

GRIAL

UAB
Universitat Autònoma
de Barcelona

U
UNIVERSITAT DE BARCELONA
B

U
Universitat de Lleida

UOC
www.uoc.edu

Grup de Recerca Interuniversitari en Aplicacions Lingüístiques



Jornada del Processament Computacional del Català

26 DE MARÇ DE 2009, PALAU ROBERT (BARCELONA)

Línies de recerca (I)

- *Desenvolupament, avaluació i aplicació de recursos lingüístics de gran cobertura:*
 - *Gramàtiques:*
 - *Dependències*
 - *HPSG.*
 - *Lèxics:*
 - *Bases de dades verbals amb informació sintàctico-semàntica*
 - *WordNet*
 - *Perífrasis verbals*



Línies de recerca (I)

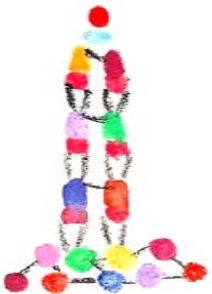
– *Corpus: anotació i explotació*

- *A nivell de mot:*
 - *categoria morfo-sintàctica*
 - *sentit*
- *A nivell de complements:*
 - *rols semàntics*
 - *categoria morfo-sintàctica*
 - *funció sintàctica*
- *A nivell d'oració:*
 - *semàntica de la construcció*
 - *aspectualitat*
 - *modalitat*



Línies de recerca (I)

- *Adquisició d'informació:*
 - *Esquemes de subcategorització*
 - *Restriccions de selecció*
- *Extracció de terminologia*
- *Traducció automàtica i automatitzada:*
 - *Avaluació*
 - *Memòries de traducció*



Línies de recerca (II)

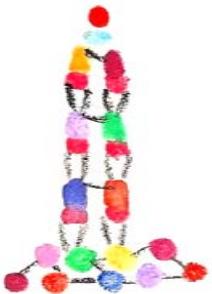
- *Desenvolupament, avaluació i aplicació de recursos lingüístics de gran cobertura:*
 - *Gramàtiques:*
 - *Dependències*
 - *HPSG.*
 - *Lèxics:*
 - *Bases de dades verbals amb informació sintàctico-semàntica*
 - *WordNet*
 - *Perífrasis verbals*



Línies de recerca (II)

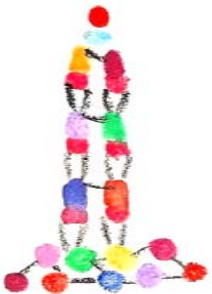
– *Corpus: anotació i explotació*

- *A nivell de mot:*
 - *categoria morfo-sintàctica*
 - *sentit*
- *A nivell de complements:*
 - *rols semàntics*
 - *categoria morfo-sintàctica*
 - *funció sintàctica*
- *A nivell d'oració:*
 - *semàntica de la construcció*
 - *aspectualitat*
 - *modalitat*



Línies de recerca (II)

- *Adquisició d'informació:*
 - *Esquemes de subcategorització*
 - *Restriccions de selecció*
- *Extracció de terminologia*
- *Traducció automàtica i automatitzada:*
 - *Avaluació*
 - *Memòries de traducció*



Projectes

- **Sensem: Banc de dades sintàctic i semàntic**
(Ministerio de Ciencia y Tecnología - BFF2003-06456 - 2004-2006)
- **Ampliació de la BD lèxica i el corpus sintàctico-semàntic de semàntica oracional del espanyol SenSem**
(Ministerio de Educación y Ciencia HUM2007-65267 - 2007-2009)
- **Anotació semàntica dels noms i adjectius de SenSem**
(Ministerio de Ciencia e Innovación FFI2008-02579-E/FILO - 2008-09)



Projectes

- ***KNOW***: desenvolupament de tecnologies multilingües a gran escala per a la comprensió del llenguatge

(Ministerio de Educación y Ciencia - TIN2006-1549-C03-02 - 2006-09)

- ***WordNet 3.0***

(Ministerio de Educación y Ciencia - HUM2006-27968-E - 2007-08)



Eines i recursos

- **Corpus trilingüe paral·lel** (anglès-espanyol-català) i anotat morfosintàcticament
<http://grial.uab.es/cpg/>
- **Lèxic verbal multilingüe** (català-castellà-francès-basc) amb informació sobre: esquemes de subcategorització, la semàntica associada a aquests esquemes, rols semàntics i exemples d'ús
<http://grial.uab.es/multi/>
- **Gramàtica de dependències**



Eines i recursos

- **Corpus trilingüe:** <http://grial.uab.es/cpg/>
 - *Especialitat: informàtica*
 - *Paral·lel (anglès-espanyol-català)*
 - *Anotat morfosintàcticament*
- **Lèxic verbal:** <http://grial.uab.es/multi/>
 - *Multilingüe (català-castellà-francès-basc)*
 - *Esquemes de subcategorització*
 - *Rols semàntics*
 - *Semàntica oracional*
 - *Exemples d'ús*
- **Gramàtica de dependències**



Eines i recursos

(En desenvolupament)

- **Corpus SenSem**

- **Registre periodístic**

- **Anotat amb informació:**

- *morfosintàctica*

- *estructura argumental*

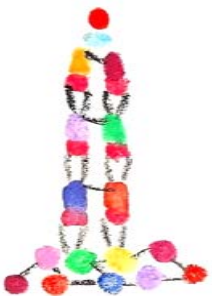
- *semàntica de la construcció*

- *freqüències*

- *aspectualitat*

- *modalitat*

- **Paral·lel (català-castellà) a nivell de mot**



Futur

- *Lèxics:*
 - *Base de dades lèxica verbal **SenSem** (bolcatge de dades del corpus SenSem)*
 - *WordNet català 3.0*
 - *Diccionari multilingüe (català-castellà-romanès) de **perífrasis verbals***
- *Corpus:*
 - *Corpus paral·lel català-castellà-anglès a nivell de mot (glosses WordNet 3.0)*
- *Gramàtiques: HPSG*

