es poden corregir amb l'ajuda d'altres obres lingüístiques de referència*: són els aspectes referits a la interpretació de la norma gramatical tractats insuficientment pels diccionaris –que es resolen a les gramàtiques–; i els aspectes de tipus paralingüístic, que es poden resoldre amb l'ajuda d'obres més específiques, com ara tractats de puntuació, diccionaris de sigles o abreviatures o llibres d'estil.

Per contra, *hi ha aspectes de la revisió de textos que no es poden corregir amb l'ajuda de diccionaris ni altres obres lingüístiques de referència*: són els aspectes que afecten fonamentalment problemes d'adequació, de coherència i de cohesió, i les qüestions considerades pròpiament d'estil. En aquests casos és el corrector qui ha de resoldre les incoherències o contradiccions amb la norma sobre la base de la seva pròpia competència lingüística; els diccionaris poden actuar només com a obres de consulta informativa amb un valor referencial molt més secundari.

Hem vist també que alguns dels aspectes revisables per diccionaris no troben una resposta adequada per mancances estructurals i informatives dels diccionaris existents. La producció de diccionaris específics des del punt de vista gramatical (diccionaris de règim o de formants) complementaria el panorama lexicogràfic amb eines molt vàlides per a la correcció de textos.

Un diccionari general de llengua per a un usuari-corrector continuaria la via lexicogràfica d'elaborar diccionaris específics per a usuaris, iniciada amb els diccionaris d'aprenents de llengua. En el disseny del prototipus d'aquest diccionari, que nosaltres apuntem succintament, és del tot necessària la intervenció dels correctors professionals.

Finalment, cal tenir en compte que l'abast 'normatiu' de les obres de referència que són els diccionaris sempre pot veure's restringit per un conjunt de normes més específiques determinades per la instància que encarrega la correcció, i que poden tenir una vigència igual –o prioritària– en la revisió d'un text particular.[25]

MERCÈ LORENTE
CARLES TEBÉ
*Universitat de Barcelona*

(25) Aquesta és la funció dels anomenats *llibres d'estil* d'alguns organismes, editorials o mitjans de comunicació, que solen contenir referències explícites a la majoria dels aspectes esmentats a (3), ja siguin d'ordre lingüístic o paralingüístic, o pròpiament d'estil: normes de presentació de textos, normes de puntuació, llistes de paraules (manlleus acceptats, manlleus rebutjats amb la forma substitutiva proposada, grafies dubtoses, barbarismes), índexs de noms propis, normes de presentació o disposició del text... i moltes altres qüestions. Tots aquests aspectes poden complementar –o matisar, o contradir– les informacions presentades als diccionaris de referència.

## REFERÈNCIES BIBLIOGRÀFIQUES

CABRÉ, M. T. (1992) *La terminologia. La teoria, els mètodes, les aplicacions.* Barcelona, Empúries.

CABRÉ, M. T.; LORENTE, M. (1991) *Els diccionaris catalans: de 1940 a 1988.* Barcelona, Universitat de Barcelona.

CASSANY, D. (1988) *Descriure escriure. Com s'aprèn a escriure.* Barcelona, Empúries.

CENTRE D'ÉTUDES DU LEXIQUE (1990) *La définition.* París, Larousse.

COSERIU, E. (1977) «Lo erróneo y lo acertado en la teoría de la traducción». A: *El hombre y su lenguaje. Estudios de teoría y metodología lingüísticas.* Madrid, Gredos.

FERNÁNDEZ-VILLANUEVA, M. (1991) «Tipologies textuals». A: *COM ensenyar català als adults*, suplement núm. 8, octubre 1991.

HAENSCH, G. et al. (1982) *La lexicografía.* Madrid, Gredos.

HAUSMANN, F. J. (1988) «La définition est-elle utile? Regard sur les dictionnaires allemands, anglais et français.» A: *CENTRE D'ÉTUDES DU LEXIQUE* (1990), pàgs. 225-233.

JAKOBSON, R. (1963) *Essais de linguistique générale.* París, Éditions du Minuit.

LANDAU, S. I. (1984) *Dictionaries. The Art and Craft of Lexicography.* Cambridge, Cambridge University Press.

LORENTE, M. (1991) «Funció lexicogràfica i contacte de llengües: Els diccionaris de barbarismes». A: *Actes del IXè Col·loqui Internacional de llengua i literatura catalanes (Alacant-Elx).* Barcelona, Publicacions de l'Abadia de Montserrat (en premsa).

PAYRATÓ, Ll. (1988) *Català col·loquial.* València, Universitat de València.

A. CLASSIFICACIONS GENERALS

1. Classificació per any d'edició
2. Classificació per editorial
3. Classificació funcional
    3.1. Diccionaris prescriptius
        - Generals
        - Terminologies i lèxics d'especialitat
    3.2. Diccionaris correctius
        - Barbarismes
        - Ortogràfics
        - (Fonètics)
4. Classificació pel nombre de llengües
5. Classificació per les llengües que contenen

B. SELECCIONS PER LA MACROESTRUCTURA

1. Per destinataris
    1.1. Escolars
2. Per la selecció de les entrades
    2.1. Per un període històric
    2.2. Per l'època actual: neologismes
    2.3. Per un dialecte geogràfic
    2.4. Per la categoria gramatical
    2.5. Per un sociolecte
    2.6. Per la temàtica
3. Per la forma de l'entrada
    3.1. Visuals
    3.2. De formes gramaticals
    3.3. Locucions
    3.4. Formants
4. Per l'ordenació
    4.1. Inversos
    4.2. De la rima
    4.3. Temàtica
5. D'autor i obra
6. De sinònims i antònims
7. D'homònims i parònims
8. Etimològics
9. Dialectals generals
10. Ideològics

C. SELECCIONS PER LA MICROESTRUCTURA

1. Transcripció fonètica
2. Informació morfològica
3. Categoria gramatical
4. Definició
5. Exemples
6. Il·lustracions
7. Sinònims
8. Antònims
9. Ètims
10. Dades dialectals
11. Citacions
12. Informació enciclopèdica

VICTÒRIA ALSINA
LLUÍS SOL / JOAN SOLER

# UNA TIPOLOGIA DE LA DEFINICIÓ LEXICOGRÀFICA APLICADA AL *DGLC*

## I. PER A UNA TIPOLOGIA DE LES DEFINICIONS LEXICOGRÀFIQUES

Els diversos intents d'establir una tipologia de les definicions lexicogràfiques que coneixem han partit de pressupòsits, enfocaments i finalitats diferents que han condicionat en gran manera els resultats de la classificació i, sobretot, el fet que no es pugui parlar d'una classificació universalment acceptable.

Entre els criteris que estableixen les línies mestres de la tipologia de definicions lexicogràfiques, hi ha el que articula el llenguatge de descripció lexicogràfica a l'entorn de dos eixos de caràcter metalingüístic: la metallengua de signe i la metallengua de contingut, que, considerat des d'un punt de vista prou general, estableix diferències de base pel que fa al tractament de les diverses categories gramaticals (és a dir, només els mots lexemàtics són definibles en termes de metallengua de contingut, mentre que els gramaticals han d'ésser forçosament explicats gramaticalment).[1]

Si per *definició lexicogràfica* entenem l'explicitació (mitjançant una paràfrasi) del *significat conceptual* del terme definit, podrem restringir-ne fàcilment la tipologia en el sentit que aquesta exclourà d'entrada la consideració del tractament d'aquelles partícules funcionals de la llengua (sc. conjuncions, preposicions, etc.) que com que no tenen significat conceptual són *indefinibles*.

(1) La distinció entre dos tipus de metallengua de descripció lexicogràfica diferents ha generat una certa bibliografia i més d'una terminologia per als mateixos conceptes. Ja Weinreich (1962: 39) distingeix dos tipus de metallengua, que Rey-Debove primer (1967: 143) anomena simplement *primera* i *segona* metallengua i, més tard (1971), anomena metallengua *de contingut* i metallengua *de signe*, respectivament. També respectivament, Martin (1977) parla de *definició paràfràstica* i *definició metalingüística*. A grans trets, podem dir que la metallengua *de contingut* consisteix en una anàlisi del significat conceptual del definit (de què consta la definició més pròpiament dita), mentre que la metallengua *de signe* expressa discursivament diverses propietats funcionals del definit. Tenim meta-

*Caplletra 17* (Tardor 1994), pp. 311-324

La lexicografia tradicional sol analitzar aquestes partícules fent-ne una explicació del funcionament, i, doncs, encara que aquesta *explicació* s'ha pres també per *definició* lexicogràfica,[2] l'estratègia que hi aplica no té res a veure amb la metodologia de la definició de caràcter parafràstic que analitzarem.

Cal dir, encara, que tant la lingüística filosòfica d'aquest segle (Wittgenstein)[3] com la pràctica lexicogràfica moderna coincideixen en la manera de validar empíricament l'explicitació parafràstica del significat. Es tracta de la validació per la *prova de la substitució* (a què recorre ja Leibniz en la formulació del seu principi de la identitat dels indiscernibles: *Eadem sunt quorum unum potest substitui alteri salva veritate*), la qual consisteix a substituir el definit, usat en un context lingüístic determinat, per la seva definició i d'observar si hi ha manteniment o no de significat.

Una conseqüència derivable directament d'aquesta prova de la substitució és que la pràctica lexicogràfica, a partir d'un moment determinat, s'imposarà l'endocentrisme entre definit i definició com un ideal a aconseguir. L'existència d'aquest tipus d'endocentrisme ens imposa una primera classificació de les definicions lexicogràfiques: aquestes s'ajustaran (o intentaran ajustar-se) al comportament *categorial* del definit.

Tres categories gramaticals diferents (el substantiu, el verb i l'adjectiu) seran l'objecte d'estudi del present article i el punt de partença de tres casuístiques més o menys diferents i més o menys coincidents de definició lexicogràfica, a partir del recull lexicogràfic que presenta el *Diccionari general de la llengua catalana* de P. Fabra o *DGLC*.[4]

## II. DEFINICIONS SINTÈTIQUES I DEFINICIONS PERIFRÀSTIQUES

Algunes caracteritzacions tendents a l'establiment de tipus de definicions es basen en criteris de naturalesa compositiva, i es fixen en l'estructura interna de cada definició. S'estableixen, així, diferències entre les definicions *sintètiques* (que consten d'una o més paràfrasis d'un sol mot) i les *perifràstiques* (que consten d'una o més paràfrasis de més d'un mot), o entre les *simples* (que consten d'una sola paràfrasi del definit) i les *múltiples* (que acumulen més d'una paràfrasi del definit).[5]

A. Una definició *sintètica* estableix l'equivalència amb un únic mot. Aquest únic mot definidor s'ha d'entendre com a sinònim absolut del mot definit. I, en aquest cas, tant podem prendre el mot definidor o descriptor per un element de la metallengua de contingut com per un element de la llengua objecte (sc. la llengua descrita),[6] és a dir, una remissió a la definició que es troba sota l'entrada del terme definidor en el diccionari.[7] Vegem exemples de definicions sintètiques simples i múltiples del *DGLC* relatives a les tres categories que ens interessen:

**estalzim** *m.* Sutge.
**envitricoll** *m.* Embull, embrolla.

---

llengua de contingut a (*a*) i metallengua de signe a (*b*), com es pot veure en els següents exemples del DGLC:

(*a*) **contrafloc** *m.* Vela triangular que va a l'extrem del botaló.

**emperxar** *v. tr.* Penjar en una perxa.

**numuliforme** *adj.* De forma discoïdal.

(*b*) **perquè** *conj.* Introdueix l'oració subordinada en què es diu el motiu o bé el fi o objecte de l'acció expressada en la proposició principal. [...]

(2) Martin (1977), p.e., parla de *definició* metalingüística. Rey-Debove (1967), però, ja rebutja aquesta aplicació del terme *definició lexicogràfica* quan assenyala que: *Il n'y a de définition lexicographique que s'il y a analyse conceptuelle* (pàg. 143).

(3) Quan pretén validar (vegeu Wittgenstein, 1953, I: § 558) la seva postulació que el significat d'un mot és el seu ús en el llenguatge (cf. Wittgenstein, 1953, I: § 43).

(4) No entrarem, però, a analitzar directament en aquest article les raons que fan que en el *DGLC* trobem definicions que no s'adeqüen a la prova de la substitució.

(5) La distinció entre definicions *simples* i *múltiples* la deixem només apuntada, ja que no condiciona la metodologia de la definició conceptual, expressada en metallengua de contingut, centre d'interès del nostre article.

(6) Cf. el comentari de Zgusta (1971) sobre l'aparició de sinònims del definit en la definició: *it is one of the interesting features of lexicography that such an explaining synonym can be taken as belonging to both categories at the same time* (pàg. 259 n. 124).

(7) Aquesta mena de definicions –si és que se les pot denominar com a tals– han estat sovint considerades com a *remissions* d'un terme suposadament secundari a un altre de suposadament *principal*.

**afonar** *v. tr.* Enfonsar.
**escampar** *v. tr.* Espargir, dispersar, disseminar; [...]

**bissextil** *adj.* Bixest.
**envitricollat** *adj.* Intricat, embullat, embrollat.

En el cas dels verbs, una definició sintètica necessita un descriptor equivalent tant semànticament com funcionalment al definit, això vol dir que si el definit és transitiu, el descriptor caldrà que sigui també transitiu. En canvi, si el definit és transitivo-pronominal (v. *calar-se, empassar-se*) el descriptor podrà ésser transitiu o bé transitivo-pronominal; si el definit és intransitiu o intransitivo-pronominal (v., p.e., *grifar-se*), el descriptor podrà ser, tant en un cas com en l'altre, intransitiu o bé intransitivo-pronominal:

**bleir** *v. tr.* Cremar.
**calar** [...] | Pron. Engolir. [...]
**empassar-se** *v. pron.* [...] enviar-se, englotir.
**tombar** [...] || Intr. Girar.
**dansar** *v. intr.* Ballar.
**carpir-se** *v. pron.* Decandir-se, consumir-se.
**apitrar** [...] | Pron. Penedir-se.
**grifar-se** *v. pron.* Grillar-se || Frisar.
**adir-se** [...] *v. pron.* Harmonitzar.

B. Altrament, quan l'element principal de la definició (això és, el *descriptor*) va acompanyat d'elements que el modifiquen, de cara a ajustar la seva caracterització semàntica a la del definit, tenim una definició anomenada *perifràstica*. Per ex.:

**sutge** *m.* Substància negra dividida en finíssimes partícules que prové de la combustió incompleta de les matèries orgàniques i que es diposita en les xemeneies i tubs que condueixen el fum.
**contingència** *f.* Qualitat de contingent.

**degotar** *v. intr.* Caure gota a gota. [...]
**emmidonar** *v. tr.* Impregnar de midó desfet en aigua.

**bondadós -osa** *adj.* Ple de bondat.
**bel.lipotent** *adj.* Poderós en la guerra.

## III. DEFINICIONS SUBSTANCIALS I DEFINICIONS RELACIONALS

(8) Vegeu la nota 2.

Seguint la classificació de Rey-Debove (1967), les definicions lexicogràfiques[8]

es poden dividir en dos grans grups en funció de la naturalesa de llur metallenguatge. D'una banda, tenim les definicions *substancials*, que es basen sempre (per inclusió o per exclusió) en un mot de la mateixa categoria gramatical que el definit. D'altra banda, tenim les definicions *relacionals*, en què la identitat categorial entre el definit i la perífrasi definidora no es troba en el nivell lèxic sinó que es fa mitjançant un constituent sintàctic. La definició substancial és l'únic tipus de definició possible en el cas dels substantius i els verbs, mentre que és facultativa en el cas dels adjectius i els adverbis.

Les definicions *substancials* admeten qualsevol de les modalitats tractades en l'apartat anterior i, per tant, poden ésser definicions *sintètiques* o *perifràstiques*. Les *relacionals*, per raons òbvies, només poden pertànyer al segon tipus.

En el cas de l'adjectiu, les definicions *relacionals*, poden ésser de dos tipus:

(*a*) *De relatiu*, en què la definició consisteix en una oració de relatiu. Per ex.:

**bimà -ana** *adj*. Que té dues mans.
**blegadís -issa** *adj*. Que es pot blegar.
**bulliciós -osa** *adj*. En què es mou bullícia.

Aquest tipus de definició és exclusiva dels adjectius, i, com en les definicions substancials, s'hi detecten grups de fórmules definidores que es relacionen amb determinats sufixos. Per ex.:

**aliforme** *adj*. Que té forma d'ala.
**bacciforme** *adj*. Que té la forma d'una baia.

**ballable** *adj*. Que es pot ballar.
**forjable** *adj*. Que es pot forjar.

**barbut -uda** *adj*. Que té barba [...]
**forçut -uda** *adj*. Que té molta força muscular.
**traçut -uda** *adj*. Que té traça.

Com en les definicions substancials, s'observen petites variants en les fórmules definidores del *DGLC* que no sempre responen a canvis semàntics. També s'observen fórmules definidores alternants sinònimes («Que es pot» i «que pot ésser», per exemple).

(*b*) *Preposicional*, en què la definició és introduïda per una preposició. Per ex.:

**brevirostre** *adj*. De bec curt.
**brocat -ada** *adj*. Amb brocats d'or o argent.

Aquest tipus de definició, que és compartida pels adjectius amb els adverbis, és menys freqüent que els altres que presenta l'adjectiu.

## IV. LES DEFINICIONS PERIFRÀSTIQUES SUBSTANCIALS

El lexicògraf, a l'hora de definir un mot, pot aprofitar-se de les relacions lèxiques existents en la llengua i confegir definicions mitjançant sinònims o a partir d'un descriptor hiperònim o antònim del definit. D'aquí surten tres tipus bàsics de definicions *lèxiques*[9] que podem anomenar, com ha fet algun altre autor,[10] *sinonímica*, *hiperonímica*, i *antonímica*.

El lexicògraf pot, d'altra banda, desglossar el significat que vol descriure en una sèrie de components semàntics sense prendre per descriptor cap relació lèxica reconeguda com a tal.[11] Ens trobem, llavors, davant d'una definició que anomenarem *semocomponencial*. La definició semocomponencial recorre a una sèrie de descriptors força genèrics que en molts casos es poden considerar semosintàctics, perquè donen compte de la transcategorització que ha convertit una categoria en una altra o perquè, simplement, donen compte de les propietats funcionals de l'enunciat definidor i del definit.

Un subtipus de definició tant lèxica com semocomponencial és la definició *morfosemàntica*,[12] que apareix quan és possible establir correspondències directes entre els diversos components de la definició i els diversos morfemes del definit, quan aquest és morfològicament compost o derivat.

### A. Les definicions lèxiques: sinonímica, hiperonímica i antonímica

Com que el tipus de definició que aquí anomenem sinonímica, la que consta d'un sinònim del definit, no és altre que una definició sintètica, de què ja hem parlat en l'apartat 2, no en tornarem a parlar aquí. Començarem parlant, doncs, de la definició hiperonímica i continuarem amb la definició antonímica.

#### 1. *La definició hiperonímica*

Entre els diferents tipus de definició que han estat descrits per la lexicografia teòrica destaca el que podríem considerar com a més representatiu de l'activitat lexicogràfica, el qual basa la definició lexicogràfica en la relació d'hiperonímia que s'estableix entre l'element genèric (descriptor, *genus*, etc.) de la definició i el mot definit. El significat d'aquest element genèric es restringeix, fins arribar a la relació de sinonímia entre el definit i la paràfrasi definidora, amb la concurrència d'altres precisions (*differentia*). Així, doncs, s'assumeix de manera implícita que el definit i la paràfrasi definidora han de mantenir una relació d'*implicació bilateral*, cosa que comporta subsidiàriament la intercanviabilitat o substituïbilitat de l'un per l'altre en un context determinat, sense que ni el significat de la frase ni la seva sintaxi se'n ressenteixin.[13]

Com la majoria de les relacions semàntiques, la hiperonímia[14] es verifica sobretot en elements lèxics pertanyents a la classe nominal, per mitjà d'enunciats com 'A *és un tipus de* B' o 'A *és una mena de* B'. S'han reconegut també relacions d'hiperonímia

(9) Una definició *lèxica* és, doncs, una definició que té per descriptor una relació lèxica, això és, que el descriptor és un terme que forma part d'una relació lèxica (o de sentit) amb el definit. Per al concepte de *relació lèxica*, vegeu Cruse (1986).

(10) P.e., Bosque (1982).

(11) Això no vol dir que aquest tipus de definició no es pugui basar també en una relació lèxica, però aquesta relació lèxica no apareix ocupant el lloc ni fent la funció de descriptor. Com veurem més avall, relacions lèxiques com la metonímia o la paronímia poden generar aquest altre tipus de definició.

(12) Prenem el nom de Rey-Debove (1967 i 1971), bé que amb un ús més restringit.

(13) De fet, aquesta formulació tan taxativa requereix una minuciosa matisació que per raons d'espai no podem desenvolupar aquí. Cal fer notar, només, que cal entendre aquest principi de *substituïbilitat* en un sentit més aviat ampli, que s'ha incorporat com un dels desideràtums metodològics més generals de la pràctica lexicogràfica.

(14) Preferim aquest terme, introduït per Lyons (1968), a causa de la seva major precisió enfront d'*inclusió*, que prové de la tradició de la lògica. Sobre l'ambigüitat d'aquest segon terme vegeu Lyons (1968). També considerem preferible el terme *definició hiperonímica* enfront del més estès de *definició analítica*, que podria aplicar-se a definicions basades en altres relacions semàntiques.

en elements de la classe verbal,[15] verificables mitjançant l'enunciat 'A *és una manera de* B', de la qual cosa podem inferir l'existència de relacions més o menys paral·leles en el cas dels adjectius ('*la propietat* A *és una modalitat de la propietat* B').

Per ex.:

**bivac** *m.* Campament de nit al ras.
**pústula** *f.* Petit tumor cutani que conté pus.
**estimbar** [...]‖ Pron. Caure daltabaix d'una timba. [...]
**velejar** *v. intr.* Navegar a la vela.
**bai -a** *adj.* Roig fosc, aplicat esp. al color dels cavalls.
**bla blana** *adj.* [...] ‖ Suau al tacte.[16]

En el camp de la definició verbal, la definició hiperonímica no es pot dir que sigui la més freqüent, ja que aquest tipus de definició només és viable en aquells casos en què podem parlar d'estratificació i de camps semàntics entre els verbs. Un d'aquests camps seria el dels verbs de percepció (*gustar, sentir, oir, veure*), definits per mitjà del descriptor *percebre*; un altre camp semàntic seria el de les feines del camp relacionades amb la sega (*dallar, monjoiar, ressegar, verdassegar*), definits per mitjà de l'hiperònim *segar*, etc. Tanmateix, la majoria de verbs no es poden relacionar amb veritables hiperònims, casos en què cal recórrer a una definició semocomponencial, que veurem més avall.[17]

Cal dir que les definicions hiperonímiques, que tan importants són en les definicions substantives i, en menor grau, en les verbals, en les definicions d'adjectiu són gairebé inexistents, de manera que gairebé podríem parlar d'un tipus de definició atípic.

2. *La definició antonímica*

El tercer tipus de definició lèxica és la que consisteix en la negació d'una definició sintètica o bé perifràstica que té per descriptor un antònim del verb definit. Aquesta definició només es pot donar pròpiament dins de les categories verbal i adjectiva. Per ex.:

**encertar** *v. tr.* No errar. [...]
**incògnit -a** *adj.* i *m.* i *f.* No conegut. [...]

Molt més correntment, el component de negació de la definició apareix motivat per l'estructura morfèmica del definit. Aleshores la definició antonímica és de naturalesa morfosemàntica:

**desobeir** *v. tr.* No obeir. [...]
**atípic -a** *adj.* No típic; [...]

(15) Per a una exposició dels problemes de caracterització d'aquesta relació vegeu Cruse (1986).

(16) Podem provar que *campament* és hiperònim de *bivac* amb la següent enunciació: bivac *és un tipus de* campament. Igualment podem dir que *velejar* és una manera de *navegar* o que *bai* és una modalitat de *roig*. Seguint la lògica del mètode, els co-hipònims, per exemple, de *velejar* haurien de tenir també *navegar* com a definidor genèric; vet aquí els verbs intransitius que en el *DGLC* s'han tractat com a co-hipònims de *velejar*: *barquejar, bordejar, costerejar, forejar, llevantejar, orsar, ponentejar, singlar, terrejar, voltejar*.

(17) En casos que el definit és compost, quan el verb té un significat componencial es recorre a la definició hiperonímica morfosemàntica, que explicita la relació existent entre els morfemes que componen el definit:

**colltorçar, colltòrcer** [...] *v. intr.* i *pron.* Tòrcer el coll [...]

**cornuar-se** *v. pron.* Nuar-se-li a algú el cor.

a diferència dels casos en què el significat que es defineix no és la suma dels components del mot entrada, com en el cas següent, que presenta una definició semocomponencial:

**sangglaçar-se** *v. pron.* Tenir un fort esglai. [...]

Algunes definicions de naturalesa conjuncional, això és, amb dos descriptors units mitjançant una conjunció, segueixen el patró de definicions antonímiques, de manera que s'estableix una relació lèxica de tres elements, on el definit es caracteritza mitjançant la negació dels dos descriptors, com per ex.:

**empatar** *v. intr.* No guanyar ni perdre en una partida de joc, en una votació, per haver guanyat el mateix nombre de punts, per haver reunit el mateix nombre de vots.

Sigui com sigui, la definició antonímica és poc freqüent en el *DGLC*.[18] Per Rey-Debove (1967) aquest tipus de definició: *est alors une négation de l'opposé du défini, c'est-à-dire une définition substantielle négative* (pàg. 153).

D'altra banda, hom ha considerat també antonímica (cf. Martin 1977, 365) una definició com ara:

**morir** [...] *v. intr.* Cessar de viure. [...];

tanmateix, tot i que en un cas com aquest aparegui un antònim del definit a l'enunciat definidor, aquest no fa de descriptor i tampoc no podem parlar de definició negada, sinó més aviat de definició *per inclusió negativa*. Rey-Debove (1967) ja assenyalava a propòsit de casos com aquest que:

> *Si l'on compare* Laisser = ne pas prendre *à* Laisser = négliger de prendre, *on constate que* négliger de *a même fonction que* ne pas, *et sert à nier. Mais* négliger de *présente les caractères de l'incluant général: premier mot de même catégorie grammaticale que le défini, tout le sémème du définissant dans celui du défini* (pàg. 154).

Es pot trobar, en el *DGLC*, algun correlat nominal d'aquestes darreres definicions:

**albaïna** *f.* Manca de vent, calma.

tot i que són molt més freqüents les morfosemàntiques:

**immodèstia** *f.* Manca de modèstia.

### B. Les definicions semocomponencials

En els casos de definicions semocomponencials és freqüent que hi hagi una relació de *paronímia*, això és, de terme derivat i primitiu derivatiu, entre el definit i un dels termes de la seva definició; el terme definit sol ésser el parònim i un dels elements definidors la seva base.[19] Aquesta característica no ens ha d'estranyar si considerem que, com hem dit més amunt, les definicions semocomponencials permeten donar compte de la transcategorització que ha sofert el terme definit. La paronímia, d'altra

(18) No sempre que trobem una definició negada es tracta d'una definició antonímica. En el *DGLC* trobem a vegades definicions negades que no presenten cap antònim del definit, ja que es tracta de la negació d'una definició semocomponencial:

**bellugar** [...] || Pron. [...] no estar quiet. [...]

**desservir** *v. tr.* No prestar (a algú) el servei o serveis acostumats o deguts.

(19) Seguim la terminologia de Cruse (1986: 130).

banda, és la relació dominant en les definicions morfosemàntiques, que desglossen el definit directament en els seus components morfèmics.

En un altre ordre de coses, cal assenyalar que si més amunt hem dit que una definició *hiperonímica* era la suma d'un descriptor hiperònim del definit més unes *differentia*, una definició semocomponencial és la suma d'un descriptor que regeix una determinada complementació i aquesta mateixa complementació. Aquesta condició no permet que el descriptor d'una definició semocomponencial passi la prova de la hiperonímia, ja que: *lorsqu'on pratique la lecture d'inclusion minimum, cette prédication devient fausse, alors même que la lecture totale est vraie* (Rey-Debove, 1971: 238).

1. *Els noms*

El *DGLC* té un nombre molt elevat de definicions bastides sobre fórmules regulars que en el diccionari resulten productives segons la importància quantitativa d'un determinat fenomen. En el cas de les definicions nominals, aquestes fórmules regulars representen un nombre molt elevat de casos en termes absoluts, i una gran majoria del diccionari en termes relatius.

Els elements constitutius d'aquestes definicions pertanyen a allò que Dubois i Dubois (1971) anomenen «nomenclature sémantique générique» i que: *est formée de termes comme* espèce, sorte, forme, qualité, action, resultat, se dit de, effet de, *etc.* (pàg. 59).[20] Aquests marcadors genèrics són generalment, en el discurs lexicogràfic, un tipus d'unitats a les quals no es concedeix, en general, cap valor metalingüístic especial.[21]

El tipus de marcadors descrits anteriorment, que en el *DGLC* corresponen correntment a *acció, efecte, qualitat, estat, etc.* tenen com a denominador comú el fet que permeten la transcategorització de l'element caracteritzador de la descripció semàntica (que normalment passa a ésser un verb o un adjectiu, en el cas de les definicions nominals). Així, una definició com:

**insult** *m.* Acció i efecte d'insultar; (...)

trasllada a la categoria verbal la descripció del significat *lèxic* del definit, i se serveix d'elements genèrics per donar raó del seu significat morfosintàctic. El resultat és, doncs, una definició de tipus *semocomponencial*.

Molt més correntment, però, aquesta mena de definicions estan motivades des del punt de vista de l'estructura morfològica del definit, en el qual cas podem parlar amb més propietat de definicions *morfosemàntiques*. Per ex.:

**maldiença** *f.* Acció de maldir; qualitat de maldient.
**dramatització** *f.* Acció de dramatitzar; l'efecte.
**esclavatge** *m.* Estat d'esclau.
**consultació** *f.* Acció i efecte de consultar, consulta.
**governament** *m.* Acció i efecte de governar, l'exercici del govern.

(20) Val la pena assenyalar que les relacions circulars entre aquest conjunt d'elements de metallengua són força freqüents:

**fet** *m.* Acció de fer alguna cosa; allò que hom ha fet.

**acció** *f.* Manifestació d'una força, manera d'obrar d'una causa, acompliment d'una funció. [...]

**efecte** *m.* Resultat de l'acció d'una causa.

**resultat** *m.* Allò que resulta.

**estat** *m.* Manera d'ésser, de trobar-se, en un moment donat, algú o alguna cosa.

(21) La gran majoria dels diccionaris de la llengua bastits segons els principis de la lexicografia tradicional no manifesten que calgui considerar aquestes unitats com a especials des del punt de vista de la descripció del significat.

**incubació** *f.* Acció de covar o incubar; l'efecte.
**albinisme** *m.* Estat del qui és albí, qualitat d'albí.
**combustibilitat** *f.* Qualitat de combustible.
**entonament** *m.* Estat o qualitat del qui és entonat.

Sovint algunes definicions nominals de naturalesa semocomponencial es basen en relacions semàntiques del tipus *part-tot* i de la seva inversa. Són les que podríem agrupar sota el denominador més o menys imprecís de *definicions meronímiques.* Aquesta relació semàntica (cf. Cruse, 1986) és difícil de caracteritzar des del punt de vista estrictament lingüístic i és una de les més interferides per l'univers de referències extralingüístiques. Cal fer notar, a tall d'exemple, la diversitat de relacions que hi poden resultar compreses, com *membre-grup, membre-classe* i *membre-col·lecció*, a més de la canònica *part-tot*, que n'implica altres com *peça-màquina*, etc.

Des del punt de vista lexicogràfic, aquesta relació presenta un alt índex d'aparició en els diccionaris tradicionals, i en especial en el *DGLC*, on se'n detecten molts casos que poden ésser motivats (a) o no (b) per la composició morfològica del definit.

a) **capellaneria** *f.* Conjunt de capellans; estament clerical.
**arrelam** *m.* Conjunt d'arrels.
**cosinalla** *f.* Conjunt de cosins.

b) **crina** *f.* Conjunt de crins del bescoll d'un cavall.
**clima** *m.* Conjunt de condicions atmosfèriques pròpies d'una regió, esp. en tant que afecten la vida animal i vegetal.
**trullada** *f.* Conjunt d'olives que es molen d'una vegada.

De manera semblant als descriptors genèrics *acció, efecte*, etc., la fórmula *conjunt de* no caracteritza semànticament el gènere o la classe del definit, sinó que introdueix respecte a aquest una relació de tipus *elements - conjunt d'elements*, on l'element és justament allò que més s'acosta al caracteritzador semàntic de les definicions hiperonímiques. Contràriament al cas anterior, però, no es produeix aquí la *transcategorització* de l'element semànticament caracteritzador.

Altres fórmules que s'utilitzen en el *DGLC*, juntament amb *conjunt* per a l'explicitació d'aquestes relacions, són també *reunió, colla,* etc.[22] Per ex.:

**llibreta** *f.* Reunió de fulls de paper disposats com els d'un llibre, que serveix per a fer anotacions, etc.
**flota** *f.* Reunió de persones que van plegades. [...]|| Una flota de bolets, grup de bolets units per llur base que viuen al peu dels arbres o rabasses mortes. [...]|| #Esp. Conjunt de naus de guerra o mercants d'un Estat, d'una companyia de navegació.
**prohomenia** *f.* [...] conjunt, reunió, de prohoms.

(22) *Grup*, en contra del que es pugui pensar en la llengua usual, té un sentit gairebé exclusivament tècnic al diccionari de Fabra, i es refereix quasi sempre exclusivament a classificacions botàniques i zoològiques. En comptadíssimes ocasions (cf. per exemple s.v. *diftong, incís*) podríem dir que *grup* s'utilitza sense tenir en compte aquest sentit.

**nació** *f.* Reunió d'homes, generalment de llengua comuna, que tenen institucions i costums anàlegs i un sentit d'homogeneïtat social i interès mutu; conjunt d'habitants d'un país regit pel mateix govern.
**canilla** *f.* Colla de gossos que cacen plegats.

La relació inversa d'elements que formen part (intuïtivament o no) de conjunts d'elements també hi és present. Per ex.:

**passapunt** *m.* Cadascun dels punts llargs sobre els quals es fan els traus, certs brodats, etc.
**vèrtebra** *f.* Cadascun dels ossos que componen la columna vertebral o espinada. [...]
**fibra** *f.* Cadascun dels elements filiformes que entren en la composició dels teixits orgànics o presenten en llur textura certs minerals. [...]
**corassot** *m.* Cadascun dels tanys drets que treu un arbre esporgat.
**comensal** *m.* i f. Cadascun dels qui mengen plegats a taula.
**capfoguer** *m.* Cadascun dels ferros o pedres que sostenen els tions en una llar de foc.

2. *Els verbs*

El verb és la categoria gramatical que presenta una varietat més rica de comportaments funcionals: ultra el caràcter verbal, el verb es pot caracteritzar pels trets de transitivitat, intransitivitat, impersonalitat. La definició semocomponencial dels verbs prendrà una o altra forma depenent d'aquests trets, fins al punt que a vegades recorre a un descriptor que no és sinó un auxiliar sintàctic de transitivitat (com és el cas del *fer* causatiu):

**encarcanyar** *v. tr.* Fer empassar a algú (una beguda). [...]

Aquest auxiliar de transitivitat pot arribar a constituir tota la informació nova d'una definició anafòrica:

**bullir** [...] || Tr. Fer bullir. [...]

L'ús del descriptor causatiu en les definicions d'accepcions transitives contrasta amb l'ús del descriptor de canvi *esdevenir* (o *tornar-se*) de moltes definicions d'accepcions intransitives (pronominals o no):

**atorrentar-se** *v. pron.* Esdevenir torrent.
**doblar** [...] || Intr. Esdevenir doble. [...]

En la definició de verbs adjectivals (sc. formats a partir d'un radical adjectiu) transitius es poden combinar aquests dos descriptors o bé aparèixer el descriptor factitiu *fer*:

**entebeir** *v. tr.* Fer esdevenir tebi. [...]
**sanejar** *v. tr.* Fer sa; [...].[23]

En general, en aquest tipus de definicions es fa la descomposició semàntica del definit recorrent al que anomenem descriptors *semosintàctics*, això és, que apareixen en funció del comportament gramatical que es vol descriure, i que es poden agrupar en dues llistes diferents, observant quins s'utilitzen en les accepcions transitives i quins en les intransitives. En les definicions d'accepcions transitives tindrem descriptors de la mena de *fer* (auxiliar) + INF; *donar, imprimir, efectuar, comunicar*; *treure, llevar, eliminar*; *causar*; *convertir*; *sotmetre, posar*; *tenir, deixar*, etc., i en les definicions d'accepcions intransitives descriptors com *prendre, agafar, adquirir, rebre*; *perdre, contreure, sofrir*; *esdevenir, tornar-se*; *estar, ésser*; *tenir, restar*, etc. El recurs a una mena o altra de descriptor semosintàctic és força sistemàtic (sempre amb variacions) en el *DGLC*:

**dutxar** *v. tr.* Donar una dutxa. | Pron. Prendre una dutxa.
**acomiadar** *v. tr.* Donar comiat. | Pron. Prendre comiat.
**engandulir** *v. tr.* Donar peresa. | Pron. Agafar peresa.
etc.

Quan el definit és morfològicament derivat pot aparèixer la definició *morfosemàntica*, la qual només és viable en aquells casos que el significat del lexema que es vol descriure és resultat directe de la suma del significat dels morfemes components. Per ex.:

**retastar** *v. tr.* Tornar a tastar.

on el segment «tastar» de la definició no és sinó el radical del definit, i el segment «Tornar a» de la definició proporciona una equivalència semàntica del morfema derivatiu *re-* del definit.

En un nivell de classificació més secundari trobem, encara dintre de les definicions semocomponencials, les definicions dites *conjuncionals* (les quals consten de més d'un descriptor annexats mitjançant una conjunció i que descriuen dos processos diferents que formen el concepte definit). Per ex.:

**borrufar** *v. intr.* Nevar i fer vent a la vegada.
**arrupir-se** *v. pron.* Contreure i estrènyer els músculs del cos. [...].

(23) Observem el paral·lelisme amb aquests casos quan es posen en contacte accepcions d'un tipus i d'un altre:
**aflonjar** *v. tr.* Fer tornar flonjo. || Pron. Tornar-se flonjo.
**agrir** *v. tr.* Fer tornar agre. [...] | Pron. Esdevenir agre. [...]
**emperesosir** *v. tr.* Fer esdevenir peresós. | Pron. Esdevenir peresós.
**enrogir** *v. tr.* Fer tornar roig. | Pron. Esdevenir roig.
**adequar** *v. tr.* Fer adequat. | Pron. Esdevenir adequat.
**alentir** *v. tr.* Fer més lent. [...] | Pron. Esdevenir més lent.
**etc.**

3. *Els adjectius*

Les definicions bàsiques d'aquest grup són les que consten del que podríem anomenar una «fórmula definidora» seguida d'un complement. Aquesta fórmula té la funció d'expressar la relació que hi ha entre el mot definit i el complement que la segueix (que acostuma a ser el mot del qual deriva l'adjectiu o un sinònim), de la mateixa manera que ho fan el relatiu i la preposició en les definicions de relatiu i preposicionals. Les definicions d'adjectiu substancials, doncs, quan no són sintètiques acostumen a no ser gaire més que definicions relacionals introduïdes per un adjectiu. Vegem-ne uns quants exemples:

**bacil·lar** *adj.* [...] || Pertanyent o semblant als bacils; causat per la presència de bacils.
**baromètric -a** *adj.* Relatiu o pertanyent al baròmetre, indicat pel baròmetre.
**basàltic -a** *adj.* Format de basalt, relatiu o pertanyent al basalt.
**baleàric -a** *adj.* Pertanyent a les illes Balears o a llurs habitants.
**baralladís -issa** *adj.* Amic de baralles, propens a barallar-se.
**barbós -osa** *adj.* Ple de barbes.
**barrejadís -issa** *adj.* Fàcil de barrejar.

Hi ha un grup de fórmules definidores que compleixen especialment bé la funció relacional de què acabem de parlar; són les que comencen per «pertanyent», «relatiu», «pertanyent o relatiu», «relatiu o pertanyent», «pertanyent o referent» i altres fórmules semblants, que formen un grup molt important dins les definicions d'adjectius. Per ex.:

**babilònic -a** *adj.* Pertanyent o relatiu a Babilònia.
**bacanal** *adj.* Pertanyent al déu Bacus.
**bacil·lar** *adj.* Pertanyent o semblant als bacils; [...]
**bacterial** *adj.* Relatiu o pertanyent a les bactèries; [...]
**balener -a** *adj.* Relatiu a la balena.

Aquest grup de fórmules definidores, totes diferents però amb el mateix significat, en realitat no té altra funció que la d'indicar una relació entre l'adjectiu i el seu primitiu. Si observem la definició de *balener -a*, ens adonem que l'únic que indica aquesta definició és que el mot definit és un adjectiu derivat de «balena».

Altres fórmules definidores tenen un grau més alt de contingut semàntic, com hem pogut veure amb els exemples anteriors, les quals solen respondre a l'explicitació del significat de determinats sufixos (vegeu, més amunt, s.v. *basàltic*); i, en aquests ca-

sos, s'observa que determinats sufixos i determinades fórmules definidores acostumen a anar lligats.

VICTÒRIA ALSINA I KEITH,
LLUÍS M. SOL I SALVADÓ, JOAN SOLER I BOU
*Institut d'Estudis Catalans*

## REFERÈNCIES BIBLIOGRÀFIQUES

BOSQUE, Ignacio (1982) «Sobre la teoría de la definición lexicográfica». A: *Verba. Anuario galego de filoloxía*. Vol. 9, Universidad de Santiago de Compostela, 1982, pàgs. 105-123.

CRUSE, D.A. (1986) *Lexical semantics*. Cambridge, Cambridge University Press.

[*DGLC*] FABRA, P. (1932) *Diccionari general de la llengua catalana*. Barcelona, Edhasa.

DUBOIS, J; DUBOIS, Cl. (1971) *Introduction à la lexicographie. Le dictionnaire*. Paris, Larousse.

LYONS, J. (1968) *Introduction to Theoretical Linguistics*. Cambridge, Cambridge University Press. [Traducció en castellà (1986) *Introducción en la lingüística teórica*. Barcelona, Teide.].

MARTIN, R. (1977) «Essai d'une typologie des définitions verbales dans le dictionnaire de langue». A: *Travaux de Linguistique et de Littérature*. XV, 1, Strasbourg, 1977, pàgs. 361-378.

RAFEL I FONTANALS, J. (1988a) «Sobre la normalització dels diccionaris catalans: el tractament de les sèries sinonímiques». A: *Estudis de llengua i literatura catalanes*. XVII, Montserrat, Publicacions de l'Abadia de Montserrat, 1988, pàgs. 57-112.

REY-DEBOVE, J. (1967) «La définition lexicographique; bases d'une typologie formelle». A: *Travaux de Linguistique et de Littérature*. V, 1, Strasbourg, 1967, pàgs. 141-159.

— (1971) *Étude linguistique et sémiotique des dictionnaires français contemporains*. Paris, Mouton.

SOL, Ll. (1992) *El* DGLC *i la formalització lexicogràfica dels verbs*. Tesi de llicenciatura dirigida pel Dr. Joaquim Rafel i Fontanals, Departament de Filologia Catalana, Universitat de Barcelona. [inèdit]

WEINREICH, U. (1960) «Lexicographic Definition in Descriptive Semantics». A: Householder, F.; Saporta, S. (eds.) (1962) *Problems in Lexicography*. Bloomington, University of Bloomington, pàgs. 25-44.

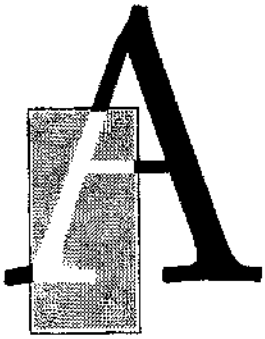
WITTGENSTEIN, L. (1953) *Investigaciones filosóficas*. Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México. [Barcelona, Crítica, 1988].

ZGUSTA, L. (1971) *Manual of Lexicography*. The Hague, Mouton.

# Part III:

# ENTREVISTES

# ENTREVISTA A ALAIN REY, DIRECTOR DELS DICCIONARIS LE ROBERT



**Alain Rey, director lexicogràfic de l'editorial Le Robert, és autor o coautor de diversos diccionaris francesos i autor de molts treballs sobre lexicologia, lexicografia i terminologia. Acaba de publicar el *Dictionnaire Historique de la Langue Francesa* del qual s'ha exhaurit la primera edició en tres mesos. El rigor del seu treball, juntament amb el de M. Josette Rey-Debove, ha donat a l'editorial Robert el prestigi que té actualment en el camp lexicogràfic. Alain Rey és avui una de les figures més representatives de la lexicografia francesa actual.**

**P. Sr. Alain Rey, parli'ns de la seva llarga experiència lexicogràfica.**

Vaig començar a treballar en lexicografia, ara deu fer uns quaranta anys, ben bé per casualitat. Jo no havia fet estudis de lingüística ni de Filologia; havia estudiat Ciències Econòmiques i Sociologia, i, posteriorment havia completat els meus estudis amb una especialització en llengua anglesa, que vaig fer als Estats Units, i en Història de l'art, concretament en l'art de l'Edat Mitjana. La meva formació pròpiament dita, doncs, no era lingüística. Però un dia vaig conèixer en Paul Robert, i ell, que tampoc no era lingüista, va engrescar-me en el seu projecte lexicogràfic; i amb la meva feina al que seria més tard el *Grand Robert* vaig entrar en el món dels diccionaris.

*Caplletra 17* (Tardor 1994), pp. 327-330

El treball del *Gran Robert* em va fer plantejar molts dels problemes que s'han de resoldre en lexicografia, ja que és un diccionari que conté tots els aspectes del lèxic, llevat només de l'aspecte fonètic. Per a mi, doncs, va significar un aprenentatge important, una iniciació general i completa en el món dels diccionaris.

Amb l'experiència del *Grand Robert* vaig iniciar el *Petit Robert*, conjuntament amb M.Josette Rey-Debove i Henry Cottez. Aquest nou diccionari ens va fer plantejar qüestions complementàries de les que ens havíem plantejat amb el *Grand Robert*. Per exemple, el tema de la relació entre els diccionaris i el públic, entre els diccionaris i la seva utilitat social. En el *Petit Robert*, que amplia la nomenclatura del *Grand Robert*, vam haver de fer una opció de model lingüístic per descriure el lèxic. Així, vam optar pel model distribucional, en contrast amb el model transformacional que Jean Dubois havia adoptat per al *Dictionnaire du français contemporain* de l'editorial Larousse. Les opcions que vam fer per al *Petit Robert* es basen sobretot en criteris d'utilitat i facilitat de consulta. Els usuaris van ser, doncs, l'eix de la nostra atenció.

Del *Petit Robert* vaig passar al diccionari de noms propis, que encara em va fer plantejar aspectes nous que no m'havia hagut de plantejar anteriorment. En aquest cas, les opcions responien a criteris de tipus cultural.

Després del diccionari de noms propis, vaig realitzar, juntament amb Cellard, el diccionari d'argot, que recull un bon nombre d'expressions d'aquest registre extretes d'obres literàries. I del diccionari d'argot passo al de locucions, que ha estat una obra molt important, tant pels seus continguts com per l'èxit comercial que ha tingut. Se n'han fet unes tretze edicions, deu de les quals contenen alguna nova aportació. Aviat n'apareixerà una altra edició que inclou sis-centes expressions noves procedents de fonts diverses, entre les quals voldria esmentar el *Dictionnaire du français parlé* elaborat per dos investigadors del CNRS i editat a Seuil. Aquesta nova edició inclourà, a més, fragments que, si bé no són pròpiament locucions, apareixen amb molta freqüència en determinats registres comunicatius com, per exemple, en la publicitat.

I del diccionari de locucions, que forma part de la sèrie dels «Usuals» del Robert, al *Diccionari Històric*, del qual després parlaré més detalladament.

En definitiva, que he participat durant molts anys en la tasca lexicogràfica sense que, a l'inici, tingués cap preparació especial per dur-la a terme. He après a fer diccionaris, fent-ne; és així. Del nostre equip inicial, M. Josette Rey-Debove, Henry Cottez i jo, només Cottez era filòleg; els altres dos no érem ni filòlegs ni lingüistes. Jo vaig entrar en la lingüística a través de la meva relació amb lingüistes de renom, com Chevalier o Pottier.

**P. Doncs, amb la seva experiència, com concep vostè la lexicografia: com una disciplina? Com un art?**

Miri, el meu objectiu essencial quan faig un diccionari és poder arribar a una mena de compromís entre els coneixements lingüístics en sentit general i el treball pràctic, que és artesanal. Per mi, la lexicografia és per damunt de tot una activitat artesanal que construeix un objecte social, que és el diccionari. Aquest objecte, però, ha de ser al mateix temps un objecte útil i una obra d'art, en el sentit més ple de la paraula.

En aquest sentit, penso que amb els Robert hem creat una mena d'estil lexicogràfic. Així ho van reconèixer Wagner i Chevalier quan va sortir el *Petit Robert.*

És important de tenir en compte que un diccionari no es pot limitar només a ser útil; cal també que sigui un producte ben fet i adequat, ja que un diccionari és una opció entre moltes possibilitats. Per exemple, el nou *Diccionari Històric* que acabem de publicar és una obra que ha fet opció per un estil narratiu, ben lluny, doncs, de l'estil dels diccionaris etimològics clàssics, que són tan feixucs de llegir. El nostre diccionari històric, si bé no està adreçat al gran públic, pot concentrar un major nombre de lectors que no pas els diccionaris clàssics.

**P. Ja que ha iniciat el tema del *Diccionari Històric*, que es va publicar a finals de 1992, faci'n una breu presentació.**

En primer lloc haig de dir que jo no en sóc l'únic autor, sinó que he comptat amb la col·laboració de Tristan Honoré, Marianne Tomi i Chantal Tanet. En segon lloc voldria remarcar que no hem fet un diccionari etimològic com el de Coromines per al català i el castellà, o com el de Pfister per a l'italià. Hem fet un diccionari històric que recull, en la línia de Wartburg, l'etimologia, però també altres informacions sobre l'evolució dels mots, sobretot informacions de tipus cultural. El nostre diccionari, en contrast amb els altres, s'adreça a un públic més ampli, perquè està redactat en un estil narratiu, com ja li he comentat, i entenedor. De fet, no devem haver fet una mala opció, perquè la primera edició ja està exhaurida. Ara treballem ja en les esmenes del que serà la segona edició. També és un fet que ha despertat un cert interès social; m'han fet diverses entrevistes a la ràdio i a la televisió sobre aquest diccionari i s'han publicat força comentaris als diaris. Gràcies a aquesta divulgació, es pot generalitzar una mica l'interès per la història dels mots. Ara, per exemple, el diari *Libération* m'ha demanat que els escrigui un article llarg sobre la història del mot «reforma». Això significa, doncs, que la publicació del nostre diccionari històric ha desvetllat un cert interès per l'evolució general dels mots. Em sembla que un diccionari com aquest pot concentrar tres grans interessos: un interès general dels francesos per la llengua; un segon interès per la llengua francesa en particular com a llengua nacional; i un tercer interès

per la pròpia història cultural. De fet els francesos sempre han tingut un interès remarcable per la llengua. La història de la lexicografia francesa així ho mostra, sobretot a partir del segle XVI. A l'Edat Mitjana el discurs sobre la llengua no existia, ningú no parlava de llengua; va ser amb el Renaixement italià que es va iniciar aquest discurs, concretament amb Dant.

**P. Parlem ara sobre el darrer diccionari de Robert: el *Diccionari del francès de Quebec*, que sembla ser que ha aixecat una certa polèmica entre els quebequesos.**

Efectivament, l'editorial Robert ha publicat aquest diccionari de la varietat quebequesa del francès. Considero que seria molt interessant de disposar de diccionaris per a cada una de les grans modalitats de la llengua francesa. La idea de fer aquest diccionari era molt antiga; jo n'havia parlat amb diversos amics de Quebec. En la meva opinió els nens quebequesos s'han d'escolaritzar amb un diccionari de la seva pròpia varietat lingüística; i els immigrants que hi viuen també han d'aprendre la llengua a partir d'aquesta varietat, que és, a fi de comptes, la que faran servir.

El *Diccionari del francès de Quebec* és una obra ben feta. Potser insisteix una mica massa en els col·loquialismes i això li dóna un cert desequilibri. Però, per damunt de tot, és una obra ben feta. Ha aixecat, en efecte, algunes crítiques negatives, però que sembla ser que ja reculen, perquè de fet es concentren en uns pocs aspectes: en algunes paraules, que són especialment conflictives i difícils d'acceptar com a mots de la llengua, i en algun aspecte del tractament lexicogràfic dels mots, com per exemple el fet de no marcar els anglicismes com a tals sinó com a mots familiars, amb l'etiqueta «fam.»; però en sentit global és un diccionari fet per quebequesos des de la perspectiva del Quebec (en aquest sentit marca com a francismes els mots que només s'usen a França però no a Quebec), sobre la base del *Micro Robert*. Que jo sàpiga, és la primera vegada que un diccionari reuneix totes aquestes característiques. I això és el més important.

M. TERESA CABRÉ CASTELLVÍ
*París, febrer 1993*

# ENTREVISTA A JOHN SINCLAIR, EDITOR EN CAP DEL COLLINS COBUILD ENGLISH LANGUAGE DICTIONARY

**John M. Sinclair és professor del Departament d'anglès modern de la Universitat de Birmingham des de 1965. Les àrees de recerca més importants en què ha treballat són el discurs oral i escrit i la lingüística computacional, amb atenció especial a l'estudi de textos. És el director del projecte COBUILD des del seu inici i l'editor principal de les publicacions del COBUILD. D'entre la seva extensa bibliografia, volem destacar per la seva importància el llibre que descriu el projecte COBUILD (Sinclair, J. (ed.) *Looking up. An account of the COBUILD Project in lexical computing*. Londres, Collins 1987), el seu darrer llibre sobre corpus (Sinclair, J. *Corpus, concordance, collocation*. Oxford, Oxford University Press 1991) i l'obra encara en premsa sobre lèxic (Sinclair, J. et al. *English word in use*. Londres, Collins).**

**P: Professor Sinclair, podria explicar-nos com va néixer el projecte COBUILD?**

El projecte va néixer cap al final dels anys setanta amb un conveni de l'editorial Collins i la Universitat de Birmingham.

L'editorial Collins tenia una gran reputació en lexicografia bilingüe i monolingüe, sobretot de l'anglès, i estava molt interessada a elaborar altres productes i obrir

nous mercats. La Universitat de Birmingham feia aproximadament uns 15 anys que desenvolupava mètodes per crear un corpus en suport magnètic. Això passava cap a l'any 1962. Així va néixer l'acord de desenvolupar sistemes informàtics i software per fer estudis de vocabulari. I jo em vaig fer càrrec de materializar aquest acord.

En aquell moment el corpus de què disposàvem era de mots simples. Amb M. Halliday, em vaig adonar que calia disposar de textos sencers organitzats en un gran corpus per poder dur a terme amb rigor estudis sobre el llenguatge i el vocabulari; que calia crear un corpus d'almenys 5 milions de mots, o probablement encara més ampli.

Cal que aclareixi que el meu interès, de fet, no és la lexicografia pròpiament dita, sinó els contextos (*collocations*) i, per fer estudis sobre contextos, necessito una gran quantitat de text, ja que la freqüència d'aparició d'un mot no és gaire alta i fins que no es disposa d'unes 10, 15 o 20 ocurrències del mateix mot no es pot considerar que un context està suficientment establert.

Tornant al COBUILD, ens vam trobar que, d'una banda, nosaltres necessitàvem software i material en suport magnètic per a les nostres recerques, i que, d'altra banda, Collins volia un diccionari. I així, el 1980 la Universitat de Birmingham va signar un primer contracte amb Collins, de 5 o 6 anys de durada, amb el compromís de produir un diccionari a partir d'un corpus, usant els mètodes que havíem desenvolupat en el meu equip. I d'aleshores ençà la col·laboració no s'ha aturat.

El primer gran projecte que va sortir del corpus va ser el diccionari COBUILD, que ha estat amplament acceptat arreu i ha rebut molt bones crítiques. Per a l'elaboració del diccionari, partíem de les idees tradicionals de la lexicografia, però el fet de partir d'un corpus ens va fer canviar de plantejaments. El diccionari COBUILD és de fet un producte molt original, per moltes raons, gairebé totes elles lligades al fet que ha estat construït a partir d'un corpus textual. He analitzat la lexicografia dels darrers 20 anys i puc dir que el COBUILD és molt més modern que els altres diccionaris, i també molt diferent. En molts aspectes. Posem alguns casos, els exemples del COBUILD vénen dels textos i les definicions surten dels contextos. La fraseologia que conté és molta i molt important. És en gairebé tots els aspectes que el COBUILD és diferent dels diccionaris existents. I aquesta especificitat rau en el fet que prové d'un corpus, d'un corpus de textos que fa evidents els fets del llenguatge.

Quan es va acabar el període del primer contracte (1986-1987) vam constatar que tant la Universitat de Birmingham com l'editorial Collins havíem quedat satisfets del treball conjunt i vam crear una nova societat en col·laboració. Disposàvem aleshores ja d'un equip d'especialistes. Aquesta nova associació entre la Universitat de Birmingham i la casa Collins va néixer per produir altres tipus de treballs, a més dels lexicogràfics, com l'elaboració d'una gramàtica i un manual d'ús de la llengua.

De fet ara ja existeixen diverses publicacions que hem fet en cooperació. Aquestes noves obres neixen amb el propòsit fonamental de proporcionar als usuaris els beneficis

d'haver usat un corpus. Els productes que hem publicat ja són força nombrosos: tenim el *Cobuild English Course*, diverses obres breus destinades sobretot a la correcció (*sillabus*) i basades en la freqüència d'aparició de determinats fenòmens (la formació de paraules, les preposicions, els determinants, les paraules que es poden confondre, etc.), dues gramàtiques (una de més completa, l'*English Grammar*, i una altra més reduïda per als estudiants, l'*Student's Grammar*) i 6 diccionaris: el diccionari general *English Language Dictionary*, l'*Student's Diccionary*, el *Dictionary of Phrasal Verbs*, el diccionari reduït *Essential English Dictionary*, el *BBC English Dictionary* (especialment dedicat al llenguatge parlat i extret sobretot dels enregistraments de la BBC) i un diccionari escolar, que encara no s'ha publicat.

El tercer tema en què hem treballat és el de l'*English Usage Book*, que es va publicar l'any passat.

Ara treballem en la introducció d'elements gramaticals en els textos mitjançant l'etiquetatge dels mots (*tagging*) i l'ús d'analitzadors morfològics i sintàctics dels textos. Aquest sistema ens permetrà recuperar després determinades estructures a partir de la selecció de categories. Ho fem amb un software procedent de la Universitat de Helsinki.

El corpus, que al començament tenia 5 milions de mots, s'ha anat ampliant amb l'escanerització de textos i amb materials de diversos estudiants. Actualment el corpus conté 120 milions de mots i és molt flexible i de fàcil accés.

Avui dia un dels propòsits del projecte COBUILD és de connectar amb una màquina que permeti seleccionar i respondre preguntes sobre 120 milions de mots de diverses procedències. Com que els textos estan classificats de diverses maneres, l'usuari pot seleccionar perfectament diversos materials a partir de diferents criteris: que procedeixin d'una font o d'una altra, escrits o orals, d'una varietat d'anglès o d'una altra, etc.

**P: Parli breument dels aspectes tècnics del projecte COBUILD.**

Poques coses puc dir en profunditat sobre els aspectes més tècnics relacionats amb la informàtica, perquè no en sóc especialista. Però puc dir que quan vam començar el projecte, no hi havia especialistes en corpus lingüístics i vam haver-ne de formar a partir de col·legues especialistes en computació.

En l'etapa inicial del projecte vam utilitzar l'ordinador central de la Universitat, que precisament el van comprar per al projecte. Era una màquina Honeywell. Ara ja és una eina de museu. Però després hi va haver una època crucial en aquest sentit que va ser quan la Universitat va decidir canviar l'ordinador que tenia per un de nou més apte per al projecte.

Cap el 1987 la política de petits ordinadors va esdevenir impopular en el món acadèmic i es va muntar un immens ordinador central que donés suport a centenars d'usuaris. Així els equips que necessitàvem informàtica vam començar a desenrotllar els nostres propis sistemes.

Des de 1993 treballem amb màquines UNIX. Busquem sistemes molt flexibles i fàcils de manejar, amb rutines especials i al màxim d'estàndards. Necessitem que la qualitat dels terminals sigui molt alta. Ara, doncs, tenim accés al gran ordinador central i disposem de dos grans servidors de la nostra xarxa local; un que conté el corpus i un altre gran servidor que conté el software.

**P: Quin és l'estat actual del projecte?**

El treball més immediat en el present és la revisió del diccionari, publicat ara fa set anys, a partir de les crítiques que hem rebut d'arreu del món, sobretot dels professors d'anglès que l'han utilitzat. El diccionari ha estat rebut en general amb gran entusiasme, però hi ha aspectes que cal millorar.

Un segon camp en el qual anem treballant és un futur diccionari de contextos (*Collocation Dictionary*), a partir de la introducció d'etiquetes (*tags*) en el corpus, etiquetatge de què abans he parlat.

També estem preparant la versió electrònica en CD-ROM dels nostres treballs. Actualment tenim un prototip que conté tres llibres: el dicionari, la gramàtica i el manual d'ús; i, a més, inclou 5 milions de paraules del corpus. D'aquesta versió experimental se n'han fet 100 exemplars, que han estat distribuïts a diverses parts del món a fi de recollir-ne les impressions i les crítiques. Serà comercialitzat per l'Editorial Collins, i espero que despertarà el mateix entusiasme que les obres impreses, si les eines per accedir-hi s'abarateixen.

Un dels grans problemes de la informació electrònica en treballs de lexicografia és com controlar la quantitat d'informació. Necessitem eines que sumaritzin la informació. Per tant considerem que cal una eina de referència que controli la informació i el vocabulari. Crec que la utilització de corpus per fer diccionaris ha fet canviar la lexicografia. Permet fer productes més sofisticats, més evidents. El lexicògraf disposa de molta informació alhora; té al seu abast sistemes d'accés a la informació més flexibles i selectius. I la introducció d'analitzadors en els textos és un pas per al tractament automàtic del llenguatge i, a la llarga, si es disposa de la mateixa anàlisi per a dues llengües, és la base de la traducció assistida o automàtica.

Un altre dels projectes importants que tenim entre mans en el marc dels acords de col·laboració amb Collins és el dels diccionaris bilingües (*bilingual bridge dictionaries*). Són diccionaris bastant originals que parteixen de la paraula en anglès i l'expliquen en la llengua de traducció. El primer *diccionari pont* que s'ha fet és l'anglès/

português-brasiler. El bilingüe anglès/espanyol és en curs d'elaboració. Per fer aquests diccionaris treballem en col·laboració amb lexicògrafs natius de cada llengua. Al final del procés de treball, el nostre equip supervisa el treball i es permet discutir amb el lexicògraf natiu determinades qüestions que poden ser problemàtiques. Un dels problemes amb què ens trobem contínuament és el calc de determinades estructures de l'anglès quan es fan les explicacions en l'altra llengua. I també ens trobem amb el fet que determinades explicacions que són naturals en anglès no ho són en altres llengües. Aquests *diccionaris pont* són força originals en relació amb els bilingües tradicionals, en el sentit que intenten reproduir el procés intel·lectual que, en consultar-los, fa el parlant no natiu amb una competència reduïda en llengua anglesa.

Un exemple real seria el següent:

> **nail** /pron/, **nails, nailing, nailed**. 1. COUNT N **A nail** es una pieza metálica pequeña, con una punta en un extremo, la cual se golpea con un martillo para meterla dentro de algo. *The mirror that hung from a nail on the wall.* 2. VB WITH OBJECT AND ADJUNCT Si se dice que alguien **nails** algo en algún sitio, lo sujeta allí con un clavo. *They nail plastic sheets over their windows.* 3 COUNT N Tus **nails** son las láminas finas y duras que recubren la punta de los dedos. *He keeps biting his nails.* 4 Si dices que alguien **has hit the nail on the head**, quieres decir que con lo que ha dicho ha acertado plenamente. 5 **a nail in the coffin** de algo: busca **coffin**.

**P: Per acabar, quin balanç faria de la utilització dels ordinadors en el tractament del llenguatge i de l'elaboració de diccionaris a partir de corpus textuals?**

Penso que la introducció dels ordinadors en el tractament del llenguatge ha fet canviar totalment el panorama de la recerca sobre el llenguatge i el sistema de treball en lexicografia.

Els corpus textuals permeten disposar d'informació no intuïtiva i, si són suficientment amplis, donen una fiabilitat a les dades que altrament no poden tenir. Tres són al meu entendre les diferències fonamentals que la utilització de corpus confereixen al treball lexicogràfic. La primera és la major qualitat i diversitat de la informació: si es disposa d'un bon corpus es poden obtenir productes lexicogràfics més sofisticats i diversos; també es poden il·lustrar les unitats de manera més fiable, tot i que hi ha diversos aspectes que no estan encara ben resolts com, per exemple, la categorització pragmàtica de les paraules. El segon avantatge que ofereixen els corpus és la possibilitat d'establir la fraseologia: amb els corpus es poden detectar moltes estructures concurrents a cavall entre la fraseologia i la combinació freqüent de determinats mots, que no es podrien detectar sense la utilització de corpus; amb un corpus ampli de dades es pot aconseguir una variació lingüística suficient per donar compte de la realitat ex-

pressiva. Les *collocations* (concurrències freqüents) són importantíssimes per descriure les llengües. La semàntica dels mots només es pot establir a partir de la seva utilització en contextos variats. La tercera característica diferencial que ofereix el treball lexicogràfic a partir de corpus és l'evidència: els corpus permeten detectar expressions, significats i fraseologia que no s'havien detectat abans i il·lustrar els mots amb molta més fidelitat a l'ús.

La introducció de sistemes d'etiquetatge en els textos obre grans possibilitats de descripció del llenguatge, encara que cal superar determinats problemes, com el que suposa la classificació pragmàtica dels mots i les expressions.

El fet de poder disposar d'eines de tractament del llenguatge i d'aprenentatge de les llengües en suport electrònic canvia notablement les possibilitats de treball, sobretot si es pot fer ús d'aquestes eines en ordinadors personals.

Del naixement del projecte COBUILD fins ara les circumstàncies han canviat molt. És un fet que cada cop hi ha més corpus de textos i que avui dia no es pot plantejar la descripció d'una llengua sense disposar de materials que aportin l'evidència dels usos, i no es pot concebre l'elaboració de diccionaris i de gramàtiques sobre la base de la intuïció.

M. TERESA CABRÉ CASTELLVÍ
*Birmingham, 11 d'agost de 1993*

# ENTREVISTA A DANIELLE CORBIN I PIERRE CORBIN, DIRECTORS DEL DIPLOMA EUROPEU DE LEXICOGRAFIA*

**Danielle Corbin i Pierre Corbin, professors de lingüística de la Universitat de Lille III i responsables del Diplôme Européen de Lexicographie que s'imparteix a la Universitat de Lille III des de 1991, són especialistes en lexicologia i lexicografia respectivament, membres de l'equip SILEX del CNRS a la Universitat de Lille i de l'equip de redacció de la revista *Lexique*. Danielle Corbin és autora de múltiples treballs sobre lexicologia teòrica i descriptiva; entre la seva àmplia producció destaca el llibre en dos volums *Morphologie dérivationnelle et structuration du lexique* (Tubinga 1987). Pierre Corbin ha publicat nombrosos treballs de metalexicografia, centrats sobretot en la descripció del panorama lexicogràfic francès.**

**P. En la primera part d'aquesta entrevista voldríem plantejar qüestions generals sobre la lexicografia. Començarem per la seva realitat. Dieu literalment en el tríptic de presentació del Diploma Europeu en Lexicografia (DEL) que «no existeix actualment a Europa cap formació específica de lexicògraf, com tampoc no existeix a França cap formació específicament lexicològica en les professions relacionades amb les indústries del llenguatge». Per què creieu que es produeix aquesta situació?**

PC. Crec que aquesta formulació és exacta. Pel que fa a França, nosaltres no coneixem cap altra formació universitària en aquest domini que la que s'ofereix a Lille.

(*) Entrevista enregistrada el febrer de 1993. Versió revisada i actualitzada el novembre de 1993.

Pel que fa a Europa, els col·legues estrangers amb qui hem treballat en el DEL comparteixen la nostra opinió.

DC. Quant a les raons d'aquesta situació, no sé de quina naturalesa són pel que fa a Europa. Ara bé, en el cas de França, es pot deure al fet que el lèxic està relativament oblidat a la Universitat: és una matèria poc ensenyada. Cert que hi va haver un lexicòleg molt important, Darmesteter, a finals del segle passat i, que, més recentment, després dels anys seixanta, hi ha hagut estudis del lèxic des d'un punt de vista polític, lligats a l'anàlisi del discurs. Però l'anàlisi pròpiament lingüística del lèxic encara ara no està gaire desenvolupada a França. Vam fer, fa algun temps, una revisió sobre l'estat de la recerca en lexicologia a les universitats franceses, i vam constatar que només existeixen dos centres veritablement orientats a l'estudi del lèxic, París XIII i Lille III.

**P. Però, per contra, hi ha una llarga tradició en lexicografia francesa que indueix a pensar que la lexicografia ha estat des de sempre una matèria d'estudi.**

PC. Hi ha hagut una desconnexió entre lexicògrafs i universitaris que només molt recentment s'ha començat a superar. Per demostrar-ho, podem prendre com a referència dos textos de Josette Rey-Debove. L'un de 1971, que pertany a la introducció del seu llibre *Etude linguistique et sémiotique des dictionnaires français contemporains,*[1] diu que, fins als anys seixanta, «aucun dialogue n'a [...] pu s'établir entre linguistes et lexicographes»,[2] i això porta Josette Rey-Debove a caracteritzar, amb una mica de malícia, el «savoir-faire» paradoxal del lexicògraf, que, tot i ser un «mediocre encyclopédiste et médiocre linguiste», sabrà fer una cosa que ni un lingüista ni un especialista d'una disciplina no sabran fer: redactar articles de diccionari.[3] Però el 1991, en el treball anomenat «La lexicographie moderne» que acaba de presentar al col·loqui de Louvain-la-Neuve sobre l'estat actual dels estudis lexicals,[4] remarca l'apropament dels punts de vista i el diàleg que s'ha instaurat entre la lexicologia, que «bénéficie des travaux lexicographiques», i la lexicografia, que «tient compte des recherches des lexicologues».[5] Per tant, no sembla tampoc que hi hagi hagut una llarga tradició d'interès per la lexicografia per part de la Universitat; actualment, el nombre de persones i d'institucions interessades pel tema és encara relativament limitat; i, a més, l'ensenyament de la lexicografia ocupa sempre un lloc ínfim en els cursos universitaris francesos. La formació en l'ofici de lexicògraf, si és que existeix, es fa empíricament a les editorials de diccionaris. Recordem que, amb un segle de distància, Émile Litré, que havia fet estudis de medicina, i Paul Robert, que era advocat, van ser dues figures de lexicògraf destacades sense cap preparació específica. D'aleshores ençà, creus que han canviat substancialment les coses?

(1) L'Haia/París, Mouton.
(2) Pàg. 14.
(3) Pàg. 43.
(4) *Travaux de linguistique* 23, «Où en sont les études sur lexique? Bilan et perspectives» (André Goosse, Jean-René Klein & Jean-Marie Pierret (eds)), pàgs. 145-159.
(5) Pàg. 146.

DC. Cal afegir, que quan es va posar en marxa el *Trésor de la langue française*, als anys seixanta, es va procedir a un recrutament de redactors per mitjà de petits anuncis a la premsa, perquè no hi havia altres sistemes, i aquest procediment, que no és pas aïllat, és encara actual per una o altra editorial.

**P. Voleu dir que aquest canvi d'actitud que reflecteix Rey-Debove no respon a l'acostament que hi ha avui dia entre la Lingüística Teòrica i la Lingüística Aplicada; i, doncs, la lexicografia com una matèria de Lingüística Aplicada.**

DC. Tinc la impressió que el que s'entén a França per Lingüística Aplicada és essencialment la lingüística aplicada a la pedagogia, a l'ensenyament de la llengua materna o de les llengües estrangeres.

**P. Així doncs podem dir que la lexicografia no es concep com una matèria de Lingüística Aplicada?**

DC. No, a França no.

PC. Per completar aquesta resposta, es pot dir que constitueix una excepció el volum d'«Hommage à Bernard Quemada» titulat «Dictionnairique et dictionnaires» que ha estat publicat a la revista *Études de linguistique appliquée*[6] que ell mateix va fundar. Es pot també fer referència a la producció per part del CREDIF[7] de diccionaris contextuals,[8] on el projecte ha estat inscrit explícitament en la perspectiva de l'ensenyament del francès com a llengua estrangera.[9] Però aquests exemples particulars no semblen invalidar la nostra afirmació: són excepcions que confirmen la regla. Si s'analitza els *Études de linguistique appliquée*, es constata que aquesta revista, amb trenta anys d'existència, només ha publicat un petit nombre d'articles sobre lexicografia, i un sol monogràfic dedicat a la utilització dels diccionaris per part dels usuaris.[10]

**P. Un altre aspecte. On penseu que s'ha de formar un lexicògraf? A la Universitat? Fora de la Universitat? De manera combinada?**

DC. Si hem decidit d'oferir el DEL a Lille III, és perquè pensem que la Universitat pot contribuir a la formació dels lexicògrafs. Creiem que aquests han de ser abans que tot uns bons lingüistes. Cal doncs formar-los en lingüística, i això és el que es fa a la Universitat i no a les editorials.

(6) Número 85-86 (Robert Galisson (ed.)), 1992.

(7) Centre de Recerca i d'Estudi per la Difusió del Francès (Ecole Normale Supérieure de Fontenay/Saint-Cloud).

(8) *Dictionnaire contextuel de français pour la géologie*, Paris, CREDIF/Didier, 1976; *Dictionnaire contextuel de français politique*, microfitxes incloses a J. L. Descamps, et al. (eds.), *Sémantique et concordances*, París, Publications de l'INALF/Klincksieck, 1992.

(9) *Sémantique et concordances*, pàgs. 1-2.

(10) Número 49. «Image et usage du dictionnaire» (Robert Galisson (ed.)), 1983.

**P. Suposo que tu penses el mateix...**

PC. Cada any, el primer curs del DEL està dedicat a una reflexió de si és legítim o no que els universitaris, no exercint l'ofici de lexicògrafs, preparin estudiants en aquest tema. Ara bé, si examinem el que diuen els lexicògrafs, ja que alguns són també metalexicògrafs, constatem que les seves opinions estan dividides pel que fa a aquesta qüestió. L'americà Sydney Landau, per exemple, nega tota pertinença de la lingüística en la formació de lexicògrafs. A *Dictionnaires. The Art and Craft of Lexicography,*[11] escriu que no només no és útil que els lexicògrafs siguin lingüistes, sinó que l'experiència demostra que val més que no ho siguin.[12] Inversament, des de l'any 1983, el lexicògraf britànic John Sinclair es pronuncia per l'ensenyament de la lexicografia a la Universitat i proposa un programa en el qual la lingüística, general i computacional, tingui un lloc rellevant.[13] L'any 1986 va aparèixer un volum dedicat a la formació dels lexicògrafs a occident,[14] en un article del qual Alain Rey,[15] després de relatar fets d'experiència semblant als mencionats per Landau,[16] proposa un disseny molt obert de lexicògraf ideal, que concentra a la vegada competències lingüístiques, amplis coneixements enciclopèdics, especialment en ciències humanes, i sabers pràctics en el domini dels oficis del llibre.[17] Però, en aquest text, Rey, pel que fa a l'aptitud de la Universitat per formar lexicògrafs, mostra reserves que mereixen ser considerades: la Universitat –diu– seria més apta per formar investigadors en metalexicografia que no pas per formar pràctics de la lexicografia.[18] És cert que la Universitat literària, que és tradicionalment el lloc del comentari del discurs segon on es glossa i s'analitza les produccions discursives existents, no té per vocació produir discursos primers. Conseqüentment, Rey postula una formació de lexicògrafs que es faria dintre d'equips lexicogràfics ja constituïts, i subratlla immediatament el caràcter idealista d'aquesta proposició ja que cap equip privat o públic no sembla pas disposat a invertir ni energies ni diners en un projecte d'aquesta mena.[19] Quan nosaltres vam introduir el DEL a Lille III, vam prendre'ns molt seriosament l'opinió de Rey. Pensem, però, que és possible de fer-hi dues rèpliques, una de pràctica i una altra de teòrica. En la concepció del DEL, ens hem esforçat per modificar les maneres de procedir i d'orientar el nostre ensenyament i el control dels estudiants, no pas, estrictament parlant, cap a l'autèntica pràctica lexicogràfica, ja que no podem reproduir les condicions reals d'una editorial de diccionaris, però sí cap a una simulació de diferents fases del treball dels lexicògrafs: documentació lèxica, estructuració de nomenclatures, redacció d'articles. Després de dos anys d'experiència, podem dir sense exagerar que som capaços de tenir sobre el lèxic un discurs alternatiu al discurs tradicional de la Universitat, i assumir el risc de l'aplicació. Així doncs, i desenvolupo la part teòrica de la meva resposta sobre les reserves de Rey, no concebem la metalexicografia universitària com una activitat separada de la lexicografia, sinó com una altra manera de fer lexicografia no sotmesa als condicionaments temporals i financers, que dóna al lexicògraf una imatge dels seus productes que el treball diari

(11) 2a ed. Cambridge/Nova York/Port Chester/Melbourne/Sydney. Cambridge University Press, 1989.

(12) Pàg. 236.

(13) «Lexicography as an academic subject». A: R. R. K. Hartmann (ed.) *LEXeter'83 Procedings. Papers from the International Conference on Lexicography at Exeter*. 9-12 September 1983. Lexicographica Series Maior 1, Tübingen, Max Niemeyer Verlag, 1984, pàgs. 3-12.

(14) Robert Ilson (ed.) *Lexicography. An emerging international profession.* The Fulbright Papers, vol. 1. Manchester/Dover (New Hampshire), Manchester University Press.

(15) «Training lexicographers: some problems», pàgs. 93-100.

(16) Pàg. 95.

(17) Pàgs. 97-99.

(18) Pàg. 99.

(19) Pàgs. 99.

no els permet donar, i que en el millor dels casos pot entrar en interacció dialèctica amb la producció de diccionaris. D'altra banda, és necessari entendre que la metalexicografia és una activitat a la qual es lliura, empíricament, tot lexicògraf quan utilitza altres diccionaris per elaborar-ne un de nou, o quan treballa en la revisió d'un diccionari: en tots dos casos, sotmet a la crítica, per esmenar-los, els textos lexicogràfics existents. I així, sempre hi ha, encara que sigui inconscientment, sense teorització ni sistematització, un component metalexicogràfic en l'activitat lexicogràfica, almenys quan aquesta s'inscriu en una tradició. Per tant, el comentari de diccionaris no constitueix en si una formació per a l'ofici de lexicògraf, ara bé no és tampoc una activitat absolutament separada ni antagònica de la pràctica lexicogràfica.

**P. És interessant aquest punt de vista... La metalexicografia concebuda com una activitat pròpiament lexicogràfica...**
**Per acabar aquest tema, quina relació creieu que hi hauria d'haver entre els centres de formació i els centres d'ocupació? Bé, en una situació òptima.**

DC. Si les universitats fossin reconegudes com a centres de formació de lexicògrafs, seria necessari que les editorials, quan recrutessin personal, agafessin els millors estudiants formats a les universitats. A més, la complementarietat entre la formació i l'ocupació hauria de poder materialitzar-se en la intervenció a la Universitat de professionals de l'empresa, que poguessin transmetre als estudiants la seva experiència com a pràctics.

**P. Això no es fa?**

DC. No, més aviat poc. Els lexicògrafs amb una experiència professional rica i, alhora, amb una perspectiva general suficient sobre la seva pràctica no són gaire freqüents, i estan massa ocupats per poder intervenir tantes vegades com caldria en una formació universitària.

PC. Actualment, a la pràctica, la relació que s'estableix entre la Universitat i els centres d'activitat lexicogràfica es concreta essencialment en les estades que els estudiants del DEL fan en alguns d'aquests centres. Aquestes estades són indispensables, perquè la Universitat només pot, en el millor dels casos, simular la pràctica lexicogràfica. Però no pot donar als aprenents de lexicògraf la competència lexicogràfica, l'experiència de les rutines professionals, ni la dels ritmes de treball (de vegades molt intensos), ni la de les decisions de redacció que s'han de prendre en un temps limitat, a costa sovint d'una certa insatisfacció intel·lectual, que cal assumir de totes a totes.

Passar-te moltes setmanes analitzant un mot perquè és particularment interessant, significatiu o difícil de tractar, és un luxe universitari que no es pot permetre qui està treballant en una activitat lexicogràfica sotmesa a les lleis del mercat. Les estades serveixen per aprendre tot això. Ara bé, la contrapartida a les condicions pràctiques de l'activitat lexicogràfica és que limita en certa manera el desenvolupament intrínsec. Els lexicògrafs professionals poden fer evolucionar els diccionaris servint-se de moltes astúcies tipogràfiques i semiòtiques que coneixen bé; però aquest desenvolupament té uns límits que només es poden depassar amb la col·laboració dels lingüistes.

Pel que fa a les innovacions més de punta que podria conèixer la lexicografia, espero més poc de les editorials que de la recerca lingüística: és per aquest motiu que cal que els futurs lexicògrafs estiguin sòlidament formats a la Universitat.

**P. Bé, passem ara a la segona part del qüestionari, que fa referència a l'activitat lexicogràfica a França. Creieu que els diccionaris que hi ha publicats omplen les necessitats dels usuaris francesos? Quins buits constateu?**

PC. Jo no conec directament les necessitats dels usuaris. Només puc mencionar les paraules dels lexicògrafs recolzades en els estudis de mercat efectuats per projectes lexicogràfics concrets. Segons aquestes opinions, la verificació ortogràfica és un dels factors més importants que fa utilitzar els diccionaris. La consulta del diccionari per qüestions de semàntica, sovint limitada a finalitats concretes lligades a les activitats escolars o a jocs de llenguatge dels mitjans de comunicació, sembla ser mínima. Els més pessimistes diuen que els components menys utilitaris dels articles (etimologies, citacions...) donen més satisfacció als lexicògrafs que no pas profit als usuaris, però que els editors troben arguments publicitaris per incloure aquest tipus de dades en un mercat amb un grau de competició fort: cap diccionari no pot oferir menys informació que els de la competència, independentment de l'ús que els usuaris facin d'aquesta informació. L'exemple de la transcripció de la pronunciació en l'Alfabet Fonètic Internacional és potser un dels punts més significatius: apareix només des de fa trenta anys en el mercat monolingüe francès, concretament en el *Dictionnaire du français contemporain*[20] o també en el *Petit Robert*,[21] i a partir d'aquí sembla que hagi esdevingut un constituent obligatori de tot diccionari que tingui un cert prestigi, encara que és sens dubte evident que es tracta d'una dada poc útil per a la majoria dels usuaris. Com en altres sectors del consum, hi hauria també, particularment sota la influència de la publicitat, un desajustament entre d'una banda allò que els compradors de diccionaris imaginen sobre les garanties de seriositat i cientificitat d'un diccionari, i d'altra banda les seves pràctiques de consulta, que variarien poc i restarien centrades en l'ortografia dels mots o en la recerca del seu sentit des d'una perspectiva més aviat referencial i enciclopèdica que no pas lingüística.

(20) París, Librairie Larousse, 1966.

(21) París, Société du Nouveau Littré, 1967.

**P. Aquesta actitud, reflecteix una crisi lexicogràfica a França? És a dir, significa que la lexicografia francesa està avui dia més preocupada per la comercialització que pels aspectes pròpiament científics?**

PC. Pel funcionament actual de les edicions privades, no veig res que m'inciti a pensar el contrari.

**P. I com expliqueu que darrere cada editorial de diccionaris hi hagi un lingüista de prestigi?**

PC. No pas en general.

**P. Com que no? Penso en Alain Rey als diccionaris Robert o en Jean Dubois a l'editorial Larousse.**

PC. Cert. Però Jean Dubois ja no és a Larousse des de fa anys. I a més, en aquestes editorials la col·laboració és més anònima, i Hachette, després de la seva tornada a la producció de diccionaris monolingües als anys vuitanta, no disposa de lexicògrafs destacats.

**P. Així, doncs, la situació ha canviat.**

PC. La situació comercial ha canviat. La pràctica editorial privada està regida avui dia per la lluita que lliuren els dos grans grups editorials francesos, Hachette i el grup de La Cité, per controlar una part del mercat. Una conseqüència observable és allò que els professionals anomenen «la rendibilització multifuncional de les obres», és a dir la multiplicació d'un diccionari font en un nombre indefinit de productes variants, més o menys identificables com a tals.

**P. Creieu que el nom d'un lexicògraf prestigiós al darrere d'una obra no fa vendre més exemplars?**

Els lexicògrafs que tu has esmentat –Jean Dubois, Alain Rey, als quals podríem afegir Josette Rey-Debove– no són comercialment prestigiosos, com ho són Littré, Larousse o Robert, que donen el seu nom al diccionari que van fer. Els lexicògrafs de què nosaltres parlem són reconeguts en la professió i també a la Universitat, perquè

són alhora metalexicògrafs que reflexionen sobre la seva pràctica i, en alguns casos, són lingüistes. Però no són, malgrat tot, coneguts pel gran públic. Potser les prestacions mediàtiques d'alguns d'ells podrien contribuir a modificar aquesta situació.[22]

DC. Per retornar a l'estat actual de l'edició lexicogràfica francesa, voldria remarcar un punt del qual encara no hem parlat: la feblesa de la nostra producció de diccionaris bilingües, cosa que connecta sens dubte amb el fet que França no és un país molt potent en l'aprenentatge de llengües estrangeres.

**P. Parleu una mica més de les raons d'aquesta situació.**

DC. Potser hagi estat la vella idea de la universalitat de la llengua francesa que ha pesat en el procés de desenvolupament de l'ensenyament de les llengües a França: això no és més que una hipòtesi que s'hauria de verificar. El que és segur és que l'ensenyament de llengües estrangeres a les escoles franceses no és satisfactori. D'un ensenyament lligat només a l'escrit en detriment de la pràctica oral, com es feia tradicionalment, s'ha passat a un ensenyament lligat a l'oral en detriment de la pràctica escrita, sense que potser hi hagi hagut un benefici significatiu per a la pràctica oral, perquè els horaris són massa reduïts i les classes massa nombroses. No sé exactament com aquest fet connecta amb la situació editorial, però podem constatar que els editors francesos produeixen pocs diccionaris bilingües i sovint col·laboren amb editors estrangers.

**P. I no creus que la nova situació europea ha fet canviar aquesta actitud?**

DC. M'agradaria molt que l'hagués fet canviar, però dubto que ho hagi fet.

PC. El que diuen alguns especialistes és que per desenvolupar una lexicografia bilingüe rica cal o bé disposar d'un poder comercial suficient per oferir als diccionaris un mercat ampli –i aquí dominen els països de llengua anglesa–, o bé ser un país on la vitalitat comercial està lligada a l'aprenentatge de llengües –podem pensar aquí en països petits en relació al nombre dels seus parlants natius, com els Països Baixos o Dinamarca, per exemple. França no pertany a cap d'aquests dos grups: el pes internacional del francès, que no és el de l'anglès, no permet desenvolupar una lexicografia bilingüe d'una importància similar i, almenys fins fa poc, l'economia de França no ha hagut de recolzar-se essencialment en el plurilingüisme dels seus habitants.

(22) Els nous diccionaris ara poden fer la funció d'una promoció mediàtica comparable a la que acompanya l'aparició d'altres funcions culturals (novel·les, pel·lícul·les, discos), que es projecte sobre l'escena de determinats lexicogràfs, per exemple Alain Rey i Josette Rey-Debove. Per altra banda, des de la tardor de 1993, Alain Rey fa, de dilluns a divendres, a les set menys cinc, una crònica etimològica a la ràdio France-Inter.

**P. Passem ara al Diploma Europeu en Lexicografia. Com va néixer la idea de posar en marxa aquest diploma?**

DC. Els nostres records sobre aquest punt són una mica vagues. La idea devia néixer en el curs de les converses que vam mantenir el 1987 amb Franz Josef Hausmann i el 1988 amb Reinhart Hartmann. Pensàvem des de feia temps en una formació francesa d'alt nivell en lexicografia, però mai no l'havíem concretat, i va ser Hartmann qui va donar l'empenta decisiva per a la creació del DEL.

**P. Quins tipus d'estudiants rebeu? De quins orígens? Amb quina formació?**

DC. Hi ha tres tipus d'estudiants. Aquells que busquen en el DEL una formació complementària en lexicologia (per exemple, perquè preparen una tesi en aquest domini) i que no volen pas orientar-se professionalment cap a la lexicografia. Hi ha també estudiants francesos, relativament poc nombrosos, que voldrien esdevenir lexicògrafs. En definitiva, els més nombrosos són els estudiants estrangers, no només europeus, els quals pensen trobar en el DEL una formació útil per al seu futur professional al seu país, i que tenen sovint projectes lexicogràfics concrets pel que fa a la seva llengua materna.

**P. De quins països són els estudiants estrangers que segueixen el diploma?**

DC. D'una banda, dels països francòfons (Magreb, Àfrica Negra, Quebec), i de l'altra, dels països de l'Europa comunitària (Alemanya, Grècia) o oriental (Rússia, Polònia).

**P. Quants anys passen a Lille els estudiants estrangers?**

PC. Això depèn: alguns s'hi passen dos anys, d'altres, només un any.

**P. Com està organitzat el diploma?**

DC. Els estudis es distribueixen en dos anys, però l'organització dels cursos permet que els estudiants segueixin tots els cursos en un sol any i que convalidin el segon curs d'una manera diferida, cosa que és particularment còmoda per als estudiants estrangers que no poden finançar-se una estada de dos anys a França.

**P. Així, doncs, cada any comença una nova edició?**

DC. Sí. Cada any consta actualment de cinc mòduls, ja que la durada de l'ensenyament varia en funció dels continguts. El conjunt dels dos anys correspon a unes dues-centes hores de curs, amb alguns ajustaments específics cada any.

PC. El primer any consta d'un mòdul d'iniciació a la història dels diccionaris francesos i al seu mercat actual, que pretén donar als estudiants informació sobre les fonts documentals. Tres altres mòduls estan orientats a la matèria pròpiament lexicogràfica: el primer sobre l'estructura del diccionari, el segon sobre l'estructuració del significat, i el tercer sobre la variació lèxica. Un darrer mòdul proposa una iniciació a la lexicografia computacional. El segon any, aquesta iniciació s'aprofundeix en un primer mòdul específic de lexicografia computacional, un segon mòdul sobre la tipologia dels diccionaris, i un tercer mòdul, anomenat «lexicografia i lingüística», que es refereix a les classes lexicals, que suposen per als especialistes el nucli de les qüestions més problemàtiques: auxiliars, determinants, pronoms, preposicions, conjuncions, connectors, morfemes no autònoms. En aquest mòdul, el programa es distribueix al llarg d'un cicle de dos anys amb la finalitat que els mateixos estudiants puguin seguir-lo dos anys seguits abordant qüestions diferents. Els dos darrers mòduls són iniciacions a la terminologia i a la lexicografia bilingüe.

**P. Podeu cobrir la docència del diploma amb els professors de la Universitat de Lille?**

DC. Actualment, els professors de Lille III es fan càrrec de sis mòduls sobre deu; però per als mòduls de lexicografia computacional, de terminologia i de lexicografia bilingüe, hem recorregut a professors estrangers, provinents d'altres països de la Comunitat Europea (Alemanya, Bèlgica, Espanya, Gran Bretanya): així, el diploma és europeu també pel que fa al reclutament dels professors. D'altra banda, la majoria d'aquests col·laboradors són membres del grup de recerca SILEX.

**P. Què és la unitat SILEX?**

DC. És un equip de recerca associat al CNRS i localitzat a la Universitat de Lille III, que s'ocupa del lèxic. Sens dubte, és l'únic equip francès reconegut oficialment que estudia aquest domini. Agrupa professors de lingüística francesa i anglesa de la Universitat de Lille III i investigadors que pertanyen a altres universitats franceses i europees.

**P. Què significa per a un professor ser membre de SILEX?**

DC. Significa que inscriu els seus treballs de recerca dins del programa de SILEX, que comporta tres eixos. El primer eix, anomenat «Morfologia derivacional», tracta de la formació dels mots. Té com a objectius l'elaboració conjunta d'una gramàtica i d'un diccionari de mots derivatius, aquest treball és en curs de realització pel que fa al francès, i s'està començant a fer, dins del mateix marc teòric, per al portuguès i el grec. En el segon eix, que s'anomena «Semàntica lèxica, semàntica gramatical i referència», actualment s'està investigant sobre la sintaxi i la semàntica de les preposicions. El tercer eix, que s'anomena «Metalexicografia i lexicografia teòrica», articula les dues orientacions anteriors amb els projectes de diccionaris.

**P. I no hi ha formació pràctica al diploma?**

DC. Sí, com ja hem dit la formació comporta una estada de dos mesos com a mínim en un centre d'activitat lexicogràfica, públic o privat. A més, els estudiants han de presentar una memòria de recerca sobre lexicologia i/o lexicografia.

**P. Qui busca el lloc de pràctiques de l'estudiant?**

DC. És l'estudiant mateix qui ha de trobar el seu centre de pràctiques, la Universitat s'encarrega només dels aspectes administratius i jurídics.

**P. I són ben acceptats?**

PC. En general, sí.

**P. Feu un seguiment de les pràctiques?**

DC. Als nostres estudiants no els demanem que presentin cap memòria de pràctiques; però, en canvi, els fem fer un informe oral no formal, i demanem als centres que els han acollit que els facin un certificat de pràctiques i que ens donin una qualificació general de cada estudiant.

**P. Què ha de fer un estudiant estranger si vol seguir el diploma?**

DC. Cal que ens escrigui personalment abans del mes de maig, abans de la represa del curs, demanant ser admès i adjuntant a la carta un *curriculum vitae* on s'indiqui els estudis universitaris que ha realitzat, l'experiència que té en lingüística i en lexicografia, les seves activitats professionals, etc., cosa que ens permet avaluar millor la seva aptitud per seguir la formació. Pel que fa a la inscripció administrativa, ha d'adreçar-se al responsable del Departament de tercer cicle del Servei d'escolaritat de Lille III. La inscripció al DEL requereix les mateixes condicions que l'acceptació al DEA: cal d'entrada una llicenciatura o un equivalent, i la inscripció es fa després de la revisió del dossier de cada candidat.

**P. Es pot seguir una primera part del diploma en lexicografia a Lille i la segona a una altra Universitat europea que també ofereixi el diploma (Exeter o Amsterdam)?**

DC. En un principi, sí. Però això encara no ha passat, i sembla bastant difícil que passi a la pràctica, sobretot per qüestions de llengua.

**P. Cal que els estudiants procedeixin de les universitats que formen part de l'acord del diploma?**

DC. Si es volen beneficiar d'una beca Erasmus, obligatòriament. Si no, l'accés és obert a tots els estudiants estrangers sense cap distinció. Per a l'any 1993-1994, el finançament europeu s'ha suspès perquè no hi hagut canvis durant els dos últims anys. Ara bé, hem demanat un nou ajut a la Comunitat Europea, allargant el consorci inicial (França, Gran Bretanya, Països Baixos, Bèlgica) a d'altres països (Dinamarca, Espanya, Grècia), i diferenciant dos tipus de centres: els centres principals, que són capaços d'oferir tota la formació, entre els quals hi haurà Barcelona; i els centres associats que només poden oferir una part de la formació. Hem previst que el finançament Erasmus serveixi principalment per als estudiants provinents dels centres associats.

M. TERESA CABRÉ CASTELLVÍ
*Lille, febrer 1993*

*Lexicografia i models lingüístics: les teories lingüístiques i el lèxic*
Lexicography and Linguistic models: Lingustic Theories and Lexicon — Toni Badia

In this paper a discussion is offered of the relation between linguistic theories and lexica. It is shown that linguistic theories imply particular approaches to the lexicon from which particular ways of conceiving the lexicographer's work can be inferred. To this end a few linguistic theoretical models are examined and their idea of the lexicon is reviewed. Obviously it is beyond the scope of this paper to discuss the approach to lexicon of all different linguistic theories developed during this century. Consequently a few models have been chosen for examination which nonetheless are quite representative of the main phases of the recent history of linguistics. After a very short discussion of the characteristics of traditional grammar, the paper starts the discussion with structuralism and then moves over to the examination of some of the linguistic theories that can be broadly considered «generative»; three basic stages in the development of Chomsky's approach are reviewed, i.e., the ones represented by *Syntactic structures, Aspects of the theory of syntax,* and *Lectures on government and binding*; and finally the discussion ends with an examination of LFG (*Lexical-functional grammar*) and HPSG (*Head-driven phrase structure grammar*) which are two of the emerging theories are that have originated in the linguistic millieus most influenced by computational lingustics.

*Lexicografia i variació funcional*
Lexicography and Functional Variation — Lluís Payrató

This paper reviews the basic concepts related to notions of *variation* and *functional variation*, in particular *register*, and it questions whether these concepts have been taken into account by the lexicographic tradition. The paper focuses on which are the most suitable marks to recognize several aspects connected with the usage of lexical items: on the one hand the field (topic), the mode (channel), the tenor (purpose), and the tone (formality), as the register's components; and on the other hand, various factors concerning cultural, enunciative, and textual grounds.

*Lexicografia i cultura: antropologia i sistema social (noms, renoms i noms de casa al Pirineu aragonès)*
Lexicography and Culture: Anthropology and Social System (names, individual nicknames and household nicknames in the Aragonese Pyrenees) — Joan J. Pujadas i Dolors Comas

In this article we analize the system of names, individual nicknames and household nicknames. The empirical information was obtained from ethnographic fieldwork in two villages of the central Pyrenees. Anthroponims serve to denominate individuals; they also classify them according to their place in the family and in the community. Finally, anthroponims evoke meanings and symbols culturally constructed.

*Lexicografia i intel·ligència artificial*
Lexicography and Artificial Intelligence — Ramon Cerdà Massó

Unlike most of traditional approaches to grammar, today's formal and computational models make use of the lexicon as an essential part of the whole system. Differences between current theoretical and computational models are more and more insignificant albeit the latter are perhaps still characterized by yielding full priority to minimalist solutions over more abstract considerations such as exhaustivity or formal elegance. The paper introduces some strategies in linguistic computation for natural language processing (NLP) according to a typology based on functional complexity and describing in every case the scope and the role performed by the dictionary: pattern matching, semantic grammars, syntactic parsers, augmented transition networks, unification formalisms, case frame grammars, etc. It ends with an exploration into lexicographically oriented procedures of conceptual dependency, which are found beyond the NLP, deeply within the domain of artificial intelligence.

*L'ús dels diccionaris escolars*
The use of L1 Learning Dictionaries — M. Paz Battaner

From a study of four cases of real use of school dictionaires a review is made of issues habitually debated in general lexicography which are present particularly in the lexicography of first-language acquisition.

Proposals are then made for the use of school dictionaries as a series of learning goals based on parallel reflection on the lexical unit, and an awareness of the assumptions involved in lexicographic conventions.

*La professió de lexicògraf*
Lexicography as a Profession — Emili Pascual

At the present time, there is no doubt that lexicography takes up an important place in the field of language science and that this fact involves the possibility of speaking about the scientific task carried out by properly trained professionals: the lexicographers.

This article reflects on the mentioned subject, paying special attention to the characteristics of the lexicographers' work.

The first section, From the Amateur to the Professional Lexicography, explains the fact that nowadays lexicography does not remain in the hands of the amateur and the enthusiast, but in those of the professional scientist'. Lexicography has changed from «the art of making dictionaries» to «the science and the technique of dictionaries».

The second section, The Establishment of the Scientific Lexicography, comments on the importance of the Exeter International Convention of Lexicography (1983), the creation of EURALEX and the union of efforts and research works with the aim of determining the lexicographical science.

The third section focuses explicitly on the several points of the lexicographer's work. In this central part of the article, lexicography is placed within the limits of applied linguistics and, therefore, they insist on the fact that the lexicographer is not basically a language theorist. The basic assignments in the treatment of the subject matters are commented, and the lexicographer's work is related to the rest of the branches of language science and the need to attend to questions that were obviated from dictionaries not long ago: treatment of macro and micro-structure, problems of definition, indications for the use of verbs, noun position and combinatory, etc.

Finally, the importance of the lexicographer in the establishment of language use and the need of collaboration with other specialized professionals is commented, as well as the computer services available for those who take up dictionary making.

*La lexicografia catalana actual*
Current Catalan Lexicography — M. Teresa Cabré Castellví

In this paper a general analysis of the evolution and current state of modern Catalan lexicography is undertaken, which starts with the setting up of a catalogue of lexicographic publications from 1940 onwards. In addition to the quantitative analysis of published dictionaries the paper introduces a typology study in order to sort out the variety and richness of modern Catalan lexicography.

In the paper relevant data is presented on the situation of Catalan lexicography with respect to the quantity of existing dictionaries, and also on the factors that have determined their variety. Finally it can serve to establish the bases to build a lexicographic plan that helps to fill the gaps and avoids the unnecessary repetitions, so that Catalan can be given the same lexicographic resorts that are available for other widespread languages.

*Un tast dels vocabularis bilingües valencians actuals*
A Choice of Present Valencian Bilingual Vocabularies — Jordi Colomina

By means of a choice of seventy words, this work shows the lexical richness and the degree of acceptance of the guidelines –defined in the Diccionari General de Llengua Catalana (General Dictionary of the Catalan Language) by Pompeu Fabra– regulating nine bilingual Valencian/Castilian and Castilian/Valencian vocabularies published between 1972 and 1993.

*La lexicografia catalana moderna a les Illes Balears*
The Modern Catalan Lexicography in the Balearic Islands — Jaume Corbera

The works on lexicography published in the Balearic Islands after the completion of the *Diccionari Català-Valencià-Balear* (1962) are described and analysed in this paper. These works are already fully inserted in the movement for the linguistic recovery undertaken the last twenty years. Especially, but not exclusively, the paper analyses the reflection on those works of the particularity of the Balearic dialect with regard to the standard language; that is to say, the degree of vindication of the specific forms and words of the insular Catalan tacitly or explicitly accepted in the standard frame.

*La lexicografia catalana a l'Alguer*
Catalan Lexicography in L'Alguer — Jaume Corbera

There are a few *alguerès* works made with lexicographic methods or pretensions, apart from some contributions more or less interesting or useful for the knowledge of the local language, intended principally for dialectologists: only two titles produce explicitly the label of «Vocabulary» or «Dictionary» and use the typical structure of this kind of works. We can add to these references an unpublished work by G. Palomba.

*El nou diccionari normatiu*
The New Normative Dictionary of Catalan — M. Teresa Cabré, Maria Bozzo, Àngels Egea, Marta Juncadella, Carolina Santamaria

This article describes the process by which the Nou Diccionari Normatiu de la Llengua Catalana (NDNLC) has been made; this dictionary brings up to date and regularizes Pompeu Fabra's Diccionari General de la Llengua Catalana (DGLC), which implies, on the one hand, adding to the dictionary new words brought about by different social habits and by the progress of science and technology, and, on the other hand, bringing the contents of the dictionary up to date.

This project has been carried out in three different phases. The first consisted in creating a general computer data base (BDOL) which contains, besides the DGLC, words belonging both to the common language and to specific fields; in this first phase a study of the DGLC was also carried out, with two aims: a) to analyse the text itself and in doing so select whatever information might be considered obsolete today, and b) to find out the criteria which Fabra used to make the dictionary, which are implicit in the text.

In the second phase the lexicographic criteria by which the dictionary was to be made were established, its entries (both words and phrases) were selected, the old articles were revised and new ones were created.

Finally, in the third phase the data base specifically for the NDNLC was created; this contains all the information of the Nou Diccionari Normatiu de la Llengua Catalana, organised into articles.

*Un corpus general de referència de la llengua catalana*
A General Corpus of Reference of the Catalan Language — Joaquim Rafel

This work presents the various aspects of the characterization of the Corpus Textual Informatitzat de la Llengua Catalana (Computerized Textual Corpus of the Catalan Language), and analyzes the present state of this project, linked to the preparation of the so-called *Diccionari del Català contemporani* (Dictionary of Contemporary Catalan) undertaken by the Institut d'Estudis Catalans (Institute of Catalan Studies).

*La terminologia catalana actual*
Current Catalan Terminology — Carles Duarte

The Catalan language currently has a series of terminological resources at its disposal which award it a privileged position compared to other languages of a similar size. This situation also allows for a fitting response to new terminological requeriments arising as a result of technological delevopment. Worthy of mention in this respect is the existence of TERMCAT, an organisation which has been a decisive factor in the promotion of terminological work. This has come about by means of the adoption of modern methodology adapted to international standards, the training of experts in terminology and the creation and publication of dictionaries. Mention should also be made of the excellent work carried out in this field by the universities and governments in Catalonia, the Community of Valencia and Andorra.

*La lexicografia bilingüe*
Bilingual Lexicography — Jordi Bruguera

A monolingual dictionary offers word definitions. A bilingual dictionary provides equivalences between a source language and a target one. It relates the words in the source language to some sort of definition by synonymy, with a different phonetics.

Can there be wholly satisfactory equivalences between the linguistic units in one language and those in another one? Does this question relate in any sense to the notion of linguistic universal? With respect to denotation, anything that can be expressed in one language can undoubtedly be said in another. The original cultural homogeneity between many languages (for example, the Indoeuropean ones) makes easy for them to syntonize with one another with respect to the universals.

The criteria that are used to choose and provide a hierarchy of lemmas with their different meanings in a monolingual dictionary can be easily applied to bilingual dictionaries. The number of entries usually depends on

the greater lexical offer of the source language because of the interest that can raise in the transfer to the target language. However, should any «speech» fact be reflected in a bilingual dictionary?

Users can expect a good bilingual dictionary to offer a large amount of vocabulary, which is usual and genuine in a variety of registers, with a detailed semantic description of every word and a collection of the idiomatisms in which the word is involved. It is also expected that the dictionary user can understand and expression in a particular language and express it in another one without altering their meaning content.

*La lexicografia general i la normativa*
General Lexicography and Normative Language — Joan Solà

As a starting point we take the real needs of the present language users, specially those professionals of linguistic services, in order to get a clear picture of which of their needs could be satisfied by a general language dictionary and which by a normative dictionary. The vagueness of the concept «normativity» is pointed out and the opinion is expressed that it is not easy to define in a precise way which of the needs mentioned ought to be the subject of normativity.

The article concludes with a series of concrete examples which illustrate the concepts presented.

*Lexicografia i correcció: La funció dels diccionaris en la revisió de textos*
Lexicography and Correction: The Role of Dictionaries in the Text Supervising Processes — Mercè Lorente i Carles Tebé

The aim of this paper is an attempt to revise the role of dictionaries in professional text supervising processes. Under the frame of text linguistics, we first explain how correction tasks are conditioned by text typology. Next, we study the macro structure of dictionaries (in order to consider what kind of lexicographical works are closer to the corrector needs) and then we study their micro structure (in order to set up what kind of information might be relevant for correction tasks). In conclusion, we finally state some lacks of current dictionaries for correction purposes, and we settle some aspects of text revision which are not expected to be solved with dictionary consultation.

*Una tipologia de la definició lexicogràfica aplicada al DGLC*

A Typology of Lexicographic Definition Applied to DGLC

Victòria Alsina, Lluís Sol i Joan Soler

This article aims at establishing a typology of lexicographic definitions by means of the study of the definitions found in Pompeu Fabra's Diccionari General de la Llengua Catalana in nouns, verbs and adjectives.

Several typologies based on different criteria are arrived at: in the first place, the difference between synthetic and periphrastic definitions is established; this classification is based on the amount of words used in the definition. In the second place, the typology which comprises substantial and relational definitions is explained; it was established by Rey-Debove (1967) and it is based on the metalanguage used in the definition. And thirdly, a classification within substantial periphrastic definitions is established, based on the defining words used: lexical and semo-componential definitions.