

# THE HISPACAT COMPARATIVE DATABASE OF SYNTACTIC CONSTRUCTIONS AND ITS APPLICATIONS TO SYNTACTIC VARIATION RESEARCH

XAVIER VILLALBA  
*Barcelona*

## ABSTRACT

The HISPACAT database of syntactic constructions in Catalan and Spanish is a dynamic comparative grammar of two closely related languages, which, from a theoretical point of view, offers us an alternative viewpoint to advance in our understanding of the “atoms” of linguistic microvariation, and to offer a snapshot of microparametric (in)variance, which will help us to predict less stable parts of the grammar, and hence more sensitive to syntactic change or interference phenomena. Moreover, this tool, which is conceived as a major empirical source for testing syntactic microvariation, may also prove helpful for researchers in bilingualism and language contact studies, and for teachers and students of Catalan or Spanish as L2.

## [1] INTRODUCTION

HISPACAT, the database of syntactic constructions in Catalan and Spanish, integrates in a bigger funded research project of the Centre de Lingüística Teòrica of the Universitat Autònoma de Barcelona, which involves a team of 16 senior and 10 junior researchers, and one head technician, Daniel Jiménez, headed by professors Carme Picallo and Josep M. Brucart.<sup>1</sup>

HISPACAT is conceived as a major empirical source for detecting syntactic microvariation, and testing hypotheses regarding bilingualism, and interference, but which may prove helpful for L2 learning as well, for it is designed as a dynamic comparative grammar of two closely related languages capable of offering

---

[1] I am really thankful to the organizers and the audience of the Workshop on Research Infrastructure for Linguistic Variation for the lively atmosphere of knowledge-sharing, and illuminating discussion. Thanks are also due to the comments of three anonymous reviewers, which have improved the paper significantly. The research included in this paper has been supported by the funded research projects BFF2003-08364-C02 (Spanish MCyT and EU FEDER) and HUM2006-13295-C02/FILO (Spanish MEC and EU FEDER) awarded to the Centre de Lingüística Teòrica of the Universitat Autònoma de Barcelona.

a snapshot of basic (in)variance patterns. HISPACAT aims to help researchers investigating what factors of the computational system and what morphosyntactic features of lexical expressions underlie the grammatical symmetries and asymmetries between Spanish and Catalan. Moreover, as the unchanging principles governing language can only be fully explained in connection to a proper understanding of the nature of linguistic variation, HISPACAT is intended to be a modest but valuable tool for building a general theory of language.

In the first part of the paper we will discuss the goals and theoretical foundations of the project, stressing the theoretical and methodological differences existing between our enterprise and current textual databases. In the second part of the paper, a general presentation will be offered of the main architecture of the database, namely the internal structure of the files, and its conceptual ontology. It will be shown that the conceptual-based design of HISPACAT, besides providing an exhaustive description of the basic features of each construction, will prove instrumental for allowing a flexible system of information queries.

Finally, the last part of the paper will be devoted to analyze several cases related to microvariation, and interference phenomena, where HISPACAT has proved to be a reliable source for establishing robust linguistic generalizations.

## [2] GOALS AND THEORETICAL FOUNDATIONS OF HISPACAT

HISPACAT was conceived as a tool within a theory-driven project aimed at the identification of the “atoms” underlying syntactic microvariation in Catalan and Spanish and at the prediction of domains most vulnerable to syntactic interference. Moreover, HISPACAT is an applied project aimed at building an empirical playground for researchers in bilingualism and L2 learning and a dynamic comparative grammar for L2 teachers/learners.

### [2.1] *Theoretical foundations*

From a theoretical point of view, the work by Richard S. Kayne – see [Kayne \(1996, 2000, 2005\)](#); see also the pioneering work by [Borer \(1984\)](#), and [Rizzi \(1982\)](#) – has emphasized that we must go beyond the concept of (macro)parameter, coined to give account of large structural differences between language groups – like the null subject ([Jaeggli & Safir \(1986\)](#)) or the polysynthetic parameter ([Baker \(1996\)](#)) – and focus attention on variation at a smaller scale, namely minor inter or cross-linguistic nuances not affecting the overall typological profile of languages, but nonetheless generating significant differences in the behavior of certain linguistic units. This microparametric framework is thus the most suitable for the comparative work on which HISPACAT is grounded.

Just to emphasize the theoretical and empirical linguistic relevance of the HISPACAT project, we will briefly consider the constructions involving directive modality in HISPACAT. The common core includes the imperative mood for affir-

mative orders (1) and the subjunctive mood for prohibitions (2), the translated use of present (3) and future tenses (4), and the exhortative subjunctive mood introduced by the complementizer (5):

- (1) a. Tanca la porta.  
close the door  
'Close the door!'  
(Catalan: (Payrató 2002, 3.4.4))
- b. ¡Cállate, estúpido!  
shut.up stupid  
'Shut up, you idiot!'  
(Spanish: (Garrido 1999, 60.2.1.5))
- (2) a. ¡No diguis aquestes coses!  
not say.SUBJUNCTIVE.2SG these things  
'Don't say that kind of things!'  
(Catalan: (Espinal 2002, 24.2.2))
- b. No se lo des.  
not him it give.2SG  
'Don't give it to him!'  
(Spanish: (Garrido 1999, 60.2.1.3))
- (3) a. ¡Ho fas i santes pasqües!  
it do.2SG and holy.PL Easters  
'Do it, period!'  
(Catalan: (Payrató 2002, 3.4.4))
- b. De noche sales conmigo o no sales.  
of night go.out.2SG with.me or not go.out.2SG  
'At night, you go out with me or you don't go out.'  
(Spanish: (Fernández-Ramírez 1951, V.46))
- (4) a. No mataràs.  
not kill.FUT.2SG  
'You shall not murder.'  
(Catalan: (Saldanya 2002, 22.5.7.3))
- b. ¿y me vas a dejar sola? ¡oh! No harás tal cosa.  
and me go.you to leave alone o not do.future.2SG such thing  
'and you are gonna leave me alone? Why, you won't!'  
(Spanish: (Fernández-Ramírez 1951, V.46))
- (5) a. ¡Que ho hagin dibuixat abans de plegar!  
that it have.SUBJUNCTIVE.3pl draw before of leave  
'They shall draw it before leaving!'  
(Catalan: (Payrató 2002, 3.4.4))

- b. ¡Que venga Juan!  
 that come.SUBJUNCTIVE.3SG Juan  
 ‘Juan shall come!’  
 (Spanish: (Ridruejo 1999, 49.1.3))

The contrast between Catalan and Spanish arises concerning the use of infinitives for conveying directives. Even though both languages share the prepositional construction in (6), they sharply contrast with respect to infinitival (7) and retrospective imperatives (8):

- (6) a. Va, nois, a treballar. (Catalan)  
 come.3PL boys to work  
 ‘Come on, boys, to the job!’  
 b. Venga, chicos, a trabajar.  
 come.2SG boys to work  
 ‘Come on, boys, to the job!’  
 (Spanish: (Hernanz 1999, 36.4.2.3))
- (7) a. \*Nens, ¡fer-me cas! (Catalan)  
 kids make.me case  
 ‘Kids, pay me attention!’  
 b. Niños, ¡hacerme caso!  
 kids make.me case  
 ‘Kids, pay me attention!’  
 (Spanish: (Hernanz 1999, 36.4.2.3))
- (8) a. ??¡Haver-ho dit abans!  
 have-it said before  
 ‘You should have said it before!’  
 (Catalan: (Payrató 2002, 3.4.4.2))  
 b. Siento mucho llegar tarde. Haber salido antes de casa.  
 feel much arrive late have left before of house  
 ‘I am very sorry for being late. You should have leave home earlier.’  
 (Spanish: (Garrido 1999, 60.2.1.6))

Interestingly, the HISPACAT database shows us that Catalan resorts to subjunctive in these cases:

- (9) a. ¡Ho haguessis dit abans!  
 it have.PAST.SUBJUNCTIVE.2SG said before  
 ‘You should have said it before!’  
 (Catalan: (Payrató 2002, 3.4.4.2))



- they are monolingual,
- they focus on the data rather than on grammatical concepts (with the noteworthy exception of Spanish AGLE), and
- they do not include negative data.

The HISPACAT database, in contrast,

- focuses on comparative data,
- is a relational database focused on the grammatical features underlying linguistic phenomena, and
- includes negative data that set the limits of grammaticality of constructions.

Regarding the first aspect, HISPACAT is designed and built in a totally contrastive way: the presentation of data aims at showing the contact points *and* the syntactic asymmetries between Catalan and Spanish, an enterprise which has not been attempted to date. We are, thus, interested in data like the following (see also the discussion in 2.1 above):

- (12) Tinc por que no vinguin. (Catalan)  
 have fear that not come.SUBJUNCTIVE.3PL  
 a. negative: 'I am afraid that they won't come.'  
 b. expletive: 'I am afraid that they should come.'
- (13) Tengo miedo de que no vengan. (Spanish)  
 have fear of that not come.SUBJUNCTIVE.3PL  
 a. negative: 'I am afraid that they won't come.'  
 b. expletive: \*'I am afraid that they should come.'

Whereas Catalan makes an ambiguous use of negation in this particular context, Spanish can only obtain the negative reading. Our main concern is determine why this is so and try to connect this fact with other apparently unrelated ones to help us understand the key factors underlying variation in this area of grammar. Therefore, we can say that, despite its database format, HISPACAT is a dynamic comparative grammar of Catalan and Spanish, which attempts to connect the valuable independent findings arrived at by the two major grammatical works of Spanish and Catalan: [Bosque & Demonte \(1999\)](#), and [Solà et al. \(2002\)](#).

Regarding the second aspect, HISPACAT is conceived as a relational database based on linguistic concepts, rather than on particular occurrences from a textual corpus, for we are interested not in the linguistic expressions themselves but on

the concepts underlying them. Obviously, this line of work presupposes the theoretical hypotheses that (micro)syntactic variation is the result of the setting of (micro)parameters – see Kayne (1996, 2000, 2005) – and that syntactic constructions are not primitive but the result of a sum of properties – Chomsky (1981); cf. the basic assumptions of Construction Grammar, as developed by Goldberg (1995).<sup>2</sup>

Finally, also in contrast to current textual corpora, HISPACAT allows the inclusion of negative data helping the reader to appreciate the grammaticality limits of constructions. For instance, in the file corresponding to the file *expletive negation in the context of comparative markers*, we include (and comment on) data like the following:

- (14) a. Juan era antes más simpático que (no) ahora. (Spanish)  
 Juan was before more friendly that not now  
 ‘Juan was more friendly then than now.’
- b. Juan era antes más simpático de lo que (\*no) es ahora. (Spanish)  
 Juan was before more friendly of it that not is now  
 ‘Juan was more friendly then than now.’

Anyone familiar with L2 learning will appreciate the advantages of including such information, unavailable in textual corpora: the provision of information about what is possible, but also about what is not (a typical shortcoming of school grammars), helps the teacher to set the properties and extent of any construction in a more precise way, while helping learners to avoid incorrect generalizations.

### [2.3] *Applied goals*

From an applied point of view, the HISPACAT database has three main goals:

- (i) providing an empirical database for research on bilingualism and L2 learning,
- (ii) providing a comparative grammar for teachers and students of Catalan or Spanish as L2,
- (iii) providing a catalog of commented examples for teachers and students of Catalan or Spanish as L2.

There is no doubt that an increasing number of researchers are approaching the phenomena of language contact, bilingualism, and L2 learning from a gram-

[2] Certainly, as one reviewer acutely remarks, the limits between constructions can hardly be stated precisely. We agree with this observation, which reflects the empirical nature of the HISPACAT database: the actual list of constructions is not a closed set, but is subject to continuous revision in the light of new evidence.

mathematical point of view. For these researchers, HISPACAT will prove instrumental for contributing to refine the descriptions and hypotheses of scholars in such areas of applied linguistics (goal 1). Moreover, HISPACAT can also be useful to teachers and students of Catalan or Spanish as a L2, since it offers a dynamic user-oriented comparative grammar of these two languages (goal 2). Finally, the inclusion of a big number of analyzed examples of both grammatical and ungrammatical constructions represents an added value for teachers and students of Catalan or Spanish as L2, for it increases the consistency of grammatical description and facilitates the development of specific teaching materials (goal 3).

### [3] ARCHITECTURE OF HISPACAT

HISPACAT is built as an ORACLE concept-oriented relational database with two key design features:

- a system of files built on a comparative basis, and including both descriptive and analytical fields, and
- an ontology of linguistic concepts, aimed at offering both a detailed description of constructions, and a robust data-retrieval system.

#### [3.1] *The files*

HISPACAT is fed through a file system including a series of fields designed to emphasize the comparative and conceptual-based nature of the database. In [Fig. 1 on the facing page](#) you have a (simplified) example of the file corresponding to *selected static locative PP headed by a 'to'*:<sup>3</sup>

#### [3.2] *The ontology*

In order to help researchers to discover the basic properties underlying each construction, an ontology was designed of 176 linguistic concepts, classified into relationships and properties, and further subdivided into lexical-grammatical and semantic-pragmatic properties and semantic and syntactic relations, as shown in [Fig. 2 on page 138](#).

Moreover, every concept is associated with a code based on their position in the ontology, which allows us to show direct relationship networks and natural classes of concepts. Henceforth, the files of the database are designed as comparative concept-based descriptions of constructions, as in the following example (the lack of Spanish examples indicates that the construction is available in Catalan only).

[3] The codes are the following: DEN-CAT=Catalan term, DEN-ESP=Spanish term, CONC=linguistic concepts, EX-CAT=Catalan examples, REF-CAT=reference of Catalan examples, EX-ESP=Spanish examples, REF-ESP=reference of Spanish examples, ANAL=analysis, BIB=bibliography, SIN-CAT=synonyms of Catalan term, SIN-ESP=synonyms of Spanish term, and REL=related constructions.

DEN\_CAT: SP de lloc estàtic encapçalat per *a* seleccionat pel predicat

DEN\_ESP: SP de lugar estático introducido por *a* seleccionado pel predicado

CONC: PROPIEDAD\_LEXICO-GRAMATICAL/CATEGORIA/PREPOSICIÓN/ESTATIVA;  
RELACION\_SINTACTICA/SELECCION/ARGUMENTO; PROPIEDAD\_LEXICO-  
GRAMATICAL/CATEGORIA/VERBO/REGIMEN; PROPIEDAD\_LEXICO-  
GRAMATICAL/CATEGORIA/VERBO/CLASE\_SEMANTICA/AGENTIVO

EX\_CAT: La Queta viu a la platja; Residim a Tàrrega; Habitareu a la setena planta.

REF\_CAT: GCC 11.3.1; \_\_\_\_; \_\_\_\_

EX\_ESP: \_\_\_\_

REF\_ESP: \_\_\_\_

ANÀL: La preposició *a* en català és una preposició feble que pot expressar lloc estàtic o dinàmic. En espanyol peninsular només pren aquest darrer valor. Per expressar el locatiu estàtic, l'espanyol emprà *en*. En català, la preposició *en* ocupa el lloc de la preposició *a* en els parlars meridionals (com ara en valencià) o bé davant de demostratiu o quantificador la primera síl·laba del qual s'iniciï amb una vocal: *Viu en aquesta casa*. Davant de l'article definit alternen *a* i *en*: *Viuen al cotxe/en el cotxe*. En balear, normalment es diu *Viu a aquella casa*.

En tots els parlars catalans, quan el lloc és abstracte o metafòric, el verb en general selecciona *en*: *Viuen en la indigència*. Si l'objecte de la preposició és un SN escarit, també s'usa *en* (*Viuen en pisos*).

Davant del nom *casa* (*de*), l'espanyol emprà la preposició estàtica *en*, mentre que el català, en general, no. Així, en espanyol tenim *Amaneció en casa* i en català *Va pernoctar a casa*.

BIB: Ruaix 1994; GCC 11; 15; Moll 1968; GDLE 10

SIN\_CAT: Preposició de lloc no dinàmica *a*.

SIN\_ESP: Preposició de lugar no dinàmica *a*.

REL: SP de lloc estàtic seleccionats pel predicat; SP de lloc estàtic encapçalat per *en* seleccionat pel predicat.

FIGURE 1: Example of HISPACAT file.

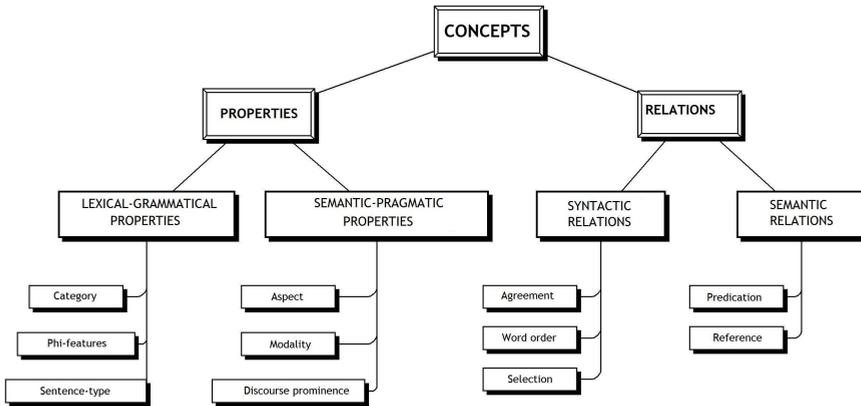


FIGURE 2: Outline of the HISPACAT ontology

## (15) Selected Static locative PP headed by a ‘to’

- EX-CAT: *Residim a Tàrrega.* (‘We live in Tàrrega.’); *Habitareu a la setena planta.* (‘You will dwell in the seventh floor.’)
- EX-ESP:
- CONC:
  - (i) lexical-grammatical-property / category / preposition / static
  - (ii) lexical-grammatical-property / category / verb / selection / inergative
  - (iii) lexical-grammatical-property / category / verb / eventive-structure / stative
  - (iv) syntactic-relation / selection / argument
- ANAL: Catalan preposition *a* ‘to’ is a weak preposition that may express a static location or movement. In Peninsular Spanish only the last value is possible. To express the static locative meaning, Spanish must resort to *en* ‘in’. When preceding a demonstrative or quantifier beginning with an initial vowel, Catalan changes to *en* ‘in’: *Viu en aquesta casa* ‘Lives at this house.’ In front of the definite article, *a* ‘to’ and *en* ‘in’ alternate: *Viu en el cotxe/en el cotxe* ‘They live at the car.’ In Catalan, when the place is metaphorical, the verb selects *en* ‘in’: *Viu en la indigència* ‘They live amidst poverty’. If the argument of the locative preposition is a bare noun, the preposition *en* ‘in’ is chosen (*Viu en pisos* ‘They live in flats.’). When preceding the name *casa* ‘home’, Spanish selects the static preposition *en* ‘in’, whereas Catalan

resorts to a ‘to’: Sp. *Amaneció en casa* ‘(S)he saw the new day at home’ and Cat. *Va pernoctar a casa* ‘(S)he stay the night at home’.

Importantly, besides its theoretical grounds, the CONC field provides HISPACAT with a strong tool for carrying out complex conceptual searches. For example, HISPACAT allows the user to get the list of the constructions sharing, for example, the concept *locative* while not including the concept *movement* (Fig. 3):

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	Victor Acedo Matellán	Dijous, 04 Maig, 2006	Objecte directe que denota locatiu	Objeto directo que denota locativo	/propietats_lexicogramaticals/categoria/verb/règim/transitiu; /propietats_lexicogramaticals/categoria/verb/estructura_eventiva /dinàmic; /propietats_lexicogramaticals/categoria /verb/estructura_eventiva/duració; /propietats_lexicogramaticals /categoria/verb/estructura_eventiva/típic; /relacions_sintàctiques /selecció/argument/locatiu; /relacions_semàntiques/predicació /primària/verbal

FIGURE 3: Boolean search.

Moreover, beyond standard boolean searches, the conceptual-based design of the database allows us to inspect the files through a conceptual tree based on the ontology. Thus, we can move from the top of the tree to a particular node and get the files including the concept corresponding to this node (Fig. 4 on the next page).

It must be emphasized that the rich system of conceptual searches is complemented with the possibility of textual searches, incorporating thus the advantages of textual corpora (Fig. 5 on page 141).

#### [4] HISPACAT AT WORK

After this quick glance at the basic architectural features of the HISPACAT database, now we will briefly discuss some immediate applications both at the theoretical and applied level.

##### [4.1] *Syntactic microvariation: theory and practice*

HISPACAT was originally designed as a tool for the study of syntactic microvariation, which is the main interest of the general research project. Its development has crucially raised many insights in this area, as we have discussed in 2.1. Now consider another case, that of the construction SELECTED STATIC LOCATIVE PP

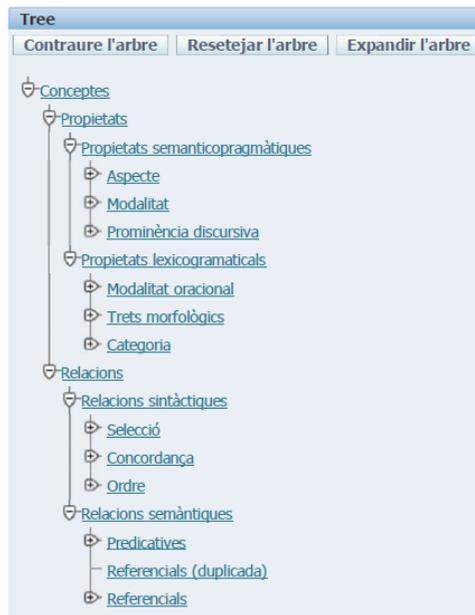


FIGURE 4: The conceptual tree.

HEADED BY A ‘TO’, which has been presented before in (15):<sup>4</sup>

(16) Selected Static locative PP headed by a ‘to’

- EX-CAT: Residim a Tàrrrega. (‘We live in Tàrrrega.’); Habitareu a la setena planta. (‘You will dwell in the seventh floor.’)
- EX-ESP:
- ANAL: Catalan preposition *a* ‘to’ is a weak preposition that may express a static location or movement. In Peninsular Spanish only the last value is possible. To express the static locative meaning, Spanish must resort to *en* ‘in’. When preceding a demonstrative or quantifier beginning with an initial vowel, Catalan changes to *en* ‘in’: *Viu en aquesta casa* ‘Lives at this house.’ In front of the definite article, *a* ‘to’ and *en* ‘in’ alternate: *Viu en el cotxe/en el cotxe* ‘They live at the car.’ In Catalan, when the place is metaphorical, the verb selects *en* ‘in’: *Viu en la indigència* ‘They live

[4] This brief example is not intended to exhaust the issue. As one reviewer points out, a more accurate description would be that “in Spanish ‘a’ can be locative, but only refers to the contact with a line (a perimeter or otherwise), while in Catalan the same preposition can refer to this type of contact or to the inclusion of an object inside the area of a bidimensional object”. For a detailed discussion, see [Fábregas \(2007\)](#).

## Cerca exemples

Número	Exemples catalans	Exemples espanyols
5.13	No vam trobar cap vi que ens agradés; No l'hem vist enlloc; No fa gaire bona cara; No m'ha dit gran cosa	No encontramos <b>ningún</b> vino que nos gustara; No lo hemos visto en <b>ningún</b> sitio; No me ha dicho gran cosa
5.21	Amb prou feines entén el que diem; Ens parlem tan poc que a penes si ens truquem dos cops l'any; -Truquen. Deu ser per a tu. - A l'igual; Amb prou feines va venir <b>ningú</b>	Apenas tuve tiempo de nada; A duras penas llegó a hablar con nadie
5.23	M'empia que toqui res sense dir-m'ho; És estrany que li costi tant de confiar en <b>ningú</b> ; Era un disbarat que confiés 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 insólito que a Luis se le ocurran semejantes tonterías; Fue una locura venir siquiera
5.24	Sense fer cap mena de treball ni d'exercici, ¿com vols aprovar?; En comptes de trucar a <b>ningú</b> , mira si pots solucionar el problema tu sola; Haurem d'intervenir abans que ningú demani la paraula; Busca l'arracada fins que aparegui enlloc	Antes de mover un dedo por semejante personaje piénsatelo bien; Antes de tocar nada, lávate las manos; Es mejor que te marches sin decir nada a nadie; Escuchó toda la conversación sin que palabra alguna le sorprendiese en <b>ningún</b> momento

FIGURE 5: Textual search.

amidst poverty'. If the argument of the locative preposition is a bare noun, the preposition *en* 'in' is chosen (Viuen en pisos 'They live in flats.'). When preceding the name *casa* 'home', Spanish selects the static preposition *en* 'in', whereas Catalan resorts to a 'to': Sp. Amaneció en casa '(S)he saw the new day at home' and Cat. Va pernoctar a casa '(S)he stay the night at home'.

The analysis is purposely descriptive and devoid of theoretical apparatus and specialized terminology, and focuses on establishing the contrasting use of locative prepositions in both languages. Moreover, the description included, when combined with the examples, helps the learner of Catalan as L2 to capture the basics of Catalan grammar concerning locative complements, while helping him or her to avoid one of the standard errors of Catalan L2 students: the use of preposition *en* 'in' –\**Viu en Barcelona* '(S)he lives in Barcelona'– instead of preposition *a* 'to' –*Viu a Barcelona* '(S)he lives in Barcelona'.

The educational benefits of HISPACAT extend to other aspects as well. On the one hand, HISPACAT provides the students of Catalan or Spanish as a L2 with a user-oriented and custom-built comparative grammar which, unlike any currently available tool, stresses the knowledge of the mother-tongue language to appraise the complex grammatical aspects of the second language. On the other hand, the inclusion of a big number of analyzed examples of both grammatical and ungrammatical constructions which can be retrieved through a series of complex and robust query systems, helps the teacher to obtain ready-made examples of all kinds of constructions to illustrate his or her grammatical explanations in the classroom, and facilitates the development of specific teaching materials.

[4.2] *Comparative approach: syntactic interference*

A second major class of applications of the HISPACAT database concerns the creation of an empirical playground for testing hypotheses about syntactic interference in cases of bilingualism or L2 learning. Note a particularly clear example of interference reflected in the following construction (I emphasize the relevant part):

(17) quantifier cada ‘each/every’ in temporal expressions which apparently aren’t distributive’

- EX-CAT: Cada diumenge vaig al cinema
- ANAL: [...] 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* ‘(S)he goes to school everyday’). [...]
- COMMENT: *Indeed, maybe Spanish allows such a construction (it is clearly found in the Spanish spoken by Catalan speakers).*

After the description of the differences concerning the use of this quantifier in standard Catalan and Spanish, the *comment* field includes valuable information about the interference of the Catalan system in the Spanish dialect spoken in Catalonia.

Consider now an example of interference the other way around (I emphasize the relevant part):

(18) phase adverbs *ya* ‘already’ and *todavía* ‘still’ in immediate postverbal position

- DEN-CAT: adverbis de fase *ya* i *todavía* en posició immediatament postverbal
- DEN-ESP: adverbios de fase *ya* y *todavía* en posición inmediatamente postverbal
- EX-CAT:
- REF-CAT:
- EX-ESP: Está *ya/todavía* aquí ‘(S)he is already/still here’
- REF-ESP: GCC 26.2.2.1

- ANAL: Both Spanish and Catalan allow the adverb-verb ordering: Sp. *Ya/Todavía está aquí* '(S)he is already/still here'; Cat. *Ja/Encara és aquí* '(S)he is already/still here'. The existence of the other order in Spanish seems to be related to the existence of a higher (left) position for the verb, which one can appreciate in other constructions like Sp. *Tiene usted mucha suerte* 'You are very lucky' o Sp. *Es quizá cierto*, where the verb may precede both the subject and certain adverbs in Spanish, but not in Catalan (or at least less easily).
- COMMENT: *As it is usually the case, one finds the spurious copy of the Spanish order in Catalan, particularly in formal written texts.*

Crucially, this is the kind of linguistic information that can hardly be found in monolingual grammars or textual databases, but is best suited to contribute to help testing empirical hypotheses about grammatical symmetries and asymmetries stemming from language contact.

#### [5] CONCLUSIONS

From the preceding discussion we can conclude that the HISPACAT Catalan-Spanish contrastive database of constructions is linguistically and conceptually based, and grounded on the theoretical framework of syntactic microvariation. Its main theoretical goals are the identification of the "atoms" of linguistic microvariation, and offering a snapshot of those grammatical areas more vulnerable to syntactic variation and interference. From an applied point of view, HISPACAT is conceived as a major database for contrasting methods and hypothesis in bilingualism and L2 learning, and as a dynamic comparative grammar and catalog of commented examples for teachers and students of Catalan or Spanish as L2. To attempt these goals, HISPACAT develops an ontology of linguistic concepts decomposing the basic features of each particular construction, while allowing a flexible and robust system of concept queries.

#### REFERENCES

- Baker, M. 1996. *The polysynthesis parameter*. Oxford/New York: Oxford University Press.
- Borer, H. 1984. *Parametric syntax: Case studies in semitic and romance languages*. Dordrecht: Foris.
- Bosque, I. & V. Demonte. 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa.
- Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

- Espinal, M. T. 2002. La negació. In J. Solà, M. R. Lloret, J. Mascaró & M. P. Saldanya (eds.), *Gramàtica del català contemporani*, vol. 3, 2727–2797. Barcelona: Empúries. Chapter sintaxi-24.
- Fernández-Ramírez, S. 1951. *Gramática española*, chap. 4 El Verbo Y la Oración. Madrid: Arco/Libros.
- Fábregas, A. 2007. The exhaustive lexicalisation principle. In Monika Bašić, Marina Pantcheva, Minjeong Son & Peter Svenonius (eds.), *Tromsø working papers on language & linguistics*, vol. 34, 2, 165–199. Nordlyd. Special issue on Space, Motion, and Result.
- Garrido, J. 1999. Los actos de habla. las oraciones imperativas. In I. Bosque & V. Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. III, 3879–3928. Espasa, Madrid.
- Goldberg, A. E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Hernanz, M. L. 1999. El infinitivo. In *Gramática descriptiva de la lengua española*, vol. II, 2197–2356. Espasa, Madrid.
- Jaeggli, O. & K. Safir. 1986. The null-subject parameter and parametric theory. In *The null-subject parameter*, 1–44. Dordrecht: Kluwer.
- Kayne, R. S. 1996. Microparametric syntax. some introductory remarks. In J. Black & V. Montapanyane (eds.), *Microparametric syntax and dialectal variation*, ix–xvii. Amsterdam/Filadelfia: John Benjamins.
- Kayne, R. S. 2000. *Parameters and universals*. New York: Oxford University Press.
- Kayne, R. S. 2005. Some notes on comparative syntax, with special reference to english and french. In G. Cinque & R. S. Kayne (eds.), *Handbook of comparative syntax*, 3–69. New York: Oxford University Press.
- Payrató, L. 2002. L'enunciació i la modalitat oracional. In J. Solà, M. R. Lloret, J. Mascaró & M. P. Saldanya (eds.), *Gramàtica del català contemporani*, vol. 2, 1149–1220. Barcelona: Empúries.
- Quer, J. 2002. Les relacions temporals i aspectuals. In J. Solà, M. R. Lloret, J. Mascaró & M. P. Saldanya (eds.), *Gramàtica del català contemporani*, vol. 3, chap. sintaxi-25, 2799–2866. Empúries.
- Ridruejo, E. 1999. Modo y modalidad. el modo en las subordinadas sustantivas. In I. Bosque & V. Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. II, 3209–3252. Madrid: Espasa.

Rizzi, L. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Saldanya, M. Pérez. 2002. Les relacions temporals i aspectuals. In J. Solà, M. R. Lloret, J. Mascaró & M. P. Saldanya (eds.), *Gramàtica del català contemporani*, vol. 3, chap. sintaxi-22, 2567–2662. Empúries.

Solà, J., M. R. Lloret, J. Mascaró & M. Pérez Saldanya (eds.). 2002. *Gramàtica del català contemporani*. Barcelona: Empúries.

AUTHOR CONTACT INFORMATION

Xavier Villalba

Universitat Autònoma de Barcelona

Dept. de Filologia Catalana

Facultat de Lletres (edifici B)

ES-08193 Cerdanyola del Vallès

Spain

[Xavier.Villalba@uab.cat](mailto:Xavier.Villalba@uab.cat)