

The HISPACAT comparative database of syntactic constructions and its applications to syntactic variation research

Workshop on Research Infrastructure for Linguistic Variation

Xavier Villalba

Centre de Lingüística Teòrica, Universitat Autònoma de Barcelona
Xavier.Villalba@uab.cat

September, 18th 2009

Overview

The goals of the HISPACAT database

Theoretical background of HISPACAT

Architecture of HISPACAT

HISPACAT at work

Conclusions

Goals of HISPACAT

- ▶ HISPACAT is a **theory-driven** project aimed at
 - ▶ identification of the 'atoms' underlying syntactic microvariation in Catalan and Spanish
 - ▶ prediction of domains most vulnerable to syntactic interference
- ▶ HISPACAT is an **applied** project aimed at building
 - ▶ an empirical playground for researchers in bilingualism and L2 learning
 - ▶ a dynamic comparative grammar for L2 teachers/learners

The project and the team

- ▶ HISPACAT is part of the funded research project HUM2006-13295-FILO (Spanish MEC and UE FEDER) awarded to the Center of Linguistic Research of the Universitat Autònoma de Barcelona (UAB)
- ▶ HISPACAT involves 26 senior/junior researchers, and one head technician, Daniel Jiménez

Theoretical background of HISPACAT

- ▶ constructions as bundles of properties
- ▶ microparametric approach to syntactic variation
- ▶ standard Catalan and Spanish language data from the two major descriptive grammars:
 - ▶ *Gramática descriptiva de la lengua española*. [Spanish]
 - ▶ *Gramàtica del català contemporani*. [Catalan]
- ▶ plus two main 'official' corpora:
 - ▶ Corpus de Referencia del Español Actual. [Spanish]
 - ▶ Corpus Textual Informatitzat de la Llengua Catalana. [Catalan]

Basic construction example: expletive negation in irrealis contexts

- ▶ Catalan: ambiguous use of negation:

- (1) Tinc por que **no** vinguin.
 - a. negative: 'I am afraid that they won't come.'
 - b. expletive: 'I am afraid that they should come.'

- ▶ Spanish: nonambiguous use of negation:

- (2) Tengo miedo de que **no** vengan.
 - a. negative: 'I am afraid that they won't come.'
 - b. expletive: *'I am afraid that they should come.'

The HISPACAT basic architecture

- ▶ HISPACAT is a **concept-oriented** relational ORACLE DB
- ▶ HISPACAT counts with an **ontology of 176 linguistic concepts**
- ▶ HISPACAT contains around **500 construction files**
- ▶ HISPACAT counts with a **robust data-retrieval system** both for concepts and text
- ▶ HISPACAT is build on truly **comparative basis**

The files: simplified version

locative possessive pronoun

EX_CAT: Era sota meu

REF_CAT: GCC 7.5.2.8

EX_CAST:

REF_CAST:

ANÀL: The possessive locative pronoun is only possible in Catalan, and it surfaces as the complement of simple and complex locative prepositions replacing a genitive PP (cf. *prop teu/prop de tu*). This construction sounds odd with inanimate referents (cf. *sota de la taula/??sota seu*).

The Conceptual Ontology: basic concepts

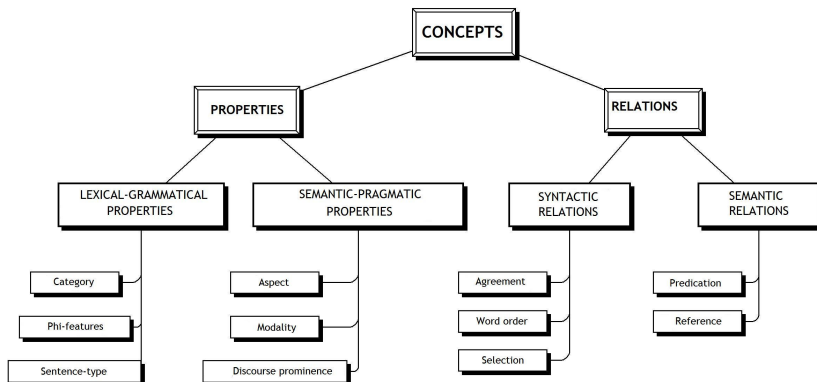


Figura: Ontology: basic concepts

Concept-based description of constructions

Selected Static locative PP headed by a 'to'

▶ EX-CAT:

1. Residim a Tàrrega. ('We live in Tàrrega.')
2. Habitareu a la setena planta. ('You will dwell in the seventh floor.')

▶ CONC:

1. LEXICAL-GRAMMATICAL-PROPERTY/CATEGORY/PREPOSITION/STATIC;
2. SYNTACTIC-RELATION/SELECTION/ARGUMENT;
3. LEXICAL-GRAMMATICAL-PROPERTY/CATEGORY/VERB/SELECTION/INERGATIVE;
4. LEXICAL-GRAMMATICAL-PROPERTY/CATEGORY/VERB/EVENTIVE-STRUCTURE/STATIVE

Concept-based boolean searches

- Row text contains 'expletiva'
- Row text contains 'negació'

Numero	Autor	Data	Denominació Catalana	Denominació Espanyola	Conceptes Clau
5.4	Maria Teresa Espinal	Dimecres, 17 Novembre, 2004	Negació expletiva en la subordinada induïda per una preposició temporal	Negación expletiva en la subordinada inducida por una preposición temporal	/propietats_lexicogramaticals/categoria/preposició/de_temps; /propietats_lexicogramaticals/categoria/adverbi/de_negació; /propietats_lexicogramaticals/trets_morfosintàctics /mode/subjuntiu; /propietats_lexicogramaticals /trets_morfològics/mode/indicatiu; /propietats_semanticopragmàtiques/modalitat/irreal
5.3	Maria Teresa Espinal	Dilluns, 27 Setembre, 2004	Negació expletiva en el context d'oracions exclamatives quantitatives	Negación expletiva en el contexto de oraciones exclamativas cuantitativas	/propietats_lexicogramaticals/trets_morfològics /modalitat_oracional/exclamativa; /propietats_lexicogramaticals/categoria/adverbi/de_negació; /propietats_lexicogramaticals/categoria/quantificador /quantitatiu; /propietats_semanticopragmàtiques/modalitat /irreal
5.6	Maria Teresa Espinal	Dimecres, 17 Novembre, 2004	Negació expletiva en la subordinada induïda per una preposició negativa	Negación expletiva en la subordinada inducida por una preposición negativa	/propietats_lexicogramaticals/categoria/preposició/negativa; /propietats_lexicogramaticals/trets_morfosintàctics /mode/subjuntiu; /propietats_lexicogramaticals/categoria /adverbi/de_negació; /propietats_semanticopragmàtiques /modalitat/irreal
5.2	Maria Teresa Espinal	Dilluns, 27 Setembre, 2004	Negació expletiva en el context de marcadors comparatius	Negación expletiva en el contexto de marcadores comparativos	/propietats_lexicogramaticals/categoria/quantificador /de_grau; /propietats_lexicogramaticals/categoria/adverbi /de_negació; /propietats_semanticopragmàtiques/modalitat /irreal
5.5	Maria Teresa Espinal	Dimecres, 17 Novembre, 2004	Negació expletiva en la subordinada induïda per un verb que denota privació, oposició i valoració negativa	Negación expletiva en la subordinada inducida por un verbo que denota privación, oposición y valoración negativa	/propietats_lexicogramaticals/categoria /verb/classe_semàntica/privació; /propietats_lexicogramaticals/categoria /verb/classe_semàntica/oposició; /propietats_lexicogramaticals/categoria /verb/classe_semàntica/negatiu; /propietats_lexicogramaticals/trets_morfosintàctics /mode/subjuntiu; /propietats_lexicogramaticals/categoria /adverbi/de_negació; /propietats_semanticopragmàtiques /modalitat/irreal

Concept-based boolean searches

Filter

Column	Operator	Expression
Conceptes Clau	does not contain	moviment

Cancel Delete Apply

- Row text contains 'locatiu'
- Conceptes Clau does not contain 'moviment'

Numero	Autor	Data	Denominació Catalana	Denominació Espanyola	Conceptes Clau
6.021	Gemma Rigau	Dimecres, 31 Agost, 2005	SP locatiu encapçalat per <i>contra</i>	SP locativo introducido por <i>contra</i>	/propietats_lexicogramaticals/categoria/preposició/dinàmica /destinació; /relacions_sintàctiques/selecció/adjunt/del_predicat; /relacions_sintàctiques/selecció/argument
6.019	Gemma Rigau	Dimecres, 31 Agost, 2005	SP locatiu d'origen adjunt al predicat	SP locativo de origen adjunto al predicado	/propietats_lexicogramaticals/categoria/preposició/dinàmica/d'origen /relacions_sintàctiques/selecció/adjunt/del_predicat
7.002	Víctor Acedo Matelán	Dijous, 04 Maig, 2006	Objecte directe que denota locatiu	Objeto directo que denota locativo	/propietats_lexicogramaticals/categoria/verb/hègim/transitiu; /propietats_lexicogramaticals/categoria/verb/estructura_eventiva /dinàmic; /propietats_lexicogramaticals/categoria /verb/estructura_eventiva/duratiu; /propietats_lexicogramaticals /categoria/verb/estructura_eventiva/tèic; /relacions_sintàctiques /selecció/argument/locatiu; /relacions_semàntiques/predicació /primària/verbal
7.001	Víctor Acedo Matelán	Dimecres, 26 Abril, 2006	Complement directe que denota locatiu	Complemento directo que denota locativo	/propietats_lexicogramaticals/categoria/verb/hègim/transitiu; /propietats_lexicogramaticals/categoria/verb/estructura_eventiva /dinàmic; /propietats_lexicogramaticals/categoria /verb/estructura_eventiva/duratiu; /propietats_lexicogramaticals /categoria/verb/estructura_eventiva/tèic; /relacions_sintàctiques /selecció/argument/locatiu; /relacions_semàntiques/predicació /primària/verbal
10.08	Ester Sánchez-González	Dimarts, 09 Setembre, 2008	noms a partir de preposicions o adverbis locatius	nombres a partir de preposiciones o adverbios locativos	/PROPIETATS_LEXICOGRAMATICALS/CATEGORIA/NOM

Concept-based searches through the concept tree



Textual search

Projecte BFF2003-08364-C02 Fixtes Cerca conceptes-clau Arbre conceptual **Cercar exemples** Administració d'objectes Membres del projecte

Cerca exemples

Número	Exemples catalans	Exemples espanyols	Conceptes-clau
5.13	No vam trobar cap vi que ens agradés; No l'hem vist entloc; No fa gaire bona cara; No m'ha dit gran cosa	No encontramos ningún vino que nos gustara; No lo hemos visto en ningún sitio; No me ha dicho gran cosa	/propietats_lexicogramaticals/categoria/quantificador/indefinit;/propietats_lexicogramaticals/tres_morfologies/modaltat_oracional/declarativa/negativa;/relacions_semantiques/referencials/accesibilitat_contextual/potential
5.21	Amb prou feines entén el que diem; Ens parlem tan poc que a penes si ens truquem dos cops l'any; -Truquem. Deu ser per a tu. - A l'igual; Amb prou feines va venir ningu	Apenas tuve tiempo de nada; A duras penas llegó a hablar con nadie	/propietats_lexicogramaticals/categoria/preposició;/propietats_semanticopragmàtiques/negació
5.23	M'empipa que loqui res sense dir-m'ho; Es estrany que li costi tant de confiar en ningu ; Era un disbarat que confies en cap dels seus companys de feina	Me sorprende que haya venido nadie; Al bedel le irrita / molesta / agobia tener que decirle a nadie cómo se rellenan los impresos; Resultaba espantosa la idea de confiar en ninguno de sus amigos; Me pareció extrañísimo / muy extraño que Juan moviera un dedo por él; Es realmente insofite que a Luis se le ocurran semejantes tonterías; Fue una locura venir siquiera	/propietats_lexicogramaticals/categoria/verb/classe_semantica/factiu;/propietat_lexicogramatical/categoria/adjectiu/classe_semantica/factual;/propietat_lexicogramatical/categoria/nom/classe_semantica/factual;/propietats_semanticopragmàtiques/modaltat/real
5.24	Sense fer cap mena de treball ni d'exercici, ¿com vols aprovar?; En comptes de trucar a ningu , mira si pots solucionar el problema tu sola; Haurém d'intervenir abans que ningú demani la paraula; Busca l'arracada fins que aparegui entloc	Antes de mover un dedo por semejante personaje píensalo bien; Antes de tocar nada, lívale las manos; Es mejor que le marches sin decir nada a nadie; Escuchó toda la conversación sin que palabra alguna le sorprendiese en ningún momento	/propietats_lexicogramaticals/categoria/preposició/de_temps;/propietats_lexicogramaticals/categoria/preposició/negativa;/propietats_lexicogramaticals/categoria/preposició/SUBS;/propietats_lexicogramaticals/categoria/preposició/dinamica/de_destinació;/propietats_semanticopragmàtiques/modaltat/real
5.26	Aquí fa menys fred que entloc de tota la casa; S'estima més prendre precaucions que trobar-se mal més en una situació similar; Es la primera vegada que ha fet res similar; Es massa tard per a anar entloc a aquestes hores; No m'agrada com va vestit el seu xicot, comparat a com va cap dels nols que la ronden	Margarita canta mejor que nadie; Romualdo está más interesado en pasárselo bien que en mover un dedo por sus compañeros; Es más probable que su padre sea rico que que él tenga dos dedos de frente; Rosa fabrica más palomitas de maíz de las que yo doy abasto vendiendo; Prefiero quedarme aquí que ir a ningún sitio; Ernesto prefiere ser escéptico que tragarse semejantes disparates; La mejor / peor novela que he leído nunca; El más extraordinario suceso que ocurriera en muchos años / visto hombre alguno; Es la primera / última vez que muevo un dedo por ayudar a nadie; Manolito es bastante aplicado comparado con lo que aprovecha ninguno de sus hermanos	/propietats_lexicogramaticals/categoria/quantificador/de_grau;/propietat_lexicogramatical/categoria/verb/classe_semantica/de_grau;/propietat_lexicogramatical/categoria/adverbide_grau;/propietat_lexicogramatical/categoria/quantificador/numeral;/propietats_semanticopragmàtiques/modaltat/real

Theoretical comparative insights: possessives

- (3)
- a. El retrat era pintat meu.
the portrait was painted mine
'The portrait was painted by me.'
 - b. *El retrato estaba pintado tuyo.
the portrait was painted mine

Comparative approach: contrastive analysis

locative possessive pronoun

EX_CAT: Era sota meu

REF_CAT: GCC 7.5.2.8

EX_CAST:

REF_CAST:

ANÀL: The possessive locative pronoun is only possible in Catalan, and it surfaces as the complement of simple and complex locative prepositions replacing a genitive PP (cf. *prop teu/prop de tu*). This construction sounds odd with inanimate referents (cf. *sota de la taula/??sota seu*).

Theoretical comparative insights: right-dislocation

Villalba (2008): data from a comparative textual corpus analysis

- ▶ Catalan
 - ▶ 232 occurrences
 - ▶ all categories and functions
 - ▶ equal preference for topic (re)introduction and continuation
 - ▶ similar preference for local and nonlocal antecedents
- ▶ Spanish
 - ▶ 10 occurrences
 - ▶ DP/CP and subject/object
 - ▶ strong preference for topic (re)introduction
 - ▶ similar preference for local and nonlocal antecedents

Comparative approach: syntactic interference

'quantifier *cada* 'each/every' in temporal expressions which apparently aren't distributive'

EX_CAT: *Cada diumenge vaig al cinema*

ANÀL: [...] With temporal expressions, the distributive use is clear in *Cada dia va a escola a una hora diferent* 'Every day (s)he goes to school at a different time', but not in *Cada dia va a escola* '(S)he goes to school everyday', where Spanish resorts to the nondistributive quantifier (*Todos los días va a la escuela*). [...]

COMMENT: Indeed, maybe Spanish allows such a construction (it is clearly found in the Spanish spoken by Catalan speakers).

Conclusions

The HISPACAT Catalan-Spanish contrastive database of constructions is:

- ▶ linguistically and conceptually based
- ▶ primarily theory driven:
 - ▶ identification of the “atoms” of linguistic microvariation
 - ▶ a snapshot of parametric (in)variance
- ▶ two major applied goals:
 - ▶ a major database for contrasting methods and hypothesis in bilingualism and L2 learning,
 - ▶ a dynamic comparative grammar for L2 teachers/learners.

Takk!