

Mètodes empírics en lingüística cognitiva

El cas d'un experiment psicolingüístic per a
avaluar l'estatus cognitiu de la macrodistinció
aspectual entre estats i esdeveniments

Document de treball

Marta Coll-Florit (mcollfl@uoc.edu)
Becària d'investigació doctoral de l'IN3 (UOC)

Working Paper Series WP07-004

Grup de recerca: Language Processing Group (LPG)
Coordinador del grup de recerca: Salvador Climent (UOC)

Data de recepció: juny de 2007
Data d'acceptació: juny de 2007
Data de publicació: juliol de 2007



www.uoc.edu

Internet Interdisciplinary Institute (IN3)

<http://www.uoc.edu/in3>

Parc Mediterrani de la Tecnologia
Av. Canal Olímpic, s/n.
08860 Castelldefels
Barcelona (Espanya)

Tel. 93 673 50 00

Universitat Oberta de Catalunya (UOC)

<http://www.uoc.edu/>

Av. Tibidabo, 39-43
08035 Barcelona
Espanya

Tel. 93 253 23 00

Sumari

| | |
|---|----|
| <i>Resum</i> | 4 |
| 1. Motivació de l'experiment | 5 |
| 1.1. L'ambigüitat semàntica i les diferències interlingüístiques de lexicalització..... | 6 |
| 1.2. Aktionsart com a criteri útil per al tractament computacional de la polisèmia verbal | 7 |
| 2. La problemàtica en l'estudi de l'Aktionsart: cap a una metodologia empírica | 9 |
| 3. Experiment | 13 |
| 3.1. Constitució de la mostra experimental..... | 13 |
| 3.2. Procediment i disseny..... | 16 |
| 4. Resultats | 16 |
| 4.1. Diferència genèrica en temps de processament entre esdeveniments i estats | 17 |
| 4.2. Tres nivells d'anàlisi dels resultats. El cas de <i>cumplir</i> | 19 |
| 4.2.1. Diferències en cost de processament entre els dos sentits del verb <i>cumplir</i> | 19 |
| 4.2.2. Els prototips. Els parells <i>lee-sabe</i> i <i>logra-cree</i> | 20 |
| 4.2.3. A quin prototip s'acosta més cada sentit del verb <i>cumplir</i> ?..... | 21 |
| 5. Conclusions | 23 |
| Agraïments | 24 |
| Referències bibliogràfiques | 25 |
| Annex | 28 |

Mètodes empírics en lingüística cognitiva

El cas d'un experiment psicolingüístic per a avaluar l'estatus cognitiu de la macrodistinció aspectual entre estats i esdeveniments

Marta Coll-Florit (mcollfl@uoc.edu)

Becària d'investigació doctoral de l'IN3 (UOC)

Resum

Una de les conseqüències fonamentals de l'auge de les tecnologies lingüístiques és que han posat de manifest la complexitat del fenomen lingüístic i la necessitat de desenvolupar coneixements exhaustius i interdisciplinaris sobre tots els aspectes del llenguatge. Per tal d'avançar, cal el treball empíric i conjunt de lingüistes, psicòlegs, neuròlegs i informàtics. Concretament, un dels principals problemes amb què topen les tecnologies del llenguatge és que han de resoldre l'ambigüitat intrínseca a les llengües, com és el cas de l'ambigüitat que presenten els verbs polisèmics. Per tant, ens trobem amb la necessitat de disposar de criteris formals per a identificar i destriar sentits verbals. En aquest article presentem l'aspecte lèxic o Aktionsart com a un dels criteris claus per a l'establiment de sentits verbals.

Amb l'objectiu final de proposar un model formal de representació de les categories aspectuals que sigui implementable computacionalment i que, alhora, estigui motivat cognitivament, s'ha desenvolupat un experiment psicolingüístic per tal d'avaluar empíricament l'estatus cognitiu de la macrodistinció aspectual entre estats i esdeveniments i, de manera relacionada, per a validar la relació entre aspecte i polisèmia verbal. En aquest article presentarem en detall la metodologia seguida per a confeccionar la mostra experimental, el disseny i l'aplicació de l'experiment dut a terme, així com un avenç dels primers resultats que apunten a la verificació de les hipòtesis de partida.

Paraules clau

Tecnologies Lingüístiques, Aktionsart, Polisèmia, Processament verbal, Lingüística Cognitiva, Mètodes empírics, Experiment psicolingüístic

Citació recomanada:

COLL-FLORIT, Marta (2007). *Mètodes empírics en lingüística cognitiva. El cas d'un experiment psicolingüístic per a avaluar l'estatus cognitiu de la macrodistinció aspectual entre estats i esdeveniments*. [document de treball en línia]. UOC. (Working Paper Series; WP07-004). [Data de citació: dd/mm/aa]. <<http://www.uoc.edu/in3/dt/cat/coll-florit07004.pdf>>

1. Motivació de l'experiment

El llenguatge humà és d'una importància clau en la societat de la informació i el coneixement actual en tant que constitueix un dels mitjans més potents per codificar i transmetre informació. D'aquí que sorgeixi la necessitat de crear eines i recursos informàtics que aprenguin a descodificar el sistema lingüístic i arribin a generar informació a partir del sistema lingüístic. És en aquest marc on es desenvolupen les anomenades tecnologies del llenguatge humà que consisteixen en l'aplicació del coneixement sobre el llenguatge al desenvolupament de sistemes informàtics capaços de reconèixer, analitzar, interpretar i generar llenguatge. El resultat són màquines que emulen la capacitat lingüística humana.

Seguint Lavid (2005), els objectius bàsics d'aquells que treballen amb tecnologies del llenguatge es poden sintetitzar en tres grans propòsits. D'una banda, es pretén assolir una interacció més natural i eficient entre persona i màquina de manera que, per exemple, arribem a parlar amb els nostres ordinadors a casa, al treball, dins del cotxe o a un lloc públic quan necessitem informació o ajut. D'altra banda, es vol facilitar el maneig de la informació per a afavorir l'adquisició de coneixement; és a dir, que els paràmetres de cerca, selecció, resum i presentació de la informació siguin més flexibles i es puguin a adequar a les necessitats de cada usuari. Finalment, un altre dels objectius primordials de les tecnologies del llenguatge és mantenir i desenvolupar una autèntica societat plurilingüe en la que cada usuari pugui accedir i rebre la informació en la seva pròpia llengua, tot superant les barreres de la globalització lingüística imperant.

En síntesi, necessitem ordinadors que siguin capaços de comprendre i generar llenguatge, un objectiu que no se'ns presenta fàcil d'assolir en tant que hem d'ensenyar a una màquina a realitzar tasques lingüístiques que els humans executem de forma tàcita o implícita. I és que hem de tenir present que els humans som capaços de generar i produir un nombre infinit d'expressions carregades de significat apropiades al context situacional, tot basant-nos en grans quantitats de coneixement lingüístic, coneixement social i cultural, així com l'anomenat coneixement del món (el sentit comú, el coneixement de la realitat que ens envolta). Alhora, usem altres capacitats cognitives com ara la capacitat d'abstracció, la memòria o la focalització de l'atenció. Per tant, per tal de formalitzar l'enorme complexitat que amaga qualsevol llengua i arribar a crear programes informàtics que la reproduïxin, el primer pas serà disposar de coneixement explícit i interdisciplinari sobre el funcionament del llenguatge, així com explicar de quina manera es materialitzen en les diverses llengües els principis genèrics subjacents a la capacitat lingüística humana.

1.1 L'ambigüitat semàntica i les diferències interlingüístiques de lexicalització

Un dels principals problemes amb què topen les tecnologies del llenguatge és el de resoldre l'ambigüitat inherent a les llengües. Una oració que admeti més d'una anàlisi o una paraula amb diferents sentits sol ser la norma, més que no pas l'excepció. És el cas, per exemple, de l'ambigüitat que presenten els verbs polisèmics, com s'exemplifica a (1) amb el cas del verb polisèmic *tocar*.

(1) *La Lola tocà la guitarra i la trencà*

L'oració de (1) té, com a mínim, dues possibles interpretacions: o bé que la Lola va entrar en contacte físic amb la guitarra (la va fregar), va caure i es va trencar, o bé que va interpretar una peça musical amb la guitarra i ho feia tant malament que en va trencar les cordes. Aquí, per tant, veiem clarament dos possibles sentits del verb *tocar*: entrar en contacte físic amb una superfície o interpretar una peça musical. Més enllà d'aquests dos sentits, *tocar* és un verb molt polisèmic. Si en consultem els sentits en castellà al diccionari de la Real Academia Española (DRAE) comprovem que es contemplen fins a 28 significats. El que aquí ens interessa subratllar, no obstant, és que si consultem un altre diccionari, com ara el *Vox Diccionario de Uso del Español de América y de España*, ens adonem que només es recullen 18 sentits, fet que revela la diversitat de criteris en l'establiment de sentits. El problema bàsic és que molts d'aquests diccionaris es basen en la intuïció o la introspecció del lexicògraf que els crea, sense uns criteris clars i consensuats per a diferenciar sentits.

La problemàtica, però, no acaba aquí. I és que, encara que hipotèticament totes les paraules d'una llengua tinguessin només un significat, seguiríem tenint dificultats de traducció. Cada llengua conceptualitza i codifica la realitat de manera diferent. En una llengua podem tenir dos o més termes que es corresponen a un sol terme en una altra llengua, o bé conceptes que es lexicalitzen en una llengua però no en una altra. Un dels exemples més coneguts és el de la llengua dels esquimals inuits: tenen fins a 20 maneres diferents de referir-se al color blanc, mentre que en català o castellà només en tenim una ('blanc'/ 'blanco'). Lexicalitzem l'espectre cromàtic de maneres diferents. El cas és, però, que no cal comparar llengües tan llunyanes per a trobar diferències de lexicalització. Per exemple, en anglès trobem dues paraules diferents per a designar els dits de les mans (*fingers*) i dels peus (*toes*), dos conceptes que es lexicalitzen en una sola paraula en català (*dits*) i castellà (*dedos*).

Anant encara més enllà, aquestes diferències interlingüístiques de lexicalització no només les trobem entre noms, sinó també entre verbs. Un bon exemple el tenim en els verbs anglesos *meet*, *know* i *taste* que no mantenen una relació biunívoca amb els seus respectius verbs en castellà i català. En podem veure el mapa semàntic a la Taula 1.

| | | | |
|-----------------|----------|-------|-------|
| Anglès | Meet | Know | Taste |
| Castellà | Conocer | | |
| | | Saber | |
| Català | Conèixer | | |
| | | Saber | |

Taula 1. Mapa semàntic dels verbs anglesos *meet*, *know* i *taste*, respecte al castellà i català

Observem que el verb *conocer* (castellà) i *conèixer* (català) lexicalitzen en una sola paraula tant el sentit del verb anglès *know* ('conèixer, tenir informació sobre algú o alguna cosa') com el sentit del verb anglès *meet* ('conèixer algú per primera vegada'). Alhora, notem que en castellà tenim el verb *saber* que inclou el sentit del *Know* (però no del *meet*), així com també el sentit del verb anglès *taste* ('tenir gust'), un últim sentit que no incorpora el verb *saber* català. Per tant, tenim tot un entramat de patrons de lexicalització que necessitem estudiar per tal d'arribar a formalitzar el contingut semàntic de cada verb i poder establir equivalències entre llengües.

1.2. *Aktionsart* com a criteri útil per al tractament computacional de la polisèmia verbal

Perquè un sistema de PLN sigui capaç de processar o generar un text o parla que inclogui verbs polisèmics (la gran majoria), cal crear models o algorismes que determinin quin sentit del verb s'està utilitzant en cada context, procediment que es coneix amb el nom de *word sense disambiguation* (Aguirre & Edmons, 2006; Resnik 2004). Ara bé, per a poder emprendre aquesta tasca, en primer lloc, hem de saber quin és el conjunt de sentits d'un verb i, en conseqüència, cal establir criteris per a destriar, identificar i representar sentits. Per tant, ens trobem amb la necessitat d'estudiar la semàntica verbal (quins són els components bàsics del significat verbal i com els processem cognitivament els parlants) per a poder-ne establir un model de representació apte per a ser processat computacionalment.

La necessitat de creació de recursos lèxics per al Processament del Llenguatge Natural ha provocat en els últims anys un auge de la Lexicografia Computacional, disciplina de la Lingüística que s'ocupa de la formalització sintàctico-semàntica del lèxic per a aplicacions computacionals. Un símptoma evident d'aquest creixement és l'emergència de projectes que desenvolupen bases lèxiques computacionals, com és el cas de la xarxa lèxica multilingüe EuroWordNet (Vossen et al., 1998) o la més recent OntoNotes (Hovy et. al, 2006); els projectes FrameNet (Johnson & Fillmore, 2000) i PropBank (Kingsbury & Palmer, 2002) per a l'anglès; o el projecte SenSem (Castellón et al. 2006) i ADESSE (García-Miguel et al., 2005) per a l'espanyol. Aquests projectes adopten diferents perspectives a l'hora de tractar la informació lingüística: alguns tracten la semàntica lèxica de diferents categories (noms, verbs i adjectius), mentre que d'altres només inclouen informació

verbal; uns tracten exclusivament informació sintàctica (esquemes de subcategorització) i d'altres també inclouen la semàntica oracional (papers temàtics, restriccions de selecció o estructura eventiva). Més enllà d'aquests diferents enfocaments, trobem una problemàtica comuna de tots els recursos a l'hora de caracteritzar el comportament verbal: la manca de criteris rigorosos i no-ambigus per a diferenciar sentits prou genèrics per a ser implementats computacionalment i, d'aquesta manera, poder tractar de manera adient la polisèmia verbal.

Des de la lingüística teòrica es considera que una forma essencial per a descriure el comportament verbal és la que es basa en l'Aspecte lèxic o *Aktionsart* (Vendler 1967), tota aquella informació que fa referència a l'estructura temporal interna de l'esdeveniment descrit per un verb o predicat. És així com s'han proposat diferents tipologies de classificació eventiva (Dowty, 1979; Pustejovsky, 1991; Tenny, 1994; Levin & Rappaport, 1995; Marín, 2000; De Miguel, 2004; entre d'altres) que es presenten com a diferents reformulacions de les classes aspectuals tradicionals proposades per Vendler (1967): estats, activitats, realitzacions i assoliments. És sorprenent, però, comprovar com la majoria d'autors (tant els que s'ocupen de l'estructura eventiva de tota l'oració, com els que estudien només la informació aspectual a nivell estrictament lèxic), adopten la peça lèxica verbal, i no el sentit, com a nivell més bàsic de classificació aspectual. Per tant, obvien el fet que la majoria de verbs són polisèmics i, de manera relacionada, que els diferents sentits d'un verb poden incloure diferent informació aspectual.

És en aquest punt en què considerem que l'aspecte lèxic és un dels criteris claus a tenir en compte en l'establiment de sentits verbals. En conseqüència, pensem que és fonamental disposar d'un model formal de representació de les categories aspectuals que sigui implementable computacionalment i que, alhora, estigui motivat cognitivament. Per tant, un dels objectius bàsics del nostre treball serà qüestionar-nos quins són els mecanismes cognitius que utilitzem per a conceptualitzar les diferències aspectuals, per tal d'oferir una proposta més clara i explicativa de les categories aspectuals i, de manera relacionada, de la polisèmia verbal.

Ens situem en el marc de la Lingüística Cognitiva, una aproximació a l'estudi lingüístic que va emergir als anys 80 a la costa oest dels Estats Units amb el propòsit fonamental de descriure la relació entre llenguatge i ment (Lakoff, 1987; Langacker, 1987; Talmy, 1985, 2000; Johnson, 1987; Croft & Cruse, 2004; Evans & Green, 2006; entre d'altres). Concretament, es parteix de dos supòsits bàsics: d'una banda, es reivindica que els processos cognitius que governen el llenguatge són els mateixos que els d'altres habilitats cognitives, com ara l'abstracció, la perspectiva, la memòria o la capacitat d'atenció. Per altra banda, s'afirma que el coneixement lingüístic és bàsicament estructura conceptual; és a dir, que no només la representació semàntica és bàsicament conceptual, sinó que també ho és la representació sintàctica, morfològica i fonològica.

Una de les principals crítiques que es van fer inicialment a la Lingüística Cognitiva és que mancava de rigor empíric en tant que es basava sobretot en intuïcions i interpretacions dels

investigadors. És així com en els últims anys s'han dirigit tots els esforços a desenvolupar mètodes empírics capaços de fer de la lingüística cognitiva un autèntica ciència (Cuycken et al. 1996; Boroditsky, 2000; Tyler & Evans, 2001; Gries, 2005; González-Márquez et al., en premsa; entre d'altres), tot basant-se en la recerca feta fins ara des de la psicologia cognitiva i la neurologia. Aquest interès bàsic de la lingüística cognitiva actual per trobar evidències empíriques del funcionament cognitiu del llenguatge és la motivació bàsica del treball que es presenta en aquest article. Concretament, exposem el disseny, implementació i primers resultats d'un experiment psicolingüístic que s'ha portat a terme amb l'objectiu d'avaluar l'estatus cognitiu de l'aspecte lèxic, així com per a obtenir evidències empíriques de la relació entre aspecte i polisèmia.

L'estructura de l'article és la següent. En primer lloc, per tal d'emmarcar teòricament l'experiment, exposarem l'estat de la qüestió en l'estudi de l'aspecte lèxic, així com aquelles qüestions actualment objecte de debat que fonamenten les hipòtesis de partida del nostre experiment (apartat 2.). Acte seguit presentarem els paràmetres que s'han tingut en compte per a dissenyar la mostra experimental i el procediment utilitzat per a la implementació (apartat 3). Finalment, analitzarem els resultats de l'experiment pel que fa a la primera hipòtesis i avançarem una síntesi dels resultats relatius a la polisèmia verbal (apartat 4). Així mateix, per tal d'il·lustrar amb detall la metodologia d'avaluació multinivell que s'ha seguit per a verificar la segona hipòtesi, ens centrarem en el cas del verb polisèmic *cumplir* (apartat 5.).

2. La problemàtica en l'estudi de l'aspecte lèxic: cap a una metodologia empírica

Com bé s'ha indicat a l'apartat anterior, un dels àmbits d'estudi més reeixits en semàntica verbal és l'aspecte lèxic o Aktionsart, tota aquella informació relacionada amb l'estructura temporal interna de la situació descrita per un verb o predicat (Vendler, 1967; Dowty, 1979; Verkuyl, 1993; De Miguel, 1999; Sasse, 2003; entre d'altres).

La distinció aspectual més genèrica és la que s'estableix entre estats i esdeveniments: els verbs eventius impliquen una o més fases temporals successives des d'una situació inicial a un estat resultant (ex. *destruir*, *córrer*), mentre que els verbs estatus impliquen una única fase temporal estable (ex. *saber*, *tenir*). Així mateix, es postulen subclasses dels esdeveniments. D'una banda, trobem la distinció entre esdeveniments tèl·lics (ex. *arribar*, *reconèixer*), que expressen una culminació temporal de l'esdeveniment, i els esdeveniments atèl·lics, exempts de límit temporal (ex. *jugar*, *córrer*). Finalment, alguns autors també estableixen diferents tipus aspectuals en funció de la duració de la situació descrita pel verb, un paràmetre que serveix per a diferenciar dos tipus d'esdeveniments tèl·lics: els assoliments (ex. *aturar-se*, *trobar*), que són momentanis o puntuals, i les realitzacions que, contràriament, són duratives (ex. *construir una casa*, *pintar un quadre*).

Per tal d'establir una tipologia clara dels tipus aspectuals i poder determinar si un verb pertany a una o altra classe aspectual, s'han proposat tota una sèrie de criteris o proves sintàctiques. Es parteix del supòsit lexicalista que la informació aspectual codificada a la peça lèxica verbal determina la realització sintàctica dels seus arguments. Seguint aquesta assumpció, tots els verbs que pertanyin a una mateixa classe aspectual compartiran un mateix comportament sintàctic. Així, en funció dels contextos oracionals en què pugui aparèixer el verb (contextos que indiquen [+/-dinamicitat], [+/- delimitació] o [+/-duració]), podrem classificar la peça lèxica en una o altra classe aspectual. No obstant, si analitzem en detall aquests criteris notem que, si bé disposem de proves bastant precises per a identificar la delimitació d'un predicat, no passa el mateix a l'hora de discriminar entre un verb estatiu i un de dinàmic, fet que ens obligarà a qüestionar-nos la validesa de les proves aspectuals i, en últim terme, la pròpia naturalesa de les classes aspectuals.

Per tal d'exemplificar perquè els tests no ens permeten establir un límit clar i precís en la distinció aspectual primigènia entre estats i esdeveniments, partirem d'un dels criteris més citats a la bibliografia per a identificar la dinamicitat d'un predicat: la incompatibilitat dels estats d'aparèixer amb la forma progressiva (<estar + gerundi>). Segons aquesta prova, els estats, en tant que no avancen en el temps ni expressen cap tipus de canvi, són incompatibles amb la forma progressiva, una construcció que indica dinamicitat. Tanmateix, si apliquem la prova a verbs classificats com a estats a la bibliografia, com ara *tener*, *conocer*, *querer* o *ignorar*, comprovem que, contràriament al que prediuen els tests, aquests verbs sí admeten la forma progressiva, com podem observar als exemples que es recullen a (2), extrets del *Corpus de Referencia del Español Actual* (CREA).

- (2) a. *El club no estaba teniendo ningún éxito.*
b. *Él siente miedo porque estoy conociendo mucha gente.*
c. *Te estoy queriendo de esta desafortada e imprudente manera*
d. *El Ejecutivo Federal está ignorando el artículo 8º constitucional*

Per tant, aquests verbs considerats tradicionalment estats poden expressar dinamicitat. Davant d'aquestes dades, podem optar per diferents vies d'interpretació. Una possible opció és simplement desestimar la prova com a no vàlida per a identificar els estats. Una altra possibilitat és acceptar els resultats del test i afirmar que aquests verbs realment no són estats. Finalment, tenim una tercera opció, que considerem més raonable, que és la d'assumir que una mateixa peça lèxica verbal pot conjugar propietats de més d'una classe aspectual i, en conseqüència, entenem que el test no s'ha aplicat adequadament. Notem, per tant, que la problemàtica dels tests rau, en primer lloc, en la pròpia concepció de les categories aspectuals, concebudes des d'una visió objectivista de la semàntica; és a dir, com a classes amb fronteres clarament definides, amb condicions necessàries i suficients per a tots els elements que formen part de la classe, i amb només dos graus de pertinença a la classe: membre o no-membre.

Tanmateix, com bé s'ha demostrat des del marc de la lingüística cognitiva, la gran majoria de categories s'estructuren al voltant de prototips (Taylor, 1995, 2002; Langacker, 1987; Lakoff, 1987; Geeraerts, 1989; Corft & Cruse, 2004), fet que es tradueix en categories difoses que es poden fusionar una dins l'altra, amb membres més centrals i d'altres més perifèrics, i amb atributs que no són compartits per tots els membres de la categoria. D'aquesta manera, considerem que una possible solució per a resoldre els problemes dels tests aspectuals és abandonar la concepció aristotèlica de les classes i partir de la hipòtesi que les categories aspectuals s'estructuren en base a prototips. Així, s'entén que la funció dels tests aspectuals és postular els membres més centrals de cada categoria i detectar a quin prototip s'assembla més cada verb o, per a ser més precisos, cada sentit verbal. I és que cal tenir present que la majoria de verbs són polisèmics i, de manera relacionada, poden incloure sentits que expressen diferent informació aspectual.

Un bon exemple de verb polisèmic és el verb *contener*. Al nostre entendre, aquest verb té com a mínim dos sentits que difereixen aspectualment, un sentit clarament estatiu ('incloure'), que no accepta contextos dinàmics com la forma progressiva o l'imperatiu (3); i un altre sentit més dinàmic ('frenar l'acció d'un cos') que, contràriament, sí que accepta aquests contextos (4).

- (3) a. *El libro está conteniendo un apartado de ejercicios.*
 b. * *[Libro,] contén un apartado de ejercicios!* (imperatiu)
- (4) a. *El ejército está conteniendo la revolución de los campesinos.*
 b. *[Ejército,] contén la revolución de los campesinos!* (imperatiu)

Si observem aquestes dades comprovem que és més adient adoptar el sentit, i no el verb, com a objecte de classificació aspectual. Així, direm que *contener* és un verb que té com a mínim dos sentits, un de més pròxim al prototip estatiu, i un altre de més pròxim al prototip eventiu. Tot i aquestes evidències de la realitat de la polisèmia aspectual, és sorprenent comprovar com la majoria d'autors que tracten l'Aktionsart no parteixen d'un llistat inicial de sentits a l'hora de classificar el verb en una o altra classe aspectual, sinó que s'adopta la peça lèxica com un tot, i la diferenciació de sentits queda implícita o confusa. De fet, són els recursos computacionals els primers en evidenciar la necessitat de diferenciar i formalitzar de manera sistemàtica els sentits d'un verb. És així com trobem recursos computacionals, com ara EuroWordNet (Vossen et al. 1998) o SenSem (Castellón et al. 2006) en què es parteix d'un llistat inicial de sentits per als verbs polisèmics i en què s'assigna informació aspectual per a cada sentit, no per peça lèxica. Tot i així, hi ha un a manca de criteris clars per a distingir sentits, així com per a definir el propi concepte de 'sentit'. És potser aquesta dificultat metodològica la que ha portat a gran part dels estudis sobre aspecte lèxic a menystenir el fet que un mateix verb pot tenir una lectura més o menys dinàmica en funció del sentit que li assignem lèxicament.

En canvi, és més freqüent trobar la constatació dels efectes del context en els canvis de la interpretació aspectual d'una oració; l'anomenada composicionalitat de l'Aktionsart (Verkuyl 1989, 1993; Mourelatos, 1978; Tenny, 1994; De Miguel, entre d'altres). Així, es diu que un OD definit (p.ex. menjar una poma) o un SP com "fins a X temps / lloc " delimiten una situació o esdeveniment, mentre que un OD indefinit (p.ex. menjar pomes) no ho fa. Per altra banda, s'afirma que els temps verbals també tenen conseqüències en la interpretació aspectual d'una oració: el present i l'imperfet impliquen un aspecte atè-lic, mentre que el passat simple i el pretèrit perfet impliquen un aspecte perfectiu. Ara bé, si observem els exemples de (2) i (3) per al verb *contener*, notem que els verbs polisèmics poden presentar diferències aspectuals més enllà d'aquests "efectes aspectuals del context", en tant que el context és idèntic: a les dues oracions el verb es conjuga en present, l'OD és delimitat i no hi ha modificadors adverbials.

La qüestió que queda pendent de resoldre és la realitat psicològica de l'aspecte lèxic; és a dir, l'observació de si els resultats dels tests aspectuals tenen un correlat en la forma en com els parlants processen cognitivament les categories aspectuals. Per tant, en primer lloc necessitem evidències empíriques del fet que els parlants realment estableixen diferències cognitives entre classes aspectuals i, en segon lloc, que estructurin aquestes categories en base a prototips. Finalment, manca demostrar que existeixen diferències en cost de processament entre els diferents sentits aspectuals d'una peça lèxica verbal, fet que evidenciaria la realitat cognitiva de la polisèmia aspectual. Aquestes seran les qüestions d'estudi fonamentals del present article a partir d'un experiment psicolingüístic.

Adoptarem com a marc de treball l'experiment realitzat per Gennari & Poepple (2003) per a l'anglès. Aquests autors parteixen de la hipòtesi que els verbs eventius, en tant que més complexos semànticament, tarden més temps en ser processats que els verbs estatus. El fonament d'aquesta hipòtesi és que processar el significat d'un verb estatiu requereix l'activació d'una única situació no-dinàmica, mentre que processar el significat d'un esdeveniment dinàmic implica tenir activades més subestructures o subsituacions implicades, com ara estat inicial, canvi i estat final. Així, en termes de processament, s'activen més estructures semàntiques en els verbs eventius. Per tant, el parlant tardarà més en processar el significat. Els resultats de l'experiment dut a terme per aquests autors mostren que, efectivament, en anglès els verbs eventius tarden més temps en ser processats que els verbs estatus. Basant-nos en aquestes resultats, el nostre primer objectiu és esbrinar si aquesta hipòtesi també es pot afirmar per a l'espanyol. Concretament, la formulació de la hipòtesi de partida del nostre experiment és la següent: en espanyol, els prototips de verbs estatus tarden més temps en ser processats que els prototips de verbs eventius, de la mateixa manera que passa en anglès.

De manera relacionada, un altre dels nostres objectius bàsics d'estudi serà observar si existeixen correlats empírics en temps de processament de les diferències aspectuals dels sentits d'un verb polisèmic, així com determinar de quina manera s'actualitzen aquestes diferències en el

processament de l'oració, una qüestió que considerem molt interessant per tal com no ha estat tractada anteriorment de forma experimental. Així, establim una segona hipòtesi segons la qual hi ha verbs polisèmics que lèxicament codifiquen sentits que difereixen aspectualment més enllà dels efectes del "context aspectual" (tipus quantificacional de l'OD, adverbis i temps verbal), unes diferències que es poden detectar en cost de processament: el sentit eventiu d'un verb polisèmic, al ser més complex semànticament, tardarà més temps en ser processat que el sentit estatiu.

3. Experiment

En aquest apartat presentem en detall els paràmetres que s'han tingut en compte per a confeccionar les oracions de la mostra experimental, per tal de copsar-se la seva complexitat i transcendència teòrica (3.1.). Així mateix, s'exposarà breument el disseny de l'experiment i la metodologia seguida per a la seva implementació (3.2.).

3.1. Constitució de la mostra experimental

El nostre objecte d'estudi són 5 verbs polisèmics molt freqüents en espanyol: *cumplir*, *querer*, *contener*, *comprender* i *salir*, cadascun dels quals permet interpretacions estatives i eventives. Partim de la hipòtesi que cada lema codifica com a mínim dos sentits que difereixen aspectualment, un sentit estatiu i un sentit eventiu. A (5) podem observar les definicions per a cada sentit que s'han extret del *Diccionario de la lengua española* (RAE) i del Corpus SenSem (Castellón et al. 2006).

- (5) a. *Cumplir*
Sentit dinàmic: 'Ejecutar, llevar a efecto'.
Sentit estatiu: 'Tener algo o alguien las cualidades requeridas'.
- b. *Contener*
Sentit dinàmic: 'Frenar la acción de un cuerpo'.
Sentit estatiu: 'Encerrar dentro de sí una cosa a otra'.
Sentit estatiu/dinàmic: 'Reprimir un deseo o sentimiento'.
- c. *Comprender*
Sentit dinàmic: 'Entender algo o a alguien'.
Sentit estatiu: 'Incluir en sí algo'.
- d. *Querer*
Sentit dinàmic: 'Amar, tener cariño o inclinación a una persona'.
Sentit estatiu: 'Anhelar, pretender o necesitar algo'.
- e. *Salir*
Sentit dinàmic: 'Desplazarse de dentro a afuera'.
Sentit estatiu: 'Costar una cosa cierta cantidad de dinero'.

Per a cada parell de sentits d'un verb polisèmic s'ha confeccionat un parell corresponent d'oracions experimentals, procurant que estiguin equilibrades pel que fa a les variables susceptibles d'afectar el temps de processament, per tal d'assegurar-nos que les possibles diferències que es puguin trobar estiguin causades únicament pel significat verbal. Aquestes variables són: estructura argumental, longitud de l'oració (mateix nombre de paraules i síl·labes) i freqüència els ítems lèxics. Pel que fa a aquest últim paràmetre s'ha seguit el *Frequency Dictionary of Spanish* (Davies, 2006), que recull les 5000 paraules més comunes de l'espanyol¹. Així mateix, s'han tingut en compte tots aquells factors contextuais susceptibles de modificar la lectura aspectual d'una oració: tots els verbs es conjuguen en el mateix temps verbal (present) i totes les oracions presenten un OD definit i estan exemptes de modificadors adverbials.

Per tal d'exemplificar el procés, presentem tot seguit els criteris que s'han tingut en compte per a elaborar les oracions experimentals corresponents a cadascun dels sentits de *cumplir* (que a partir d'ara es designaran com a "oracions-A"). Podem observar a (6) com les dues oracions són iguals en longitud (5 paraules i 12 síl·labes), les dues són transitives, tenen subjecte explícit i un OD definit, i el temps verbal és present.

(6) Oracions-A per a *cumplir*:

- a. Sentit dinàmic: *El comerciante cumple sus objetivos.*
- b. Sentit estàtic: *El candidato cumple los requisitos.*

Ahora, tots els ítems lèxics de cada oració presenten una freqüència equiparable², tal i com es mostra a la Taula 2 (cf. Annex per a consultar els paràmetres que s'han seguit per a la resta d'oracions-A de l'experiment).

| | | | | | |
|---------------|----|-------------|--------|-----|------------|
| Cumplir_event | El | comerciante | cumple | sus | objetivos |
| Freqüència | 1 | 2442 | 363 | 14 | 2027 |
| Cumplir_estat | El | candidato | cumple | los | requisitos |
| Freqüència | 1 | 2148 | 363 | 20 | 2344 |

Taula 2. Oracions per als dos sentits de *cumplir* i freqüències dels ítems lèxics

¹ El *Frequency Dictionary of Spanish* (Davies, 2006) es basa en un corpus de 20 milions de paraules que inclouen diferents tipus de textos: ficció, no-ficció i converses actuals.

² El número es refereix a la posició de la paraula a la llista de freqüències de (Davies, 2006), per tant, a menor índex, major freqüència. Seguint la metodologia de Gennari & Poepple (2003) s'ha establert que un interval de 400 posicions permet designar ambdues freqüències com a equiparables als efectes de l'experiment.

Per altra banda, per tal de verificar que la possible diferència en temps de processament entre els dos sentits d'un verb es deuen a factors aspectuals, per a cada oració_A elaborem dues oracions exactament iguals canviant només el verb principal: un verb prototípicament estatiu i un verb prototípicament eventiu. La finalitat és comprovar si cada oració_A s'acosta més en temps de processament al seu respectiu prototip estatiu (oració_B) o al prototip eventiu (oració_C). A continuació (7) podem veure les oracions B i C per als dos sentits del verb *cumplir*.

(7) Sentit dinàmic:

Oració_A: El comerciante *cumple* sus objetivos

Oració_B: El comerciante *logra* sus objetivos → esdeveniment

Oració_C: El comerciante *cree* en sus objetivos → estat

Sentit estatiu:

Oració_A: El candidato *cumple* los requisitos

Oració_B: El candidato *lee* los requisitos → esdeveniment

Oració_C: El candidato *sabe* los requisitos → estat

Pel que fa a la tria dels verbs prototípicament estatus i prototípicament eventius, s'han tingut en compte diversos factors. En primer lloc, s'ha comprovat a partir de corpus el grau de dinamicitat dels verbs escollits. S'han agafat els 5 criteris més freqüents que es citen a la bibliografia per a determinar la dinamicitat d'un predicat (forma progressiva, "después de haber + v", "obligar + v", interpretació habitual i adverbial *poco a poco*), i s'ha observat quins contextos accepta cada verb, i amb quina freqüència, al *Corpus de Referencia del Español Actual* (CREA). Els verbs que accepten més contextos s'han escollit com a prototips eventius i els verbs que rebutgen més contextos s'han escollit com a prototips estatus. Alhora, s'han tingut en compte els següents paràmetres: el prototip eventiu / estatiu ha de tenir el mateix nombre de síl·labes i una freqüència similar al verb de l'oració_A; ha de ser un verb monosèmic o, si més no, que tingui sentits que pertanyin a la mateixa classe aspectual (és a dir, que siguin monosèmics aspectualment parlant); a més, l'oració resultant ha de ser gramatical i semànticament coherent.

Finalment, per tal de verificar si en espanyol els verbs eventius tarden més que els verbs estatus en ser processats, tal i com han demostrat Gennari & Poepple per a l'anglès, compararem el temps de processament dels verbs de B i C. Per tant, tindrem tres nivells d'anàlisi de les dades:

- Es comparen els dos sentits aspectuals d'un verb polisèmic (les dues oracions A)
- Es compara cada sentit aspectual d'un verb polisèmic respecte al seus prototips estatus i eventius (cada oració A respecte a B i C).
- Es comparen verbs estatus respecte a verbs eventius (verbs de B i C).

3.2. Disseny i procediment

S'ha realitzat un experiment d'auto-ritme de lectura implementat amb el programa informàtic E-Prime, v.1.1. (Shneider et al. 2002), amb l'objectiu d'evidenciar la diferència en temps de processament entre esdeveniments i estats. Pel que fa al software utilitzat, l'E-Prime, és una de les aplicacions més usades actualment per a generar experiments en laboratoris i centres de recerca en neurociència i psicolingüística. Concretament, l'E-Prime permet recollir i processar dades lingüístiques amb una precisió de mil·lisegons, fet que el converteix en una de les eines bàsiques per a realitzar experiments basats en temps de reacció, com és el cas de les tasques de decisió lèxica o l'experiment d'auto-ritme de lectura que presentem en aquest article.

En quant a la implementació específica de l'experiment, tenim un total de 33 oracions, ordenades de forma arbitrària, que es presenten al subjecte dividides per constituent sintàctic (subjecte, verb i oració) al centre de la pantalla³. Els participants han de clicar Espai per anar llegint cada constituent, i el programa enregistra amb una precisió de mil·lisegons quant de temps tarden en clicar. Després de cada 5 oracions se'ls presenta una pregunta de comprensió, per tal que prestin atenció a l'hora de llegir. Els subjectes no coneixen l'objecte de l'experiment, sinó que es creuen que està relacionat amb aquestes preguntes de comprensió. S'ha aplicat l'experiment a 20 parlants d'espanyol, treballadors de la Universitat Oberta de Catalunya.

4. Resultats

En termes generals, els resultats de l'experiment verifiquen la primera hipòtesi segons la qual els verbs eventius, en tant que més complexos semànticament, tarden més temps en ser processats que els verbs estatus. Concretament, trobem que en un 81% dels casos analitzats els esdeveniments tarden més en ser processats que els estats, amb una diferència de 277 ms de mitjana. En quant a la segona hipòtesi, els resultats també apunten al fet que les diferències en cost de processament no només es donen entre verbs, sinó també entre sentits aspectualment divergents d'un verb polisèmic. Concretament, del total de verbs polisèmics analitzats, en un 80% dels casos el sentit eventiu tarda més temps que l'estatiu en ser processat, més enllà dels anomenats efectes del context aspectual. Tanmateix, considerem que és necessari delimitar amb més precisió quin és el paper de l'aspecte i quin el de la freqüència i la prototipicitat de sentits en el cost de processament, per tal de poder extreure conclusions més sòlides pel que fa a les diferències en temps de processament entre els diferents sentits aspectuals d'un verb polisèmic.

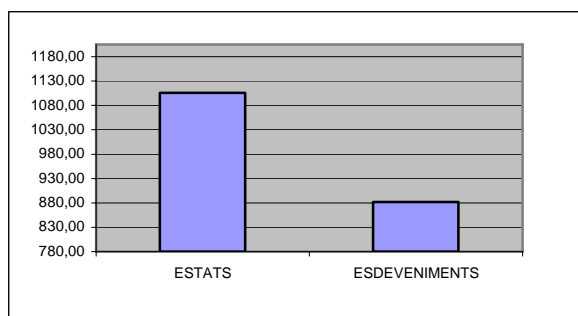
Per altra banda, és important apuntar que en aquells casos en què no trobem una diferència significativa en cost de processament entre estats i esdeveniments, un dels dos verbs o sentits objecte de comparació pertany a l'anomenada classe dels verbs psicològics (*respetar*, *crear*, sentit psicològic de *comprender* i *contener*), fet que ens porta a concloure que els verbs psicològics o de

sensació no són prototips, ni estatus ni eventius, una qüestió que creiem convenient analitzar més a fons i de manera específica en futurs experiments.

Tot seguit presentarem en detall els resultats per a la primera hipòtesi en tant que és el fonament de l'experiment que s'ha portat a terme (4.1). Per altra banda, per tal d'exemplificar el mètode d'avaluació que s'ha seguit per a verificar la segona hipòtesi, ens centrarem en l'anàlisi multinivell dels resultats del verb polisèmic *cumplir* (4.2).

4.1. Diferència genèrica en temps de processament entre verbs eventius i verbs estatus.

Es disposa d'un total de 11 parells de verbs prototípicament estatus-eventius, que corresponen a les oracions B i C de cadascun dels verbs polisèmics estudiats: *buscar-crear*, *pedir-tener*, *publicar-incluir*, *iniciar-incluir*, *subir-valer*, *expresar-carecer*, *llegar-estar*, *provocar-respetar*, *escuchar-respetar*, *leer-saber* i *lograr-crear*. Si analitzem les dades de forma global, el resultat és que els verbs eventius tarden una mitjana de 224,13 ms més que els estatus en ser processats (cf. Gràfica 1).



Gràfica 1. Diferència en temps de processament entre estatus i esdeveniments

Si analitzem les dades amb més detall, trobem que són 9 els parells en què els verbs eventius presenten un major cost de processament que els verbs estatus (és a dir, en un 81 % dels casos), mentre que en dos dels parells no s'aprecia una diferència rellevant. Allò que ens interessa subratllar és que en aquests dos últims parells el verb eventiu és el mateix, *respetar*, fet que ens fa pensar que per alguna raó que ara no entrarem a analitzar, *respetar* no és un bon exemple de prototip estatus.

En els parells verbals en què la diferència entre esdeveniments i estatus ocorre, aquesta diferència és significativa estadísticament, com bé revela l'aplicació del *test* ANOVA, que es recull a la Taula 3.

³ El fet de dividir les oracions ens permetrà analitzar les dades a diferents nivells: el temps de processament

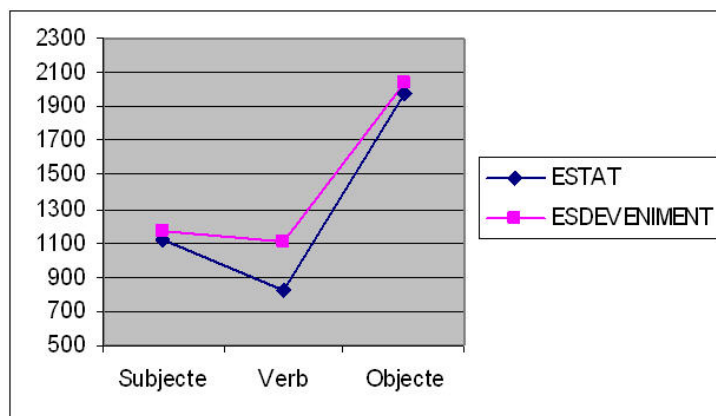
| Parells de Verbs | F | Valor crític per a F | P |
|------------------|--------|----------------------|------|
| Busca-Cree | 4,3389 | 4,0981 | 0,04 |
| Publica-Incluye | 4,7998 | 4,0981 | 0,03 |
| Inicia-Incluye | 4,2259 | 4,0981 | 0,04 |
| Sube-Vale | 4,3622 | 4,0981 | 0,04 |
| Llega-Está | 4,2521 | 4,0981 | 0,04 |
| Lee-Sabe | 4,2301 | 4,0981 | 0,04 |
| Expresa-carece | 2,9963 | 2,8424 | 0,09 |
| Pide-Tiene | 3,0247 | 2,8424 | 0,09 |
| Logra-Cree | 0,0052 | 2,8424 | 0,9 |

Taula 3. ANOVA per als parells de verbs en què hi ha diferència entre estats i esdeveniments.

Concretament, podem observar que en 6 casos (*busca-cree*, *publica-incluye*, *inicia-incluye*, *sube-vale*, *llega-está*, *lee-sabe*) la diferència entre esdeveniments i estats és significativa al 0'05 % de confiança (és a dir, que F és més gran que el valor crític a un nivell de probabilitat del 0'05%) i en 2 casos (*expresa-carece*, *pide-tiene*) la diferència és significativa a un 0'1% de confiança. L'única excepció és el parell de verbs *logra-cree*: tot i que el verb eventiu tarda 12 ms més que el verb estatiu en ser processat, aquesta diferència és de l'ordre de l'1% i, per tant, no podem saber si els resultats es deuen, o no, a l'atzar. Aquestes dades ens porten a pensar que els verbs prototípicament no-duratus (assoliments), com és el cas de *lograr*, presenten diferències en cost de processament respecte als verbs eventius duratus (realitzacions, activitats i estats), fet que intuïtivament sembla raonable. És per aquest motiu que considerem convenient experimentar, en un futur immediat, amb més subclasses aspectuals i observar si existeixen diferències en temps de processament entre esdeveniments duratus i no-duratus, així com entre esdeveniments delimitats i no-delimitats.

Per altra banda, és interessant comprovar (veure Gràfica 2) que en aquells casos en què el verb eventiu tarda més que l'estatiu en ser processat (tots els casos excepte els parells amb *respetar*), trobem diferències significatives en temps de processament entre verbs (mitjana de 277,04 ms) però no entre subjectes i objectes exactament iguals. Per tant, en tant que s'han controlat les possibles variables que poden afectar el temps de processament, podem concloure que les diferències que trobem a la posició verbal es deuen al significat aspectual de les peça lèxica verbal.

de cada constituent o del total de l'oració.



Gràfica 2. Temps de processament per constituent en la diferència entre estats i esdeveniments

En definitiva, segons els resultats de l'experiment, podem verificar la hipòtesi segons la qual els verbs eventius tarden més temps en ser processats que els verbs estatus en espanyol, com han demostrat Gennari & Poeple (2003) per a l'anglès. Amb l'objectiu de confirmar aquesta hipòtesi, en un futur immediat està previst estendre l'experiment a més parells de verbs.

4.2. Tres nivells d'anàlisi dels resultats. El cas de *cumplir*

Per tal d'exemplificar la metodologia seguida en l'anàlisi de les dades, presentarem en detall els resultats de l'experiment per al verb *cumplir*. Com bé s'ha apuntat anteriorment, el procediment d'anàlisi s'ha dut a terme a tres nivells diferents:

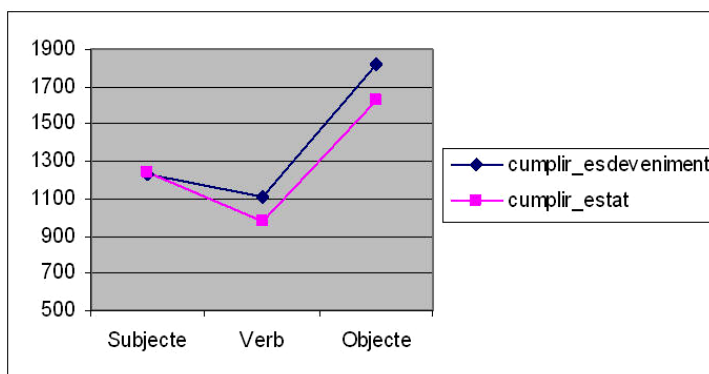
- 1) En primer lloc observarem si existeixen diferències en cost de processament entre el sentit estatiu i el sentit dinàmic del verb polisèmic *cumplir*, tot comparant entre sí el cost de processament de les dues oracions_A (cf. apartat 4.1.1.).
- 2) Acte seguit comprovarem si els prototips estatiu / eventiu que s'han escollit per a cada sentit (és a dir, els verbs de les oracions B i C) realment difereixen en cost de processament; és a dir, cal verificar que són bons exemples de les categories aspectuals objecte d'estudi. En el cas de *cumplir* s'han escollit com a prototips els parells de verbs *leer-saber* i *creer-lograr* (cf. 4.1.2.).
- 3) Finalment, l'última qüestió que volem observar és si el temps de processament de cada sentit de *cumplir* s'acosta més al seu respectiu prototip estatiu o eventiu, per tal d'evidenciar la valència aspectual de cada sentit. Per tant, es compara cada oració A amb les seves respectives oracions B i C (cf. 4.1.3.).

4.2.1. Diferències en cost de processament entre els dos sentits del verb *cumplir*

Per als dos sentits aspectualment diferents dels verb *cumplir*, que s'han experimentat a partir de les oracions_A que es recullen a (8), trobem que l'accepció eventiva presenta un cost de

processament més gran que l'estativa, tal i com s'havia hipotetitzat. En podem veure els resultats il·lustrats a la Gràfica 3.

- (8) a. El comerciante cumple sus objetivos. (sentido eventivo)
b. El candidato cumple los requisitos. (sentido estativo)



Gràfica 3. Diferències de processament de *cumplir*

La dada que cal subratllar és que la diferència en el temps de processament d'aquestes dues oracions es manifesta tant en el verb com en l'objecte directe, però no en el subjecte. Concretament, la mitjana de la diferència entre subjectes és només de 6 ms, mentre que entre verbs i objectes és de 135ms i 192ms, respectivament. Així, el fet que el sentit eventiu tardi més que l'estatiu a la posició verbal, implica que cada sentit del verb *cumplir* ja es comença a actualitzar en la posició verbal, després d'haver llegit el subjecte.

4.2.2. Els prototips. El parells *lee-sabe* i *logra-cree*

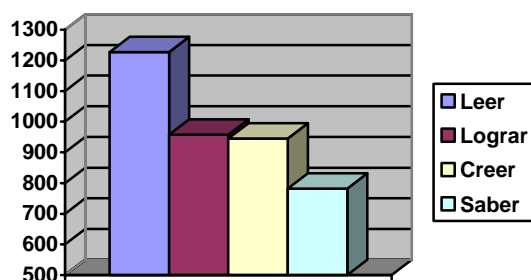
Per tal de verificar que les diferències en cost de processament detectades entre els dos sentits del verb *cumplir* es deuen a factor aspectuals, s'han elaborat dues oracions exactament iguals canviant només el verb principal per un verb prototípicament eventiu i un altre de prototípicament estatiu (seguint els paràmetres exposats a l'apartat 3.1. pel que fa a la tria dels prototips). L'objectiu final és comparar el cost de processament de cada sentit amb el cost dels seus respectius prototips aspectuals. En primer lloc, per tant, serà necessari comprovar que els verbs escollits com a prototips realment difereixen en temps de processament i, d'aquesta manera, poder corroborar que realment són membres centrals de cada categoria aspectual.

El prototips aspectuals (verbs de les oracions B i C) escollits per al sentit estatiu de *cumplir* són el verb *leer* (esdeveniment) i el verb *saber* (estat). Les dades resultants de l'experiment mostren que, efectivament, el prototip eventiu *lee* tarda una mitjana de 445ms més en ser processat que el prototip estatiu *sabe*, una diferència que es revela estadísticament significativa si apliquem el test de l'ANOVA; concretament, la diferència entre *lee* i *sabe* és significativa al 0'04% de confiança, és

a dir, que la probabilitat a l'hora de verificar que els resultats no es deuen a l'atzar és del 96 % de confiança.

Pel que fa als prototips del sentit dinàmic del verb *cumplir*, tenim el parell de verbs *logra* (event)-*cree* (estat). En aquest cas també trobem que el verb eventiu tarda més que l'estatiu en ser processat; tanmateix, la diferència en cost de processament entre aquests dos verbs és només de 12 ms. Alhora, si apliquem el test estadístic de l'ANOVA observem que en aquest cas la diferència és de l'ordre de l'1% de confiança; per tant, una diferència que no es revela significativa estadísticament. Aquestes dades ens porten a pensar que els verbs prototípicament no-duratius, com ho és clarament el verb *lograr* (de fet, la classe dels esdeveniments puntuals es designa així: *logros*), presenten diferències en cost de processament respecte als verbs eventius duratius (activitats i realitzacions), fet que intuïtivament sembla raonable. Alhora, el verb que hem escollit com a prototip estatiu, *crear*, es podria encabir dins la classe dels anomenats verbs psicològics que no només són de dubtosa classificació aspectual des del punt de vista teòric (Grimshaw, 1990; Pustejovsky, 1988; Tenny, 1994; Marín & McNally, 2005; entre d'altres), sinó que també es mostren com a no-prototips pel que fa al cost de processament, segons els dades del nostre experiment.

Per tant, els resultats semblen indicar de forma bastant clara que no és adequat parlar de categories aspectuals amb límits clarament definits, sinó que és més apropiat establir una gradació entre els paràmetres prototípics de la dinamicitat i l'estativitat. Així, sembla que la classe dels assoliments (verbs puntuals) i la classe dels verbs psicològics es situarien a mig camí entre el prototip estatiu i l'eventiu, com es pot comprovar a partir de les diferències en temps de processament dels quatre verbs analitzats en aquesta secció: *leer*, *lograr*, *crear* i *saber*, de més a menys eventiu, respectivament (com s'il·lustra a la Gràfica 4).



Gràfica 4. Temps de processament en ms per a *leer*, *lograr*, *crear* i *saber*

4.2.3. A quin prototip s'acosta més cada sentit del verb *cumplir*?

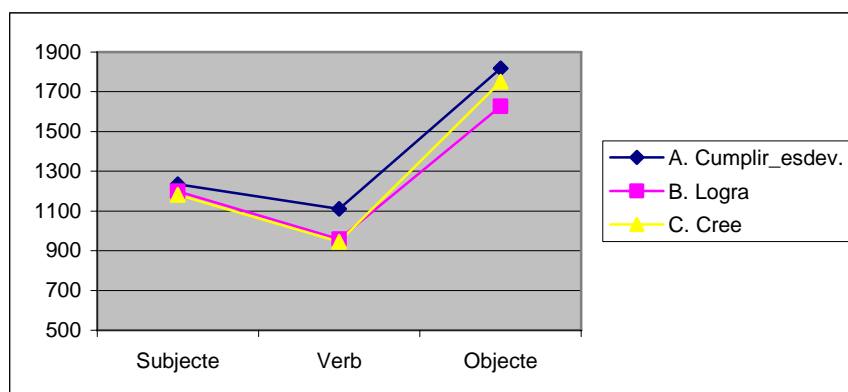
Després d'haver verificat que la diferència aspectual entre els dos sentits del verb *cumplir* té un correlat en temps de lectura (cf. 4.1.1.), i un cop comprovat el fet que també existeixen diferències

en cost de processament entre verbs més eventius i més estatus (cf. 4.1.2.), ens resta comprovar si el temps de processament de cada sentit de *cumplir* es pot equiparar als seus respectius prototips aspectuals; és a dir, verificar que el sentit que hem hipotetitzat com a dinàmic ('ejecutar, llevar a efecto') s'acosta més al seu respectiu prototip eventiu (*logra*), que a l'estatiu (*cree*); i, de la mateixa manera, que el sentit que hem hipotetitzat com a menys dinàmic ('tener algo o alguien las cualidades requeridas') s'acosta més al seu respectiu prototip estatiu (*sabe*) que a l'eventiu (*lee*). Tot seguit n'analitzarem els resultats.

Pel que fa al sentit dinàmic de *cumplir*, que hem experimentat a partir de l'oració_A que es recull a (9.a.), els resultats mostren que tarda més en ser processat que els seus respectius prototips eventiu i estatiu, que s'han experimentat a partir de les oracions B (esdeveniment) i C (estat) que es recullen a (9.b.) i (9.c.).

- (9) a. Oració_A: El comerciante *cumple* sus objetivos
 b. Oració_B: El comerciante *logra* sus objetivos → esdeveniment
 c. Oració_C: El comerciante *cree* en sus objetivos → estat

Les dades són coherents si tenim en compte que *logra*, que s'havia establert inicialment com a prototip eventiu, és un verb no-duratiu i per tant menys prototípicament eventiu que el sentit dinàmic de *cumplir*. Concretament, com es pot observar a la Gràfica 5., trobem que a la posició verbal, l'oració A tarda una mitjana de 165 ms més en ser processada que l'oració B (*logra*) i 153 ms més que l'oració C (*cree*); pel que fa a la posició d'objecte directe, l'oració A tarda una mitjana de 168 ms més en ser processada que l'oració C (*logra*), i 292 ms més que l'oració C (*cree*)⁴. Per tant, el sentit dinàmic de *cumplir* es situa a l'extrem màximament eventiu pel que fa al temps de processament.



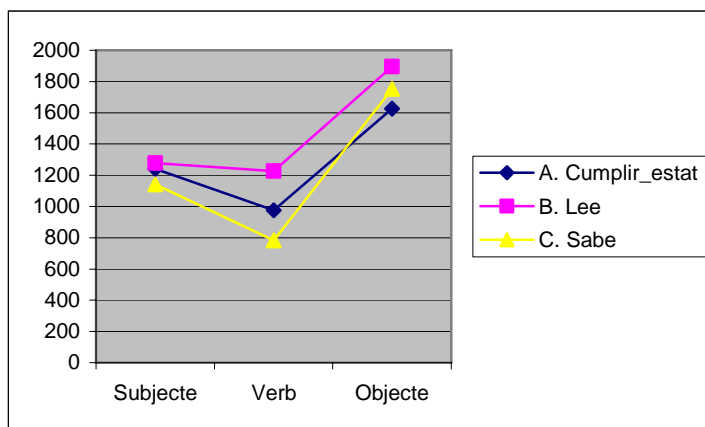
Gràfica 5. Sentit dinàmic de *cumplir* respecte als seus respectius prototips

⁴ Cal notar que l'objecte directe tarda més en ser processat en el prototip estatiu (*cree*) que en l'eventiu (*logra*). Atribuïm aquesta diferència en temps de processament al fet que l'oració C presenta una paraula de més, la preposició "en", requerida pel verb perquè l'oració sigui gramatical.

En quant al sentit estatiu de *cumplir*, l'oració experimental del qual es pot observar a (10.a.), trobem que es situa a mig camí respecte als seus respectius prototips eventiu (10.b.) i estatiu (10.c), pel que fa al cost de processament.

- (10) a. Oració_A: El candidato *cumple* los requisitos
 b. Oració_B: El candidato *lee* los requisitos → esdeveniment
 c. Oració_C: El candidato *sabe* los requisitos → estat

Tot i així, cal indicar que l'oració A s'acosta més en temps de processament al prototip estatiu (oració C), tal i com s'havia predit, tant a la posició verbal (59 ms més propera a l'estat que a l'esdeveniment) com pel que fa a l'objecte directe (144 ms més propera a l'estat que a l'esdeveniment), com es pot observar a la Gràfica 6.



Gràfica 6. Sentit estatiu de *cumplir* respecte als seus respectius prototips

5. Conclusions

En aquest article hem presentant un experiment psicolingüístic sobre el processament de la dinamicitat en las unitats verbals de l'espanyol, que parteix del propòsit fonamental de validar empíricament la macrodistinció aspectual entre estats i esdeveniments, així com la relació entre aspecte i polisèmia verbal. L'objectiu final d'aquest experiment és obtenir un model formal de representació de les categories aspectuals que no només sigui implementable computacionalment sinó que també, i sobretot, estigui motivat cognitivament.

A partir de l'aplicació de l'experiment s'ha presentat una primera anàlisi de les dades extretes. En síntesi, les dades resultants de l'experiment verifiquen una primera hipòtesi segons la qual, en espanyol, els verbs eventius tarden més temps en ser processats que els verbs estatus, tot reafirmant allò que han demostrat Gennari & Poepple (2003) per a l'anglès. De manera

relacionada, una altra de les conclusions principals que podem extreure de les dades obtingudes és que el paràmetre de la dinamicitat, a partir del qual es basteix la macrodistinció aspectual entre estats i esdeveniments, és un paràmetre gradual que s'estructura en base a prototips, amb membres que són més bons exemples que d'altres de la categoria, un fet que hem pogut constatar no només a partir de l'aplicació dels *tests* aspectuals, sinó també de forma experimental a partir de l'obtenció d'un correlat empíric en temps de processament: els verbs que s'havien elegit com a prototips eventius realment tarden més temps en ser processats que els verbs classificats inicialment com a prototips estatus; resultats que havíem predit en la hipòtesi de partida de l'experiment. Per tant, s'ha obtingut un correlat empíric de com els parlants representen i processen aquestes dues macrocategories aspectuals.

Per altra banda, hem pogut comprovar que les diferències en temps de processament no només es donen entre verbs monosèmics aspectualment parlant, sinó que també les trobem entre sentits de verbs polisèmics que divergeixen aspectualment, com bé s'ha il·lustrat en detall amb els dos sentits del verb *cumplir*. De confirmar-se aquest fet en futurs experiments, es podria dir tant que els sentits verbals són una realitat psicològica, com que l'aspecte lèxic té un paper clau a l'hora d'establir-los i delimitar-los. Alhora, de confirmar-se, aquests resultats tindrien un gran valor tant per a la desambiguació de sentits com per a la classificació verbal en les bases de coneixement lèxiques.

Finalment, la complexitat de les dades ha fet emergir tota una sèrie de factors que queden pendents d'anàlisi i experimentació per a un treball futur. Com bé hem fet notar a partir de l'anàlisi dels resultats del parell de verbs *logra-cree*, les dades apunten al fet que existeixen diferències en cost de processament entre esdeveniments puntuals (assoliments) i esdeveniments duratius (activitats i realitzacions), fet que serà objecte d'estudi específic en un futur experiment amb un nombre més gran d'assoliments. Per altra banda, els resultats de l'experiment mostren que els verbs o sentits verbals psicològics no es comporten com a prototips, ni eventius ni estatus, pel que fa al temps de processament. En conseqüència, creiem convenient expandir l'experiment a un nombre més gran de verbs psicològics per tal de comprovar que es manté la coherència en temps de processament per a aquesta classe semàntica verbal.

Agraïments

Aquest treball ha estat subvencionat amb els ajuts de l'Internet Interdisciplinary Institute (Universitat Oberta de Catalunya) i el projecte KNOW (Desarrollo de tecnologías multilingües a gran escala para la comprensión del lenguaje) del Ministerio de Educación y Ciencia (TIN2006-1549-C03-02).

Referències bibliogràfiques

- Aguiree, E & P. Edmons (2006). *Word Sense Disambiguation: Algorithms and Applications*. Ed. Springer. Series: Text, Speech and Language Technology , Vol. 33
- Boroditsky, L. (2000). 'Metaphoric Structuring: Understanding time through spatial metaphors'. *Cognition*, 75(1), 1-28.
- Castellón, I., A. Fernández, G. Vázquez, L. Alonso, J.A. Capilla (2006). 'The Sensem Corpus: a Corpus Annotated at the Syntactic and Semantic Level', *Fifth International Conference on Language Resources and Evaluation (LREC)*.
- Cuycken, Hubert, Dominiek & Sally rice (1997). 'Towards an empirical lexical semantics', dins B. Smieja and M. Tasch (eds.). *Human Contact Through Language and Linguistics*. Frankfurt: Peter Lang, pp. 35-54.
- Croft, W. & A. Cruse (2004). *Cognitive Linguistics*. Cambridge University Press.
- Davies, M. (2006). *A Frequency dictionary of Spanish: core vocabulary for learners*. London: Routledge.
- De Miguel, E. (1999). 'El aspecto léxico', dins *Gramática Descriptiva de la Lengua Española*, eds. I. Bosque y V. Demonte, 2971-3060. Madrid: Espasa Calpe.
- (2004). 'Qué significan aspectualmente algunos verbos y qué pueden llegar a significar', *Estudios de lingüística*, núm. 1, 167-106.
- Dowty, D. (1979). *Word Meaning and Montague Grammar*. Reidel: Dordrecht.
- Evans, V., & Green, M. (2006). *Cognitive Linguistics: An Introduction*. Mahwah, NJ and Edinburgh: Lawrence Erlbaum Associates/Edinburgh University Press.
- García-Miguel, J. M. & F. Albertuz (2005). 'Verbs, Semantic Classes and Semantic Roles in the ADESSE Project', *Interdisciplinary Workshop on Verb Features and Verb Classes*. Saarbrücken.
- Gennari, S., D. Poepple (2003). 'Processing correlates of lexical semantic complexity', *Cognition*, 89.
- Geeraerts, D. (1989). 'Introduction: Prospects and problems of prototype theory'. *Linguistics*, 27: 587-612.
- González-Márquez, M., I. Mittelberg, S. Coulson, & M. Spivey (en premsa). *Empirical Methods in Cognitive Linguistics*. Ithaca: John Benjamins.
- Gries, S. (2005). 'The many senses of *run*'. S. Gries and A. Stefanowitsch (eds.), *Corpora in Cognitive Linguistics*. Berlin: Mouton de Gruyter
- Grimshaw, J. (1990). *Argument Structure*, MIT Press: Cambridge.
- Hovy, E.H., M. Marcus, M. Palmer, S. Pradhan, L. Ramshaw, and R. Weischedel (2006). 'OntoNotes: The 90% Solution. Short paper'. *Proceedings of the Human Language Technology / North American Association of Computational Linguistics conference (HLT-NAACL 2006)*. New York, NY.

- Johnson, M. (1987). *The body in the mind. The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Johnson C. & C. J. Fillmore (2000). 'The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure'. *Proceedings of the Conference on Applied Natural Language Processing*, p.56-62.
- Kingsbury P. & M. Palmer (2002). 'Adding Semantic Anotation to the Penn Treebank'. *Proceedings of the Human Language Technology Conference*.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lavid, J. (2005). *Lenguaje y nuevas tecnologías. Nuevas perspectivas, métodos y herramientas para el lingüista del siglo XXI*. Madrid. Ed. Cátedra
- Levin B. & M. Rappaport Hovav (1995). *Unaccusativity at the Syntax-Semantics Interface*. Cambridge MA:MIT Press.
- Marín, R. (2000). *El componente aspectual de la predicación*. Tesis Doctoral, Universitat Autònoma de Barcelona.
- Marín, R. & L. McNally (2005). 'The Aktionsart of Spanish reflexive psychological verbs', dins Emar Maier, Corien Bary & Janneke Huitink, eds. (2005), *Proceedings of SuB9*.
- Morimoto, Y. (1998). 'El Aspecto léxico: delimitación', *Cuadernos de lengua española*, Ed. Arco/Libros, S.L., Madrid.
- Mourelatos, A. (1978). 'Events, Processes and States'. *Linguistics and Philosophy* 2:415-434.
- REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual*. <<http://www.rae.es>>
- REAL ACADEMIA ESPAÑOLA: *Diccionario de la lengua española* [en línea]. <<http://www.rae.es>>
- Pustejovsky, J. (1988). 'The geometry of events', en C.Tenny (ed.), *Studies in Generative Approches to Aspect and Lexicon*, Project Working Papers 24, Center for Cognitive Science al MIT, Cambridge.
- (1991). 'The Syntax of Event Structure', a B. Levin i S. Pinker (eds), *Lexical and Conceptual Structure*, Oxford, Blackwell, p. 47-81.
- Resnik, P. (2004). "Exploiting hidden meanings: Using bilingual text for monolingual Annotation". *Lecture Notes in Computer Science 2945: Computational Linguistics and Intelligent Text Processing*, ed. by Alexander Gelbukh, 283- 299. New York: Springer-Verlag.
- Sasse, Hans-Jürgen (2003). 'Recent Activity in the Theory of Aspect: Accomplishments, Achievements, or Just Non-Progressive State?'. *Linguistic Typology* 6.199-271. Review article.
- Schneider, W., A. Eschman, and A. Zuccolotto (2002). *E-Prime User's Guide*. Pittsburg: Psychology Software Tools Inc.
- Smith, C. (1991). *The parameter of aspect*. Kluwer Academic Press.
- Talmy, L. (1985). "Patterns of lexicalization", dins Shopen, T. (ed.): *Language typology and*

- syntactic description. Grammatical categories and the lexicon*. Vol. III, Cambridge University Press.
- (2000). *Toward a cognitive semantics*. 2 vols. Cambridge, MA: Massachusetts Institute of Technology.
- Taylor, J.R. (1995). *Linguistic categorization*, Oxford University Press, 2n ed.
- (2002). *Cognitive Grammar*, Oxford: Oxford University Press.
- Tenny, C. (1994). *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer Academic Publishers.
- Tyler, A. & V. Evans (2001). 'Reconsidering Prepositional Polysemy Networks: The Case of *Over*'. *Language*, 77, 4, 724-765
- Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, N.Y.: Cornell University Press.
- Verkuyl, H. (1989). 'Aspectual classes and aspectual composition'. *Linguistics and Philosophy* 12:39-94.
- (1993). *A theory of aspectuality. The interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press.
- Vossen, P. (1998). *EuroWordNet: A multilingual database with lexical semantic networks*. Dordrecht: Kluwer Academic Publishers.

ANNEX

Paràmetres per a la confecció de totes les oracions-A de la mostra experimental

Llistat de paràmetres:

- La mateixa estructura sintàctica
- El mateix nombre de paraules i síl·labes
- Arguments que tinguin una freqüència equiparable de tots els ítems lèxics (400 posicions de diferència al Diccionari de Mark Davies, 2006)
- El mateix context aspectual (verb present, tipus OD, sense modificadors adverbials).

1. COMPRENDER

| SENTIT 1 | El proyecto comprende | | una fase de experimentos | | | | |
|---------------------------|-----------------------|-----|------------------------------|----|------|-----|------|
| Freqüència | 1 | 604 | 306 | 56 | 2310 | 2 | 2815 |
| Síl·labes per constituent | 4 | | 3 | 9 | | | |
| Paraules | 2 | | 1 | 4 | | | |
| SENTIT 2 | El público comprende | | los nervios del protagonista | | | | |
| Freqüència | 1 | 769 | 306 | 1 | 2704 | 2+1 | 2536 |
| Síl·labes per constituent | 4 | | 3 | 9 | | | |
| Paraules | 2 | | 1 | 4 | | | |

2. QUERER

| SENTIT 1 | El maestro quiere | | a la mujer del obrero | | | | | |
|---------------------------|-------------------|--|--------------------------|----|----|-----|-----|------|
| Freqüència | | | 57 | 5 | 7 | 127 | 2+1 | 2243 |
| Síl·labes per constituent | 4 | | 2 | 8* | | | | |
| Paraules | 2 | | 1 | 5 | | | | |
| SENTIT 2 | El maestro quiere | | las dos obras del pintor | | | | | |
| Freqüència | | | 57 | 1 | 56 | 206 | 2+1 | 2218 |
| Síl·labes per constituent | 4 | | 2 | 7 | | | | |
| Paraules | 2 | | 1 | 5 | | | | |

* Tot i que hi ha una síl·laba de més, el nombre de paraules és el mateix. A més a més, cal tenir present que l'OD que té 8 síl·labes, té menys lletres (17 lletres) que l'OD que té 7 síl·labes (20 lletres). Per tant, la diferència en longitud també queda compensada.

3. CONTENER

| SENTIT 1 | La revista | contiene | dos artículos | de ética |
|---------------------------|--------------|----------|-------------------|---------------|
| Freqüència | 1 804 | 829 | 56 819 | 2 3685 |
| Síl·labes per constituent | 4 | 3 | 9 | |
| Paraules | 2 | 1 | 4 | |
| SENTIT 2 | La población | contiene | los sentimientos | de venganza |
| Freqüència | 1 900 | 829 | 20 867 | 2 3450 |
| Síl·labes per constituent | 4 | 3 | 9 | |
| Paraules | 2 | 1 | 4 | |
| SENTIT 3 | La policía | contiene | a los estudiantes | de traducción |
| Freqüència | 1 1017 | 829 | 5 20 1019 | 2 3344 |
| Síl·labes per constituent | 5 * | 3 | 0** | |
| Paraules | 2 | 3 | 5** | |

* Tot i tenir una síl·laba de més (respecte a *revista* i *población* que en tenen 4), en nombre de lletres (7 lletres) és igual que *revista* i inferior que *población* (9 lletres); per tant, la diferència en longitud queda compensada.

**El fet de tenir una paraula de més pot ser el motiu pel qual aquest sentit tarda més? En principi la diferència no es deu a la presència de "a" perquè és una paraula d'una sola lletra i és molt freqüent (de fet, si sumem les freqüències de *a* + *los* (5+20) és més freqüent que *dos* (56)).

4. SALIR

| SENTIT 1 | El preso | sale | a la calle |
|---------------------------|----------|------|--------------|
| Freqüència | 1 3069 | 111 | 5 1 248 |
| Síl·labes per constituent | 3 | 2 | 4 |
| Paraules | 2 | 2 | 3 |
| SENTIT 2 | El kilo | sale | a dos euros* |
| Freqüència | 1 2971 | 111 | 5 56 ? |
| Síl·labes per constituent | 3 | 2 | 4 |
| Paraules | 2 | 2 | 3 |

*Euros no surt al diccionari de freqüències, però sí *dólares*. No obstant, considerem que *euros* és una paraula molt freqüent.

5. CUMPLIR

| SENTIT 1 | El comerciante | cumple | sus objetivos |
|---------------------------|-----------------------|---------------|-----------------------|
| Freqüència | 1 2442 | 363 | 14 2027 |
| Síl·labes per constituent | 5 | 2 | 5 |
| Paraules | 2 | 1 | 2 |
| SENTIT 2 | El candidato | cumple | los requisitos |
| Freqüència | 1 2148 | 363 | 20 2344 |
| Síl·labes per constituent | 5 | 2 | 5 |
| Paraules | 2 | 1 | 2 |



IN3 Internet
Interdisciplinary
Institute

www.uoc.edu