

# La pluriaccessibilitat com a eix de la lexicografia hipertextual

Carme Bach, Anna Matamala

Grup IULATERM

Tot i que els diccionaris són una font d'informació valuosa, de vegades l'usuari queda insatisfet perquè no hi troba allò que busca. Pot ser que l'usuari no busqui en el diccionari adient, que l'absència sigui fruit d'un error del diccionari o bé que la informació buscada hi sigui, però costi de trobar, i que fins i tot sigui inaccessible. Un exemple clar de les dificultats d'accés a la informació lexicogràfica el configuren totes les peces lèxiques que són difícils de definir mitjançant la metallengua del contingut, bona part de les quals tenen una forta càrrega pragmàtica.

En aquest article defensem que l'explotació màxima de les possibilitats que ofereix l'hipertext facilitaria l'accés a la informació que cerca l'usuari i li permetria utilitzar de forma més satisfactòria les obres lexicogràfiques. Per a aquest treball partim del supòsit que els entorns electrònics són els que permeten explotar al màxim la hipertextualitat. Per demostrar-ho, ens proposem d'analitzar les maneres d'accedir a la informació que ofereixen diverses obres electròniques per a dues unitats amb un fort component pragmàtic: els connectors i les interjeccions. Vistes les mancances d'accés a la informació que presenten, i tenint en compte les possibilitats de l'hipertext, proposem dues aplicacions que aprofiten al màxim la hipertextualitat i que han estat desenvolupades amb la intenció de millorar l'accés a la informació lexicogràfica.

L'article s'estructura en tres parts: en la primera, reflexionarem a l'entorn del concepte d'hipertext i ens centrarem en la seva implementació lexicogràfica. En la segona, analitzarem la hipertextualitat de tres eines electròniques, prenent com a unitats d'anàlisi els connectors i les interjeccions. Ja per acabar, presentarem dues aplicacions pluriaccessibles ideades amb la voluntat d'explotar al màxim l'hipertext.

## 1. L'HIPERTEXT

### 1.1 Concepte d'hipertext

Quan parlem d'hipertext, sovint fem referència a dos conceptes propers però diferents:

- a) un sistema d'organitzar el coneixement (Cabré *et al.*, 1996)
- b) una manera d'estructurar la informació.

En aquest treball ens basem en aquest segon sentit i considerem que l'hipertext és “una estructura de la informació que organitza les entitats en forma de xarxa” (Codina, 1998). Aquesta no-linealitat de la informació és un element clau que converteix l'usuari en un personatge actiu que va navegant pels diversos blocs d'informació i va construint el seu propi camí. Podem trobar elements hipertextuals en la majoria d'obres lexicogràfiques, sigui quin sigui el seu format: coincidim, doncs, amb Gelpí (2001, 165) quan afirma que “hipertextual

no és sinònim d'electrònic”, però considerem que és en entorns electrònics on es pot treure el màxim rendiment de la hipertextualitat.<sup>1</sup>

## 1.2 Lexicografia i hipertext

Tot i que l'hipertext està estretament lligat a entorns electrònics i que les obres impreses solen tenir una lectura lineal, els diccionaris en paper es podrien considerar mostres d'hipertext, ja que empenyen l'usuari a tenir un paper actiu i a construir el seu propi trajecte, passant d'un node a un altre sense fer-ne una lectura seqüencial. De fet, segons Pruvost (2000), tot el diccionari es podria considerar un gran hipertext.

Ara bé, creiem que hi ha determinats elements on la hipertextualitat encara és més explícita. Per exemple, els índexs—que obren les portes a diferents components del diccionari—, les remissions —que es fan servir per definir algunes unitats i que permeten passar d'un article lexicogràfic a un altre— o determinades etiquetes —com les que indiquen la conjugació del verb i que remetent a uns quadres situats generalment a la part final—. Només aquests tres exemples ja ens demostren que els diccionaris en paper són mostres d'hipertext i que la vinculació es pot produir a nivells diferents:

“Using cross-reference is a method to connect and interrelate different knowledge elements or items included in a dictionary. [...] The relation can be established between two positions inside the same dictionary article, inside the same component part of the dictionary, in different component parts or even between a position inside an another outside the dictionary.” (Tarp 1999: 116)

Aquesta interconnexió dels elements permet que l'usuari vagi construint el seu propi camí de cerca i esdevingui un usuari actiu. Recuperant la terminologia de Fuentes (1995) utilitzada en referència a les vies d'accés a la informació, podríem considerar que l'usuari parteix d'un punt (nus inicial) amb la finalitat d'arribar a la informació buscada (nus final), tot passant opcionalment per elements intermedis (nusos relacionals).

En conclusió, la lexicografia en paper presenta dos elements clau de l'hipertext: la no-seqüencialitat de la informació i la construcció per part de l'usuari del propi camí. Tot i així, és innegable que la lexicografia electrònica és la que permet explotar al màxim l'hipertext, i és per això que ens centrarem en l'anàlisi dels diccionaris en suport electrònic.

## 2. ANÀLISI DE DICCIONARIS EN SUPORT ELECTRÒNIC

Per il·lustrar les possibilitats que obre l'hipertext, analitzarem tres eines electròniques: l'*Hiperdiccionari* (Hiper), que inclou el *Diccionari de la llengua catalana* d'Enciclopèdia —DLC3, la versió en CD-ROM del *Diccionari de la llengua catalana* de l'IEC (DIEC) i la versió en línia d'aquest diccionari (PDL). Ens centrarem en el tractament que aquestes eines

---

<sup>1</sup> Aquestes tres idees fonamentals que acabem d'exposar —la no-linealitat, el lector com a personatge actiu i les possibilitats del format digital— ja les esmenta Bush el 1945 i les recullen diversos teòrics posteriors, com ara Nelson (1992): “per hipertext entenc una escriptura no seqüencial, un text que es divideix i permet al lector escollir; quelcom que es gaudeix al màxim davant d'una pantalla interactiva.”

fan de dues unitats que es defineixen mitjançant la metallengua del contingut, els connectors i les interjeccions.

Per tal de valorar el tractament que aquestes eines fan dels connectors i de les interjeccions, partim de la delimitació d'aquestes unitats establerta a Bach (2001) i a Matamala (2001), respectivament. En aquests treballs s'estudia una llista d'unitats, prèviament seleccionades de corpus, que configuren el conjunt de peces lèxiques que s'utilitzen per a l'avaluació del tractament dels connectors i les interjeccions en les obres lexicogràfiques electròniques que configuren el corpus d'anàlisi d'aquest article.<sup>2</sup>

Ens fixarem en dues formes de navegació que permeten les eines electròniques hipertextuals: (i) la navegació dirigida i (ii) la navegació lliure, ja que considerem que aquests són els dos elements clau dels grans productes hipertextuals.

## 2.1 Navegació dirigida: consultes

Parlem de *navegació dirigida* quan les eines analitzades guien l'usuari tot indicant-li els trajectes que ha de seguir per obtenir de forma directa la informació que cerca. En les eines que analitzem, les consultes són el pas imprescindible per iniciar la navegació.

En els subapartats següents ens fixarem en tres maneres diferents d'arribar als connectors i a les interjeccions a partir de les consultes: (i) mitjançant el lema mateix; (ii) delimitant la cerca per categoria gramatical, i (iii) per mitjà d'elements continguts en les definicions.<sup>3</sup>

### 2.1.1 Accés mitjançant el lema

Quan es vol buscar una peça que ja es coneix, la via més fàcil és introduir directament aquesta unitat. Dos problemes poden impedir o dificultar la consulta: (a) l'absència de les unitats del diccionari i (b) la dificultat d'accedir-hi. Són aquests els aspectes que destacarem a continuació.

#### *a) Absència o presència dels connectors i de les interjeccions*

Pel que fa a l'Hiper, s'hi recullen onze connectors dels catorze que es descriuen a Bach (2001) i quaranta-dues interjeccions de les seixanta-dues unitats del corpus 1. Quant al DIEC, tant en la versió en CD-ROM com en la versió en línia, s'hi inclouen sis connectors i quaranta

---

<sup>2</sup>En el cas dels connectors, s'analitzen catorze peces lèxiques vehiculadores de la reformulació discursiva: cinc connectors de reformulació parafràstica —*és a dir, o sigui, més ben dit, en altres paraules i dit d'una altra manera*— i nou unitats de reformulació no parafràstica —*de fet, de tota manera, en tot cas, en realitat, en definitiva, en qualsevol cas, en el fons, al cap i a la fi i en fi*. Per a les interjeccions, es prenen com a referència dos corpus: d'una banda, el corpus constituït per les unitats incloses al *Diccionari de freqüències* com a interjeccions que tenen una freqüència absoluta superior a cinquanta (62 interjeccions) [corpus 1]. De l'altra, un corpus lexicogràfic més ampli format per 164 articles del DIEC i 166 de l'Hiperdiccionari, a partir dels criteris definits a Matamala (2001) [corpus 2].

<sup>3</sup>Aquestes no són les úniques vies d'accedir als lemes ni els únics filtres que es poden establir, però considerem que són els més significatius per a les unitats que ens ocupen.

interjeccions, tot i que no totes porten la categoria gramatical d'interjecció. El quadre següent resumeix les dades obtingudes:<sup>4</sup>

|                           | Hiper      | DIEC       |
|---------------------------|------------|------------|
| Connectors: 14 unitats    | 11 (78,5%) | 6 (42,8%)  |
| Interjeccions: 62 unitats | 42 (67,7%) | 40 (64,5%) |

*Quadre 1. Unitats incloses als diccionaris*

*b) Dificultat d'accés a les unitats recollides pels diccionaris*

Deixant de banda l'absència de determinades unitats dels diccionaris, volem fixar-nos en la facilitat d'accedir als lemes inclosos. En el cas de l'Hiper, totes les unitats, tant connectors com interjeccions, es recuperen fàcilment, excepte el connector “o sigui”<sup>5</sup> i les interjeccions que el diccionari lematitza amb signes de puntuació.<sup>6</sup> Només cal introduir la peça que es vol buscar (tant si és monolexemàtica com si és polilexemàtica) i es visualitza directament l'article corresponent, independentment que sigui entrada o subentrada. Això representa un avenç important respecte de la lexicografia en paper, ja que no cal que l'usuari sàpiga sota quin lema s'han inclòs determinades unitats. Així, s'accedeix directament a interjeccions com “vinga” o “vejam” —que en la lexicografia en paper s'havien de buscar a l'article lexicogràfic de “venir” i “veure”— o a connectors polilexemàtics com “al cap i a la fi”, sense necessitat de saber si s'ha inclòs a l'article de “cap” o de “fi”.

Pel que fa al DIEC en la versió en CD-ROM, la via d'accés als lemes és més problemàtica. En tots els casos, cal saber si el lema que es busca és tractat com a entrada o subentrada, ja que són tipus de cerques diferents. Això implica que si es busca la interjecció “vinga” com a entrada, el resultat de la cerca serà nul: en casos com aquest, cal especificar que s'està buscant una subentrada per accedir a l'article. El mateix succeix en relació amb els connectors.

Pel que fa al PDL, el procediment és semblant: l'usuari també ha de definir si està buscant una entrada o bé un element dins la subentrada. Els problemes resultants són els mateixos que en el DIEC en CD-ROM, amb dues particularitats: 1) pel que fa als connectors, a través del portal es poden recuperar totes les variants que presenten diferències de flexió, com és el cas de la unitat “de totes maneres / de tota manera” i, 2) pel que fa a les interjeccions, cal reproduir exactament els signes de puntuació amb què es lematitzen per recuperar-les.

En conclusió, l'Hiper facilita l'accés als lemes posant al mateix nivell entrades i subentrades a l'hora de recuperar unitats, mentre que el DIEC i el PDL exigeixen que l'usuari sàpiga si busca una entrada o una subentrada.

<sup>4</sup> Hi ha quatre unitats (*pam*, *crac*, *clac*, *pum*) que figuren a tots dos diccionaris com a substantius i dues peces (*vinga*, *vejam*) que s'inclouen com a subentrades de verbs sense cap categoria gramatical diferenciada. En tots els casos és evident el seu ús interjectiu.

<sup>5</sup> Aquesta variant no es troba ni com a entrada ni subentrada en cap de les obres analitzades, tot i que sí que es descriu el connector sota la variant menys freqüent —o sia.

<sup>6</sup> Així, cal escriure exactament “voto a ...!” i “ca, barret!” per recuperar aquestes unitats.

|               | Hiper | DIEC  | PDL   |
|---------------|-------|-------|-------|
| Connectors    | 90,9% | 83,3% | 83,3% |
| Interjeccions | 95,1% | 98,7% | 96,9% |

Quadre 2: Grau de cobertura en la cerca per lema respecte als lemes inclosos

### 2.1.2 Accés mitjançant la categoria gramatical

Si bé en el subapartat anterior ens plantejàvem la recuperació dels articles d'uns lemes preestablerts, es pot donar el cas que l'usuari vulgui recuperar lemes que desconeix a partir de trets associats als lemes. Un d'aquests trets pot ser la categoria gramatical.

Pel que fa a l'Hiper, aquest tipus de cerca és possible, tot i que els resultats no són satisfactoris per les característiques pròpies de les unitats seleccionades. Així, en el cas dels connectors, dels onze seleccionats cinc no porten cap tipus de marca de categoria gramatical i no són recuperables per aquest sistema. Cinc porten la categoria de *loc. adv.*, però, si es busquen totes les unitats que porten aquesta categoria, el resultat són 621 locucions, una quantitat de soroll que fa la consulta inoperativa (0,8% de precisió). Finalment, una unitat es considera "*locució conjuntiva*", però no s'assenyala en forma d'etiqueta de categoria gramatical sinó dins de la definició, per la qual cosa tampoc no és recuperable mitjançant aquest tipus de consulta. En el cas de les interjeccions, sembla que els resultats haurien de ser més satisfactoris, ja que aquestes unitats pertanyen a una de les categories gramaticals clàssiques previstes pels diccionaris. Ara bé, fent una consulta per aquesta categoria, recuperem 98 unitats de les 166 que formen el corpus (57% de cobertura) i sis que no hi figuren perquè no considerem interjeccions (95% de precisió). D'això es dedueix que un percentatge elevat no són recuperables filtrant la cerca per categoria gramatical.

Pel que fa al DIEC, la versió en CD-ROM no permet dur a terme aquesta cerca, mentre que el PDL sí que incorpora aquesta opció, amb els resultats següents: en el cas dels connectors, de les sis unitats incloses, només tres porten marques gramaticals. Dues s'hi inclouen com a locució adverbial i una com a conjunció, però el soroll que generen les cerques restringint per aquestes categories gramaticals és gran (cent en el primer cas, perquè és el màxim que permet recuperar la interfície, i vint-i-cinc en el segon). Quant a les interjeccions, de les 164 unitats del corpus, només 99 es poden recuperar com a interjeccions i 65 porten una altra categoria gramatical. Tot i que es recuperen una quantitat d'interjeccions destacada, hi ha unitats que queden fora d'aquesta cerca. Especialment significatiu és el cas de les interjeccions representatives (és a dir, les onomatopeïques), que solen figurar en ambdós diccionaris com a substantius: "*atxim*", "*bub-bub*", "*bum*", "*plof*" o "*xip-xap*" en són alguns exemples.<sup>7</sup>

El quadre següent resumeix les dades analitzades en termes de cobertura i de precisió:

<sup>7</sup>Això contravé les dades del DFREQ, ja que un 16% de les unitats definides com a onomatopeïes i considerades substantius a tots dos diccionaris estan recollides únicament com a interjeccions al DFREQ i un 37% reben les dues categoria gramaticals (substantius i interjeccions).

|         | Hiper | DIEC | PDL   |
|---------|-------|------|-------|
| Conn.   | 45,4% | 0%   | 0%    |
| Interj. | 57%   | 0%   | 60,3% |

*Grau de cobertura/categoria*

| Hiper | DIEC | PDL  |
|-------|------|------|
| 0,8%  | 0%   | 0%   |
| 95%   | 0%   | 100% |

*Grau de precisió/categoria*

En resum, tot i que l'Hiper i el PDL permeten recuperar les unitats filtrant per categoria gramatical, l'anàlisi demostra que els connectors no reben una etiqueta gramatical única en totes les eines (i, per tant, no es poden recuperar mitjançant aquest sistema) i també demostra que fins i tot les interjeccions reben categories gramaticals diferents que contradiuen les dades obtingudes de corpus.

### 2.1.3 Accés mitjançant la definició

Una altra manera d'accedir als lemes és a partir d'elements continguts en la definició. Aquesta opció està prevista en el cas de l'Hiper i del PDL, però no pas en el DIEC. En aquest apartat estudiarem (a) si les peces seleccionades presenten unes estructures definitòries comunes que facin possible de recuperar-les en bloc, i (b) si es poden recuperar tenint en compte el contingut.

#### a) A partir de l'estructura definitòria

Si totes les unitats d'un mateix tipus continguessin uns mateixos elements a la definició, la recuperació del conjunt en bloc seria molt més fàcil. Però això no sempre és així.

L'Hiper preveu aquest tipus de cerca, però els resultats no són satisfactoris. Així, en el cas dels connectors seleccionats, només dos inclouen definició: “és a dir” i “o sigui”. La definició del primer connector comença amb la fórmula “Expressió...”, fórmula que comparteix amb 521 entrades més del diccionari.<sup>8</sup> Per al segon, la fórmula triada és “locució conjuntiva que denota...” i hem pogut comprovar que 202 entrades inclouen “denota” a la definició. Pel que fa a les interjeccions, la varietat d'estructures definitòries és molt gran: en les unitats del corpus dos trobem 31 fórmules diferents. Els percentatges més elevats —sense tenir en compte les interjeccions onomatopeïques— corresponen a “Interjecció...” (amb un 24,2%) i “Exclamació...” (16,4%). Hem deixat de banda les interjeccions onomatopeïques perquè reben un tractament diferent i solen estar definides majoritàriament amb la fórmula “Onomatopeia...”, tot i que també hi trobem altres fórmules inicials com “Soroll” (*trinc, gloc-gloc*), “Cant” (*quiquiriquic*), “Veu” (*zim-zam*) o “So” (*dring*), per posar alguns exemples.

Pel que fa al PDL, en el cas dels connectors estudiats, cap no presenta una estructura definitòria i, per tant, no són recuperables d'aquesta manera. El panorama de les interjeccions és ben diferent: si bé hi ha diverses estructures (18 de diferents), la gran majoria d'interjeccions del corpus 2 —deixant de banda les onomatopeïques, que reben un tractament diferenciat— tenen una definició que comença per “Expressió” (70,6% de la resta), generalment “expressió usada per a”. De fet, si busquem aquesta mateixa estructura

<sup>8</sup> La fórmula inicial és “expressió amb què hom”. Ara bé, com que l'Hiperdiccionari no permet de fer una cerca més àmplia pel contingut de la definició, només hem pogut obtenir dades per a la fórmula inicial “expressió”. El resultat ens demostra el soroll que genera aquest tipus de cerca és molt gran i, per tant, la precisió de la consulta és molt baixa. El soroll es pot reduir combinant la cerca per estructura definitòria amb un filtre per categoria gramatical, però els resultats continuen sent poc satisfactoris.

definitòria, totes les unitats recuperades són interjeccions. Pel que fa a les onomatopeïques, tot i que majoritàriament estan definides com a “onomatopeia”, n’hi ha algunes que no són recuperables amb aquest criteri ja que recorren a altres fórmules inicials: “Soroll” (*trinc, tictac*), “Veü” (*quiquiriquic*) o “Crit” (*marrameu*) en són alguns exemples.

#### b) A partir del contingut de les definicions

Una altra possibilitat seria accedir als lemes a partir del contingut de les definicions. En el cas d’unitats que desenvolupen funcions pragmàtiques, es podria introduir la funció que compleix la unitat cercada com a element contingut en la definició.

Pel que fa a l’Hiper, es preveu l’opció de buscar elements dins la definició. En el cas dels connectors, hem cercat les funcions que duen a terme els CR del corpus i hem pogut observar que aquest sistema de recuperació té certa productivitat a l’Hiper, però la quantitat de soroll produïda dificulta la cerca. Si busquem “aclarament” recuperem set unitats, i una és connector. Si busquem “conclusió” recuperem 47 lemes, i només dos connectors i, si busquem “equivalència”, recuperem 25 lemes, però només dos són connectors. Per a la resta de funcions no es recupera cap dels connectors que les vehiculen: “argumentació”, “continuació”, “correcció”, “denominació”, “designació”, “distanciament”, “enumeració”, “exemplificació” i “reformulació”.

El mateix problema presenten les interjeccions: si bé es recuperen unitats a partir dels continguts, la quantitat de soroll és gran. Així, si busquem “alegria”, recuperem 57 unitats, vuit de les quals són interjeccions; si cerquem “dubte”, obtenim 59 fitxes, de les quals només quatre corresponen a interjeccions del corpus.

Pel que fa al DIEC, en el CD-ROM no es preveu aquesta opció. En la versió en línia, tot i que hi ha prevista la possibilitat de buscar per elements continguts en la definició, aquesta cerca només es du a terme en les definicions de les entrades, i no pas en les subentrades, i això fa impossible de recuperar els connectors, que estan recollits com a subentrades. Així, si es busca “conclusió” es genera una llista d’unitats lèxiques però no hi figura cap connector. En el cas de les interjeccions, se’n recuperen força perquè la majoria corresponen a entrades, tot i que continua havent-hi força soroll: per a “alegria” es recuperen 35 unitats, de les quals 7 són interjeccions i, per a “dubte”, els resultats són 35 fitxes contra 3 interjeccions que expressen efectivament dubte.

Fins aquí hem vist diferents vies de recuperació dels lemes que faciliten aquestes eines electròniques. És evident que també es podrien obtenir resultats a partir de la combinació de diferents elements d’aquest tipus —opció que preveuen l’Hiper i la versió en línia del PDL—, però sempre hi ha unitats que no es recuperen ja que no hi ha una sistematicitat total ni en l’atribució de la categoria gramatical ni en l’estructura definitòria de les unitats d’un mateix tipus.

## 2.2 Navegació lliure: trajectòries alternatives

A banda dels diversos sistemes de consulta —tres dels quals hem descrit detalladament en els apartats anteriors—, els diccionaris incorporen altres elements de caràcter hipertextual, que permeten el que hem denominat *navegació lliure*. L’usuari pot moure’s amb llibertat per tots

els hipervincles que faciliten les eines analitzades, sense seguir cap camí prèviament establert. És a dir, pot arribar a la informació associada navegant a través de les obres.

Pel que fa a l'Hiper, els lemes són accessibles des de qualsevol article lexicogràfic que els contingui, ja sigui formant part d'una definició o com a remissió pròpiament dita. Un exemple seria la interjecció “*manoi*”, que està definida com a “*Renoi!*”. Si fem doble clic sobre aquesta unitat, saltem directament a la definició de “*Renoi*” (*Renoi. Exclamació que expressa l'admiració, la sorpresa*). I des d'aquí podem passar a les definicions que el diccionari dóna de “*admiració*” i “*sorpresa*”. Pel que fa als connectors, des de la definició de “*és a dir*” podem passar a les de “*aclariment*” i “*explicació*”. Ara bé, aquesta funció no es pot aplicar a unitats polilexemàtiques. Així, tenim definida la interjecció “*Àngela (o Àngela Maria)*” amb l'expressió sinònima “*justa la fusta*”, però si volem consultar l'article lexicogràfic d'aquesta unitat ens trobem que el diccionari només permet saltar hipertextualment a cada una de les unitats simples (“*justa*”, “*la*” i “*fusta*”) però no pas a la unitat complexa, tot i que està inclosa com a subentrada al diccionari. El mateix succeix amb els connectors: en la definició del connector “*en realitat*” es remet a “*de fet*”, però no podem saltar directament a *de fet*, sinó que fent un clic sobre una de les dues unitats accedim a l'entrada “*de*” o “*fet*”.

A banda d'aquests enllaços hipertextuals, els articles lexicogràfics també inclouen uns botons que condueixen a informació addicional (botó de morfologia i fonètica, botó de traduccions angleses i botó de traduccions castellanques).

Pel que fa al el DIEC, incorpora alguns elements hipertextuals més a banda del sistema de recuperació d'entrades i subentrades. Així, igual que l'Hiper, es pot accedir directament a l'entrada de qualsevol paraula inclosa al text del diccionari fent-hi doble clic, independentment que estigui flexionada o no. Per exemple, si ens trobem a l'article “*Patatxup m. Patatxap*” podem fer clic sobre *patatxap* i accedirem directament a l'article corresponent a aquesta unitat. Ara bé, trobem un cop més que aquesta via només és possible en el cas d'unitats simples i que no es pot accedir a unitats complexes. Altres elements hipertextuals que incorpora el DIEC en CD-ROM és la possibilitat de visualitzar el desglossament de les abreviatures i l'accés a remissions específiques (a una accepció concreta, una subentrada, etc.) mitjançant el botó “*R*”.

Finalment, en el PDL la hipertextualitat és molt menor. Un cop feta una cerca, amb filtres o sense, es genera una llista d'unitats que s'adiuen amb els criteris establerts que enllacen hipertextualment amb l'article lexicogràfic corresponent. Ara bé, quan es visualitza l'article, no es pot fer clic a sobre de cap unitat ni accedir a l'article corresponent. L'única via alternativa d'accedir a la informació és a través del corpus textual, però només permet interrogar sobre formes simples i, per tant, les que tenen més d'un element —com tots els connectors i una part significativa d'interjeccions estudiades en aquest article— no disposen d'aquesta via addicional

En resum, a banda dels sistemes de consulta que presenten els diccionaris, també trobem enllaços hipertextuals en el text dels diccionaris en CD-ROM i en línia, amb la limitació que només donen accés a unitats simples.

Fins aquí hem vist les diverses formes d'accedir a la informació sobre els connectors reformulatius i les interjeccions que preveuen les eines electròniques que hem analitzat.



Veiem que l'accessibilitat a les dades associades a aquestes unitats de caràcter més aviat pragmàtic és limitada. Com a generalització podem dir que:

- a) faciliten l'accés directe al lema sempre que les unitats cercades no siguin polilexemàtiques,
- b) possibiliten altres formes d'accés a la informació però amb una precisió bastant baixa i un grau de cobertura, quan n'hi ha, del voltant del 55%.

Per aquestes raons, en la segona part de l'article presentem dues aplicacions que pretenen millorar la localització de la informació associada a les unitats lèxiques de caràcter pragmàtic a partir de la pluriaccessibilitat que permet l'hipertext.

### 3. LA PLURIACCESSIBILITAT, EIX DE LA LEXICOGRAFIA HIPERTEXTUAL

Entenem que una obra lexicogràfica és pluriaccessible quan obre diversos camins d'accés a la informació, en oposició als diccionaris que només tenen un camí de cerca i que Hausmann i Wiegand (1989: 338) anomenen “monoaccessibles”. Exemples d'obres lexicogràfiques pluriaccessibles solen ser els diccionaris hipertextuals en suport electrònic —com els que hem analitzat en aquest article—, tot i que el grau de pluriaccessibilitat divergeix segons l'explotació que es fa de l'hipertext.

En aquest article propugnem que la pluriaccessibilitat és un element clau de la lexicografia hipertextual i la considerem l'eix vertebrador sobre el qual s'haurien de construir els diccionaris hipertextuals electrònics. També defensem que aquesta pluriaccessibilitat s'ha d'explotar al màxim tenint presents les característiques pròpies de les unitats descrites.

Així, en el cas de les unitats seleccionades en aquest treball, creiem que seria útil oferir eines que milloressin la navegació dirigida —preveient sistemes de recuperació de lemes adequats als connectors i a les interjeccions— i també la navegació lliure —fent possible que la informació continguda sigui accessible des de múltiples punts.

Pel que fa a la navegació dirigida, proposem un nou accés directe a aquestes peces prenent com a punt de referència les funcions que desenvolupen, sense deixar de banda l'accés tradicional per lema. D'aquesta manera, es podrien recuperar en bloc totes aquelles peces que vehiculen unes mateixes instruccions funcionals i que en els recursos lexicogràfics actuals solen ser difícils de recuperar conjuntament.

Pel que fa a la navegació lliure, proposem noves relacions entre els elements aprofitant al màxim l'hipertext, tenint present que el resultat ha de ser una eina útil i no pas un cúmul d'informacions que facin perdre l'usuari.

#### 3.1 Propostes d'aplicació

En aquest apartat presentarem dues aplicacions hipertextuals que han estat concebudes de bon principi com a eines en suport digital i que presenten de forma pluriaccessible la informació

associada als connectors i a les interjeccions, respectivament.<sup>9</sup> Es tracta de *prototips parcials* (Lorente, 1996) perquè descriuen un aspecte concret del lèxic català i perquè indirectament poden oferir pautes per a la renovació de la descripció d'unitats pragmàtiques. Com a tals, no són comparables amb els diccionaris generals descrits prèviament, tot i que la filosofia que ha guiat la seva concepció es podria traslladar fàcilment a altres tipus d'unitats i a altres tipus de diccionaris.

Les aplicacions que presentarem a continuació són:

- (i) *ALCOR*, una eina destinada als lexicògrafs que inclou connectors reformulatius en català, i
- (ii) *INTCA*, una eina en fase d'elaboració que inclou interjeccions en català i en anglès, i que està destinada a professionals de la llengua que volen produir textos.

En aquest apartat aprofundirem en els elements hipertextuals que inclouen, tant en l'accés als lemes com en l'accés a la informació associada a aquests lemes. Els trets comuns més destacables de les dues propostes són que ambdues eines es basen en treballs previs de caràcter teoricodescriptiu<sup>10</sup> i que la informació és accessible des de molts punts diferents gràcies a l'hipertext.

Les aplicacions es presenten en un format hipertextual que permet presentar les dades de les aplicacions de forma plàstica, de tal manera que els usuaris poden adaptar-les a les seves necessitats. Són aplicacions hipertextuals i dinàmiques en el sentit que no es limiten només a la selecció per part de l'usuari de pàgines HTML prèviament construïdes, sinó que l'usuari hi interactua: les pàgines d'entrada a les aplicacions són capaces de presentar la informació de maneres diferents en funció de la interacció amb l'usuari.

#### *a) Navegació dirigida*

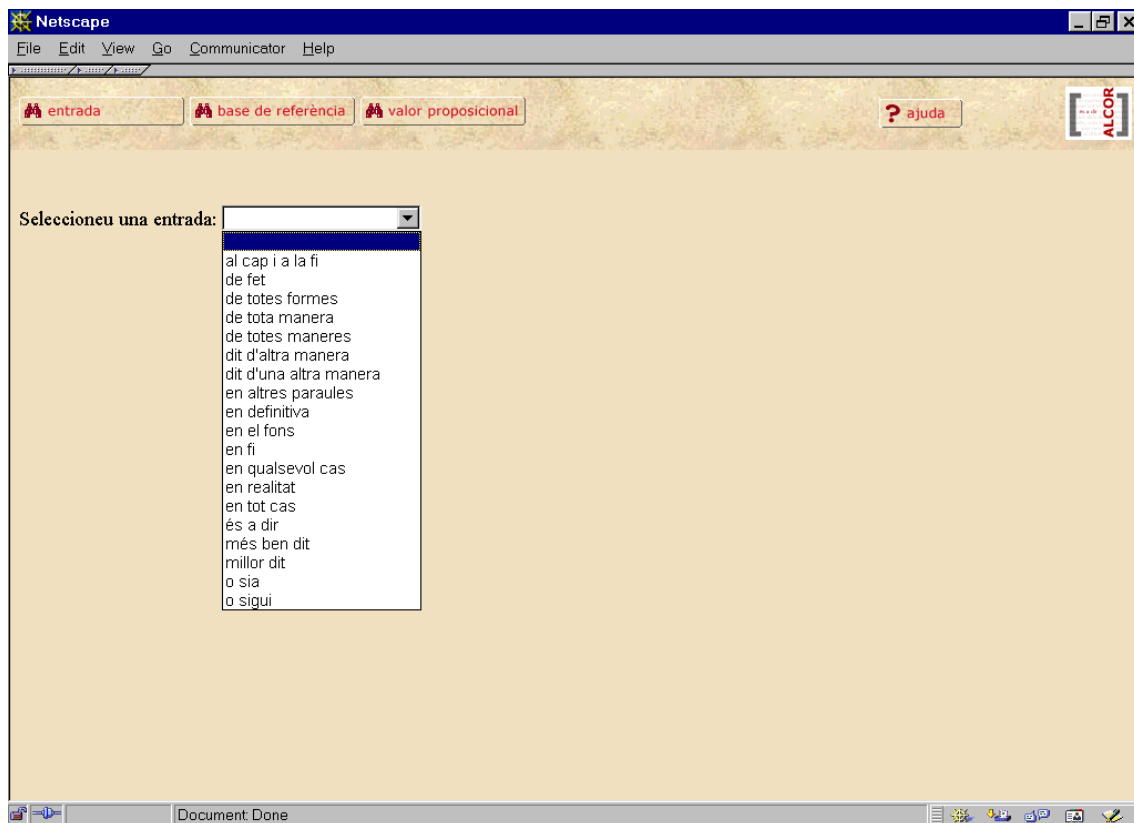
- Pel que fa al lema, totes dues eines ofereixen accessos directes a la peça lèxica que seleccionen, és a dir, cada connector i cada interjecció presenten una entrada pròpia: *o sigui, de fet, ai, ves*, etc.
- Pel que fa a la funció, les dues eines possibiliten accessos diferents a la informació de les peces lèxiques perquè, com ja hem destacat anteriorment, pretenen descriure dos tipus d'unitats diferents.

*ALCOR* presenta, d'una banda, un accés directe als lemes, en totes les seves variants. D'aquesta manera, l'usuari pot accedir a la informació associada a cada connector a partir de qualsevol de les diverses formes que adopta en el discurs.

---

<sup>9</sup> Una altra aplicació que aprofita al màxim l'hipertext i que s'ha desenvolupat en el marc de l'IULA és el DSVC de Bernal (2000).

<sup>10</sup> *ALCOR* forma part de la tesi doctoral de Bach (2001). Pel que fa a *INTCA*, el disseny es presenta a Matamala (2001) i es preveu implementar-lo a Matamala (en curs).



*Figura 1: Menú que ofereix ALCOR per a l'accés directe als lemes*

De l'altra, pel fet de ser una eina pensada inicialment com a prototip de treball per als lexicògrafs, ofereix també un accés als lemes a través de totes les bases lèxiques plenes que formen cada connector (figura 2).



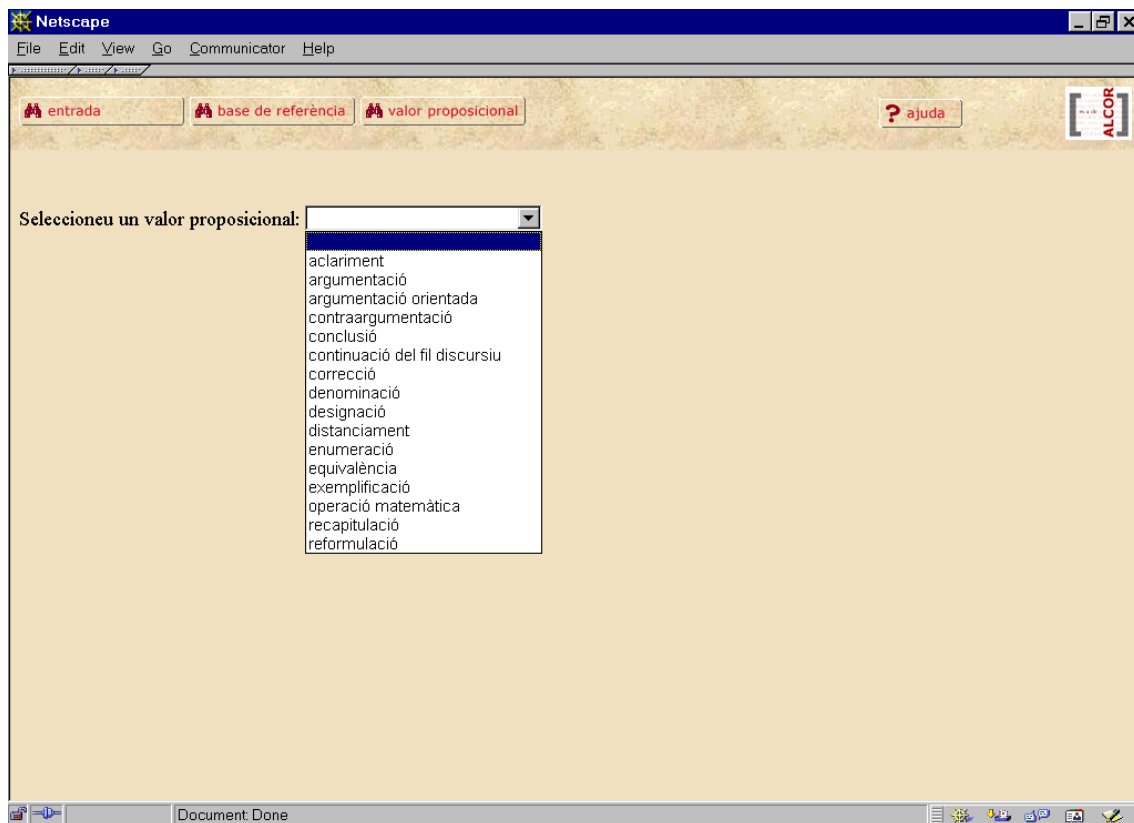
Figura 2: Menú que ofereix ALCOR per a l'accés als lemes des de la base de referència

Així es pot seleccionar la *base de referència* "manera" i es possibilita l'accés a tots els connectors reformulatius formats per aquesta base (figura 3).



Figura 3: Accés als lemes des de la base de referència “manera”

Finalment, *ALCOR* ofereix un altre accés directe a les peces lèxiques a través de la funció bàsica que comparteixen totes les ocurrences d'un mateix connector —denominada "valor proposicional": *aclariment*, *exemplificació*, *correcció*, etc. La resta de trets funcionals del connector, que configuren la seva perifèria funcional i que s'actualitzen en interacció amb el context, es descriuen dins l'entrada de diccionari.



*Figura 4: Menú que ofereix ALCOR per a l'accés per funcions*

Si es fa clic en aquestes funcions, n'apareix la definició i tots aquells exemples de connectors del corpus analitzat a Bach (2001) que les vehiculen.

*INTCA* proposa un accés per lemes en dues llengües, anglès i català, ja que es tracta d'una eina bilingüe.

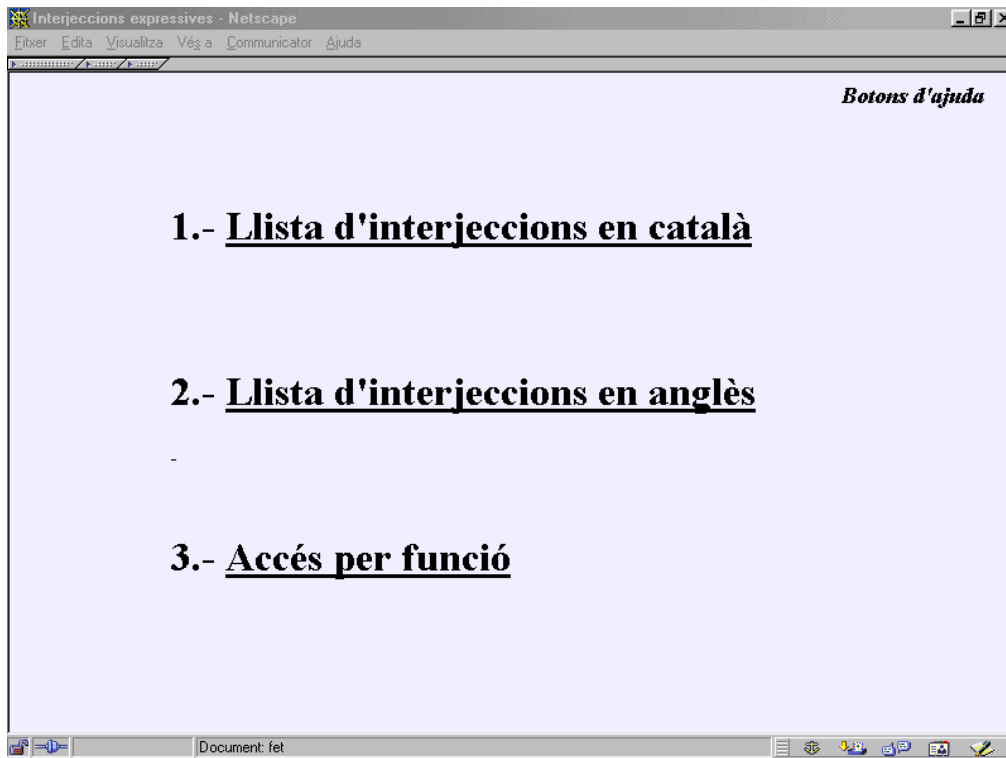
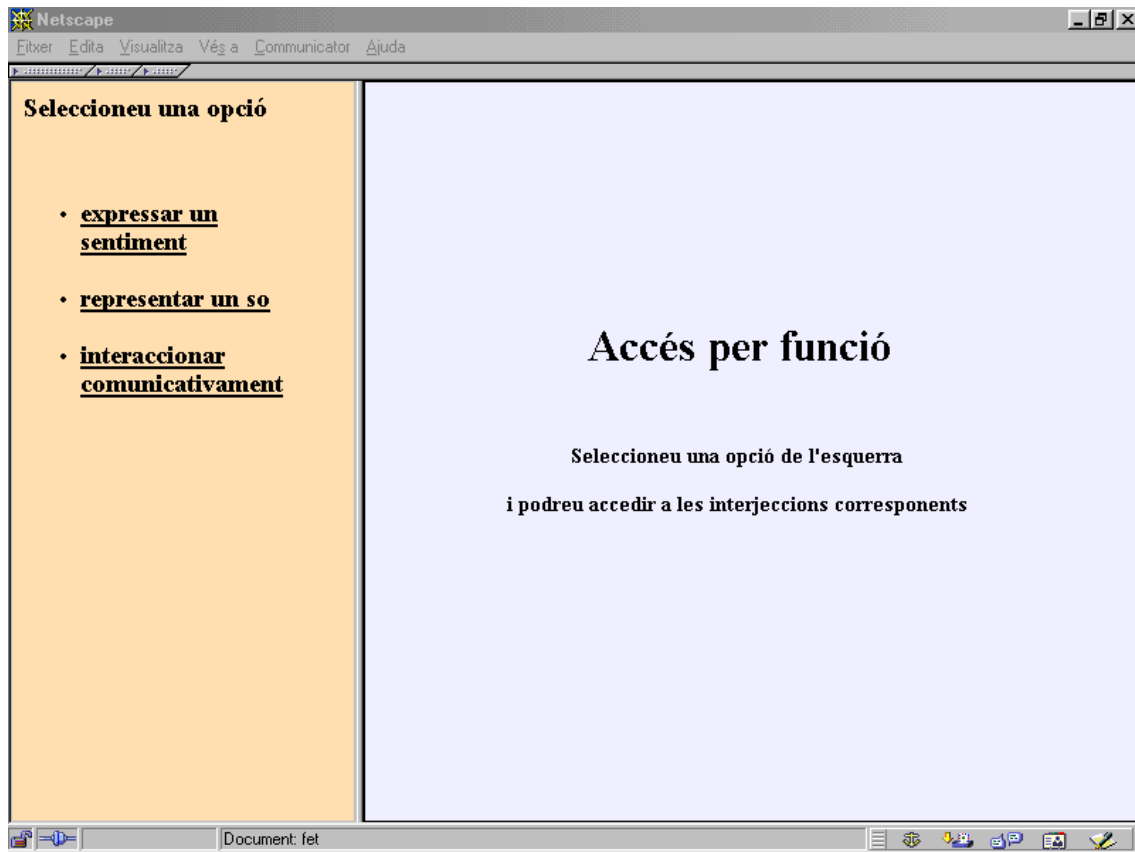


Figura 5: Menú inicial d'INTCA per a l'accés a la informació associada als lemes

També proposa un accés per funcions, amb diferents nivells de decisió. En un primer nivell, l'usuari selecciona una funció general i decideix si busca unitats per *expressar un sentiment*, *representar un so* o *interaccionar comunicativament*.



*Figura 6. Accés per funcions generals proposat per INTCA*

En un segon nivell, l'usuari tria una funció més específica que anomenem "element bàsic" —en el cas de les interjeccions expressives podria ser *alegria*, *tristesa* o *por*, per posar uns quants exemples— i, si vol, pot restringir la cerca per "elements de variació" —per exemple, les marques funcionals. El resultat és una llista d'interjeccions que s'adiuen amb les condicions expressades i que, al mateix temps, són enllaços hipertextuals que ens porten a l'entrada corresponent en aquella unitat.



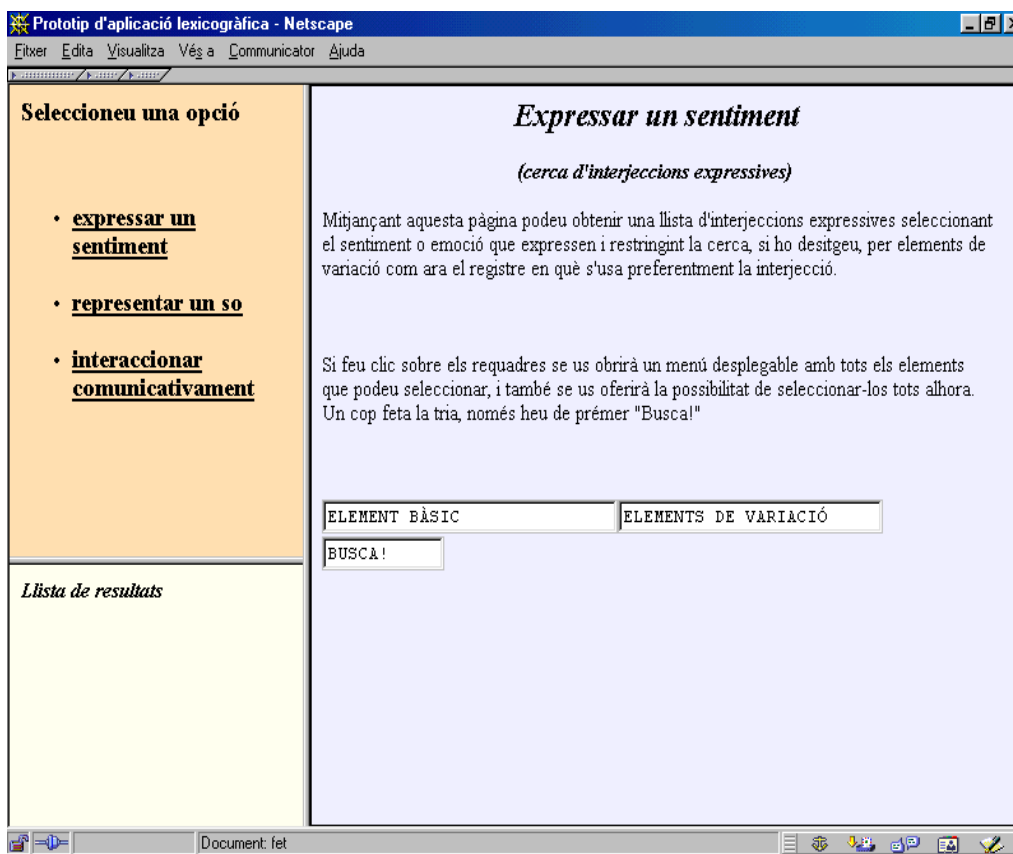


Figura 7. Pantalla d'accés a les interjeccions expressives.

#### b) Navegació lliure



A banda d'oferir diverses vies d'accés als lemes, les possibilitats de navegació per les eines són múltiples.


ALCOR explota al màxim les possibilitats de l'hipertext en la presentació de la informació associada a cada lema.

Les etiquetes que fan referència a la categoria gramatical de les unitats descrites es vinculen hipertextualment amb la seva definició i amb totes aquelles unitats descrites que les comparteixen. D'aquesta manera es pot fer clic sobre l'etiqueta "connector" i s'accedeix a la definició proposada i a tots els connectors que s'han implementat a l'eina.

De la mateixa manera, les instruccions funcionals associades a cada connector permeten també l'accés a la seva definició i a tots els connectors reformulatius implementats que les vehiculen.

Finalment, des de la pantalla de la informació associada als lemes, es pot fer clic a tres botons diferents que també permeten la navegació per l'eina:

-  **freqüència** per consultar la freqüència d'aparició del connector respecte al corpus utilitzat (que pot recalcar-se sempre que es decideixi canviar el corpus de partida),
-  **contextos d'aparició** per consultar els contextos d'aparició del connector i les unitats amb què es combina i,

-  **format discursiu** per accedir a una altra fitxa en què hi ha la mateixa informació però desplegada.

En el cas de l'*INTCA*, es preveu que l'usuari podrà consultar de manera hipertextual una gran quantitat d'elements associats al lema, com ara la categoria gramatical, el tipus d'interjecció, les variants formals del lema, la freqüència de l'accepció, i també propostes sinonímiques (en el cas de la versió monolingüe) i de traducció (en el cas de la bilingüe) amb els diferents graus d'equivalència explicats. Tota aquesta informació serà consultable des de diferents punts de l'eina de manera que, partint de l'article lexicogràfic d'una interjecció, l'usuari podrà accedir a altres articles lexicogràfics i a presentacions alternatives de la informació. També es preveu vincular aquesta eina amb la informació disponible a la xarxa mitjançant enllaços a diccionaris de referència. Finalment, vista la importància de l'entonació en aquest tipus d'unitat, està previst incloure-hi arxius d'àudio i de vídeo a tall d'exemple.

En resum, aquestes eines pretenen obrir el ventall de rutes fins als lemes tot potenciant l'accés per funcions —via especialment útil en el cas d'unitats amb un caràcter marcadament pragmàtic—, i també pretenen fer accessibles des de múltiples punts les diverses informacions contingudes. En resum, volen explotar l'hipertext per ser eines autènticament pluriaccessibles.

#### 4. SÍNTESI

En la primera part de l'article hem fet un repàs de la concepció d'hipertext i de la seva relació amb la lexicografia en qualsevol dels seus formats. Tot seguit, per tal de detectar els encerts i les mancances de les implementacions actuals de l'hipertext en lexicografia, hem analitzat els sistemes de navegació que preveuen l'Hiper, el DIEC i el PDL per a connectors i interjeccions, unitats definibles mitjançant la metallengua del signe. Aquesta anàlisi ha posat de relleu que la implementació de l'hipertext en aquestes eines no permet recuperar satisfactòriament la informació lexicogràfica associada a aquestes unitats. En vista d'això, hem plantejat que una implementació adequada de l'hipertext en lexicografia possibilitaria la creació d'obres pluriaccessibles, és a dir, d'eines que facilitarien la recuperació de la informació obrint múltiples vies d'accés. Com a exemple, hem presentat dues eines (*ALCOR* i *INTCA*) que recullen connectors i interjeccions, respectivament, i que s'han dissenyat seguint aquesta idea.

## 5. BIBLIOGRAFIA

- BACH, C. (2001) *Els connectors reformulatius catalans. Anàlisi i proposta d'aplicació lexicogràfica*, Barcelona: Institut Universitari de Lingüística Aplicada. Tesi doctoral.
- BALASUBRAMANIAN, V. (1994) *State of the art review on hypermedia issues and applications*, URL: <http://cbl.leeds.ac.uk/nikos/tmp/hypemedia/hypemedia.html>. [Consulta: 27 d'abril del 2002]
- BERNAL, E. (2000) *Els sufixos verbalitzadors del català. Relacions semàntiques i diccionari*, Barcelona: Institut Universitari de Lingüística Aplicada. Tesi doctoral.
- BUSH, V. (1945) *As we may think*. Per a més informació, podeu consultar: <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>. [Consulta: 27 d'abril del 2002]
- CABRÉ, M. T., LI. DE YZAGUIRRE, M. LORENTE, J. MOREL i A. ROJO (1996) "A hypermedia project on terminology: the design of a telematic hypertextual dictionary", a *TKE'96 Terminology and Knowledge Engineering*. Frankfurt: Indeks Verlag, 424-430.
- CODINA, L. (1998) *Introducción a la navegación y a la teoría hipertextual*. Publicació electrònica disponible a <http://camelot.upf.es/~lcodina/public01.htm>. [Consulta: 27 d'abril del 2002]
- FUENTES MORÁN, T. (1995) *Die Grammatik des Spanischen in zweisprachigen Wörterbüchern des Sprachenpaares Spanish-Deutsch für Deutschsprachige*. Augsburg: Lexicographica. Tesi doctoral, 50-110.
- GELPÍ, C. (2001) "Lexicografia bilingüe, especialitzada per la temàtica i hipertextual", a *Estudis de Lexicografia 1999-2000*, Barcelona, Institut Universitari de Lingüística Aplicada, 159-174.
- HAUSMANN, F.J. i H.E. WIEGAND (1989) "Component parts and structures of general monolingual dictionaries: a survey", a HAUSMANN, F.J. (ed.) *Wörterbücher: ein internationales Handbuch zur Lexicographie*. Berlín: Walter de Gruyter, 328-360.
- LORENTE, M. (1996) "La lexicografia teòrica: els prototips de diccionaris", *Estudis de llengua i literatura catalanes*, Barcelona: Publicacions de l'Abadia de Montserrat, 339-363.
- MATAMALA, A. (2001) *Les interjeccions als diccionaris. Teoria, anàlisi i primera proposta de representació lexicogràfica*, Barcelona: Institut Universitari de Lingüística Aplicada. Treball de recerca.
- \_\_\_\_\_ (en curs) *Les interjeccions com a unitats dinàmiques: descripció en textos i proposta d'aplicació*. Tesi doctoral.

NELSON T. (1992) *Literary machines: the report on, and of, project Xanadu concerning world processing*, Sausalito: Mindful Press.

PRUVOST, J. (2000) *Dictionnaires et nouvelles technologies*, París: PUF.

TARP, S. (1999) “Theoretical foundations of the so-called cross-reference structures”, a WIEGAND, H.E. (ed.): *Mediostrukturen in Printwörterbüchern*. *Lexicographica*, 15. Tübingen: Max Niemayer Verlag, 114-137.

#### OBRES LEXICOGRÀFIQUES CONSULTADES

*Diccionari de la llengua Catalana*, Barcelona: Institut d’Estudis Catalans. 1995, recurs electrònic. [DIEC]

*Hiperdiccionari català-castellà-anglès*, Barcelona: Enciclopèdia Catalana, recurs electrònic. [HIPER]

*Portal de dades lingüístiques*, Barcelona: Institut d’Estudis Catalans, versió en línia. [Consulta: 27 d’abril del 2002] [PDL]