

The syntax of Comparative Correlatives in French and Spanish

Anne Abeillé, University Paris 7 Robert Borsley, University of Essex
Maria-Teresa Espinal, University Autonomous of Barcelona
abeille@linguist.jussieu.fr <rborsley@essex.ac.uk> teresa.espinal@uab.es

Introduction¹

Long neglected as belonging to the "periphery", Comparative correlatives (CC) have been much studied recently: Culicover & Jackendoff (1999) propose (for English) that they are a special construction with a symmetric syntax and an asymmetric semantics. Borsley (2004) argues that they are one of a number of non-standard head-adjunct structures (with the first clause as a syntactic adjunct). Den Dikken (2005) proposes a universal syntactic analysis of CCs as involving a subordinate (relative) clause adjoined to a main clause and claims that no special construction is needed. We present here some new data from Romance languages showing that:

- CCs require specific constructions,
- two syntactic patterns are available for CCs: an asymmetric pattern, as in English, Spanish (2a) or Italian (3a), or a symmetric pattern, as in Spanish (2b), or Italian (3b),
- one language can have the two patterns.

French appears to have only one construction (1), but, depending on the speakers, it can be analysed as belonging to the symmetric or the asymmetric pattern.

- | | |
|-------------|---|
| (1) French | Plus je lis (et) plus je comprends
the more I read (and) the more I understand |
| (2) Spanish | a Cuanto más leo, (tanto) más entiendo
how-much more I-read, (that-much) more I understand
The more I-read, the more I understand |
| | b Más leo (y) más entiendo
more I-read (and) more I-understand |
| (3) Italian | a Quanto più leggo, (tanto) più capisco
how-much more I-read, (that-much) more I-understand
The more I read, the more I understand |
| | b Più leggo (e) più capisco
More I-read (and) more I-understand |

For reasons of space, in the rest of the paper, we leave aside Italian, which does not seem to differ from Spanish in any substantial way.

1. The syntactic properties of French CC

1.1 The structure of each clause

In each clause, the fronted phrase can be AP, AdvP, NP or PP and must begin with a comparative form (*plus, moins, mieux, meilleur, pire*), or a predicative preposition (*en, de*):

- | | |
|-----|---|
| (4) | a [Plus brillante] _{AP} est l'interprétation, [plus profond] _{AP} est le ravissement de l'auditeur
the more brilliant is the interpretation, the more profound is the listener's ravishment |
| | b [Plus vite] _{AdvP} vous diagnostiquez, [meilleur médecin] _{NP} vous êtes
the faster you diagnose, the better a doctor you are |
| | c [Plus] tu te reposeras, [en meilleure forme] _{PP} tu seras à ton retour
the more you'll rest, in better shape you'll be on your return |

It cannot begin with a determiner (5a) or a non predicative preposition (5b):

- | | |
|-----|---|
| (5) | a *[Plus vite] _{AdvP} vous diagnostiquez, [un meilleur médecin] _{NP} vous êtes
the faster you diagnose, the better a doctor you are |
| | b * [Plus] tu sors, [avec plus de gens] _{PP} tu parles |

¹ We want to thank for their comments or judgements Olivier Bonami, Paul Cappeau, Annie Delaveau, Claire Blanche-Benveniste, Ángel Gallego, Danièle Godard, Brenda Laca, Oscar García-Marchena, Jaume Mateu, François Mouret, Georges Rebuschi, Louisa Sadler, Marie-José Savelli, Dan Van Raemdonck.

the more you go-out, with more people you talk

Fronted *plus* (or *moins*) can exhibit 'quantification at a distance' over an NP or AP, like other French degree adverbs (*combien*, *tant*, *beaucoup*...cf Obenauer 1983) □

- (6) a Plus l'interprétation est [brillante]_{AP}, plus le ravisement est [profond]_{AP}
- b [Moins d'argent] vous avez, [plus de mal] vous avez pour vivre
the less money you have, the more trouble you have for living
- c Moins vous avez [d'argent]_{NP}, plus vous avez [de mal]_{NP} pour vivre

Both clauses are usually head-filler phrases (as in English, cf Borsley 2004), with an unbounded dependency (7a), they can also be head-subject phrases (7b):

- (7) a Plus vous voulez avoir [de calme], [plus loin] il faut que vous alliez -
the more you want to have quietness, the further you have to go
- b Plus il y aura de policiers, [moins de crimes] seront commis
the more there will-be policemen, the less crimes will-be committed

Both clauses must be finite (8a-b), or both verbless (8c-d):

- (8) a Je crains que plus je mange, plus je grossisse
I fear that the more I eat, the more I get-fat
- b * Je crains de plus manger, plus grossir
I'm afraid of more eating, more getting fat
- c Plus belle la mariée, plus heureux le mariage
the prettier the bride, the happier the marriage
- d Plus d'amis, plus de joie
the more friends, the more happiness

Sentence (8c) is a case of copula omission, while sentence (8d) consists of two NPs.

1.2. The relationship between the two clauses

Let us call the first clause C1 and the second clause C2.

As noted by Beck 1997, C&J 1999, CC are interpreted like conditional sentences. (1) means 'Si je lis plus, alors je comprends plus' (If I read more, then I understand more). They have a fixed ordering, like if-then clauses (cf Borsley 2004), but differ from them in that C1 is presupposed in CC, and not in if-then clauses (# means unfelicitous):

- (9) a S'il pleut, alors Jean sera content (mais il ne pleut jamais)
if it rains, then J will-be happy (but it never rains)
- b Plus il pleuvra, plus Jean sera content (# mais il ne pleut jamais)
the more it will-rain, the happier J will-be (but it never rains)

The syntax of CC also differs from that of conditional sentences. C1 can have future morphology (10b), C2 cannot be an imperative nor a question in CC (11), whereas it is possible with if-then clauses (12) □

- (10) a * Si Jean courra, alors il sera fatigué
if Jean will-run, then he'll-be tired
 - b Plus Jean courra, plus il sera fatigué
the more Jean will-run, the more he'll-be tired
- (11) a * Plus tu lis, plus apprends □
the more you read, the more learn !
 - b * Plus tu lis, plus comprends-tu □
the more you read, the more do you understand □
 - c Est-ce que plus on lit, plus on apprend ?
is it that the more one reads, the more one learns ?
- (12) a Si tu cours, alors ne te fatigue pas !
If you run, then don't get tired
 - b Si Jean court, alors qui l'aidera ?

If Jean runs, then who will help him ?

The only way to ask a question is to embed the whole CC under an interrogative marker (*est-ce que*), as in (11c). We thus conclude that French CC must have a propositional content.²

We now test whether C1 can be analysed as syntactically subordinate in French CC. To test whether C1 is a subordinate clause, we use clitic subject inversion, extraction, and verbal mood. Clitic inversion is ruled out in subordinate clauses (13b), but it is possible in C1 (13c)□

- (13) a Peut-être viendra-t-il
Maybe will-he come
- b *Je pense que peut-être viendra-t-il
I think that maybe will-he come
- c (Paul a peu de temps). Aussi plus vite commencera-t-il, plus vite aura-t-il fini
Paul has little time left. So the faster starts he, the faster is he done

It is also expected that if C2 is a main clause (and C1 an embedded clause), one can extract a complement out of C2 without extracting anything out of C1 (cf 14a). Extraction is indeed possible out of French CCs, but only out of both clauses simultaneously□(4b):

- (14) a C'est un livre, que si tu veux, je lirai
it is a book that if you want , I will read -
- b C'est un livre que plus tu lis -, plus tu apprécies
this is a book that the more you read -, the more you like
- c * C'est un livre dont plus tu le lis, plus tu te souviens -
this is a book of which the more you read it, the more you remember -
- d * C'est un livre dont plus tu te souviens , plus tu l' apprécies
this is a book of which the more you remember , the more you like it

It is also expected that if C2 is a main clause, its verbal mood will be selected (in embedded contexts) independently of the verbal mood of C1 (15b). We thus test CC embedded under a verb triggering the subjunctive mood (16). It is impossible to have the selected subjunctive form in C1 only, which means that it is not the case that C1 is a main clause and C2 an embedded clause. With respect to subjunctive in C2, there is variation among speakers. Some speakers accept it only when there is also a subjunctive form in C1 (and reject 16b), while others can have subjunctive in C2 only (and accept 16b):

- (15) a Il faudrait que l'on reçoive / *reçoit des aides
it must that one receives (subjunctive / * indicative) some aids
 - b Il faudrait que si on en a besoin, on reçoive des aides
it must that if one needs (indicative), one receives (subjunctive) des aides
- (16) a * Il faudrait que plus on en ait besoin, plus on reçoit d'aides
it must that the more one has-subj need for it the more one receives-ind aids
 - b % Il faudrait que plus on en a besoin, plus on reçoive d'aides
it must that the more one has-ind need for it, the more one receives-subj aids
 - c % Il faudrait que plus on en ait besoin, plus on reçoive d'aides
it must that the more one has-subj need for it, the more one receives-subj aids

We call speakers who want the same mood in both clauses speakers A, and those who don't speakers B. Some speakers (usually speakers B) also accept a plus-clause as an adjunct clause, after an ordinary clause, outside CC□constructions:

- (17) % Ça risque d'empirer, plus le temps passe
Things may get worse, the more time is passing

In this case, as in the "reverse" CC construction in English, it is clear that the first clause is the main clause, and the second clause is a subordinate clause.

We conclude that the syntax of French CC is symmetric with respect to clitic inversion and to

² We assume that in verbless CC, each clause also has a propositional content (cf 8d).

extraction, but that for some speakers (our B speakers) there is syntactic asymmetry based on verbal mood. For A speakers, the syntax is completely symmetric and CC can be analysed as a subtype of coordinate phrases (with some specific constraints), while for B speakers, the syntax is less symmetric and a comparative clause can also be used as a subordinate clause (with some specific constraints).³

2. Spanish Comparative Correlatives

As in French, we have observed some speaker variation. We rely on Sánchez 2005's data for the asymmetric pattern, and on our informants for the symmetric pattern (which Sánchez ignores).

2.1. Internal structure of each clause

The fronted comparative begin with a comparative form (*más*, *menos*, *mejor*, *menor*, *mayor*, *peor*) which can be premodified by *cuanto* (in C1), or *tanto* (in C2) :

- (18) a (Cuantos) más libros leo, (tantas) más cosas entiendo
how-much-mpl more books I read, thus more-fpl things I understand
- b Cuanto más prescribas, [mejor médico]_{NP} serás
how-much more you-prescribe, better doctor you'll-be

The comparative phrase can begin with a preposition, but not with a determiner:

- (19) a Cuanto más sales, [de mejor humor]_{PP} te encuentras
how-much more you-go-out, of better mood you-are
- b * Cuanto más prescribas, [un mejor médico] serás
how-much more you-prescribe, a better doctor you'll-be

As in French, both clauses can also be verbless:

- (20) a Más bonita la novia, más feliz la boda
the prettier the bride, the happier the marriage
- b Más estudiantes, más exámenes (por corregir)
the more students, the more exams (to grade)

3.2 The relationship between the two clauses

In Spanish, two different CC constructions can be identified: the first one (with *cuanto*) disallows 'y' insertion and displays asymmetry in mood or extraction, while the second one (without *cuanto*) permits 'y' insertion and requires syntactic similarities between the two clauses (same mood, and parallel extraction):

- (21) a Cuanto más leo (*y) (tanto) más entiendo
how-much more I-read (*and) (that-much) more I-understand
- b Más leo (y) más entiendo
the more I-read (and) the more I-understand

The *cuanto* clause is a subordinate clause, and can be used outside the CC as an ordinary adjunct clause in (22a), unlike plain comparative clauses (22b):

- (22) a Entiendo más, cuanto más leo
I-understand more, the more I-read
- b * Entiendo más, más leo

According to Sánchez (2005), the subjunctive is triggered only in C2 (in the asymmetric pattern); according to our informants, the same mood is required in both clauses in the symmetric pattern.

- (23) a Es posible que cuantos más libros {lees/leas} más {*sabes/sepas} del asunto.
It is possible that how-much more books you-read (ind/subj), more you-know (*ind/subj) on the subject
- b Es posible que más libros {*lees/leas} y más {*sabes/sepas} del asunto.

According to Sánchez, extraction is possible out of C2 only, but not out of C1 only (in the

³ A difference with other coordinate constructions (Savelli 1995) is that gapping is impossible: * *Plus Paul lit Proust, et plus Marie Balzac.* (the more P reads Proust, the more M Balzac). This can be explained by a symmetric constraint requiring both clauses to be verbal (or both verbless).

asymmetric construction) according to our informants extraction is not possible out of one clause only in the symmetric pattern

- (24) a. Dime de quién_i [[cuanto más lo conoces] menos te fías -_i]
 Tell me whom the more you know him the less you trust -
 b. *Dime a quién_i [[cuanto más conoces -_i] menos te fías de él]
 Tell-me whom the more you-know -, the less you trust him
 c. *Dime de quién más lo conoces y menos te fías -
 d. *Dime a quién más conoces - y menos te fías de él

With the symmetric pattern, extraction is possible only out of both clauses :

- (25) Este es un tipo de aceite del que más uno compra - y más utiliza- en las ensaladas
 This is a type of oil of which more one buys and more one uses in the salads

We conclude that the symmetric CC in Spanish (without *cuanto*) is a non standard type of coordinate construction, and the asymmetric CC (with *cuanto*) is a non standard type of subordinate construction, with the *cuanto*-clause being the subordinate clause.

There are further differences between the two patterns. The order of both clauses is fixed with the symmetric pattern (for a given meaning) but, for some speakers, it is freer with the asymmetric pattern:

- (26) a. % (Tanto) más entiendo, cuanto más leo
 the more I read the more I understand
 b. % Más me parezco a Scarlet Johanson, cuanto más me maquillo
 the more I make up, the more I resemble SJ

Another difference is semantic. In the asymmetric pattern, C1 is not presupposed, while in the symmetric pattern, C1 is presupposed or asserted:

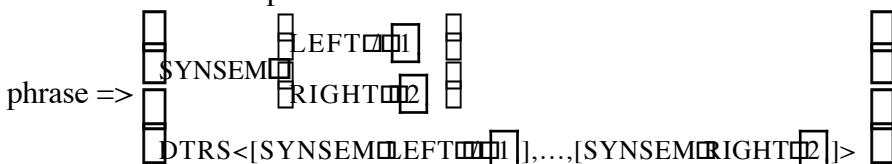
- (27) a. Más me maquillo y más me parezco a Scarlet Johanson (# pero no me maquillo)
 the more I make up and the more I resemble SJ (but I don't make up)
 b. Cuanto más me maquillo, más me parezco a Scarlet Johanson (pero no me
 maquillo)
 How-much more I make up, the more I resemble SJ (but I don't make up)

3. HPSG Analysis

3.1. Internal structure of each clause

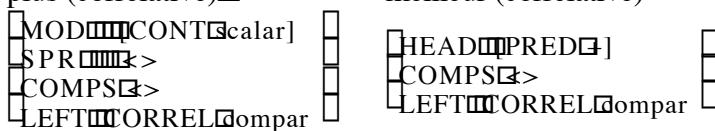
We rely on an EDGE feature (cf. Bonami et al. 2004), which is part of SYNSEM and has two values LEFT and RIGHT (each with their own *left* and *right* values). We define a LEFT feature [CORREL string] to identify the comparative correlative forms in the lexicon, and to percolate the information on the left edge of the clause. We define the EDGE feature principle as a default principle (which can be violated by specific constructions such as CC):

- (28) EDGE feature Principle:



The comparative forms in CC are specifiers or adjuncts to various categories (like other degree quantifiers) with a MOD feature selecting a scalar predicate (cf Abeillé and Godard 2003), and a specific feature [LEFT CORREL compar].⁴ The Spanish markers *cuanto* and *tanto* are analysed as specifiers selecting a comparative phrase, with two specific LEFT CORREL values :

- (29) plus (correlative) meilleur (correlative)



⁴ For an HPSG analysis of quantification at a distance, see Abeillé et al. 2005.

(30)	más (correlative)	cuanto (correlative)	tanto (correlative)
	$\boxed{\text{MOD}} \boxed{\text{CONT}} \boxed{\text{scalar}}$ $\boxed{\text{SPR}} \boxed{(\text{cuanto} \sqcap \text{tanto})} >$ $\boxed{\text{COMP}} \boxed{>}$ $\boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{ompar}}$	$\boxed{\text{SPEC}} \boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{ompar}}$ $\boxed{(\text{INDEX} \sqcap \text{msg})}$ $\boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{cuanto}}$	$\boxed{\text{SPEC}} \boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{ompar}}$ $\boxed{(\text{INDEX} \sqcap \text{msg})}$ $\boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{tanto}}$

We assume that the conjunctions (*et*, *y*) and some prepositions inherit the LEFT CORREL feature from their complement⁵.

For Spanish asymmetric CC, and for French B speakers, we define a special type of adjunct clause (with a specific MOD feature):

(31)a	compar-clause \square (French B speakers) \Rightarrow	$\boxed{\text{HEAD}} \boxed{\text{finite}} [\text{MOD} \boxed{\$} \text{[finite]}]$ $\boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{ompar}}$ $\boxed{\text{HD-DTR}} \boxed{[\text{HEAD} \text{MOD} \boxed{\$} \text{one}]} >$
b	<i>cuanto</i> -clause \Rightarrow	$\boxed{\text{HEAD}} \boxed{\text{finite}} [\text{MOD} \boxed{\$} \text{[finite]}]$ $\boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{cuanto}}$ $\boxed{\text{HD-DTR}} \boxed{[\text{HEAD} \text{MOD} \boxed{\$} \text{one}]} >$

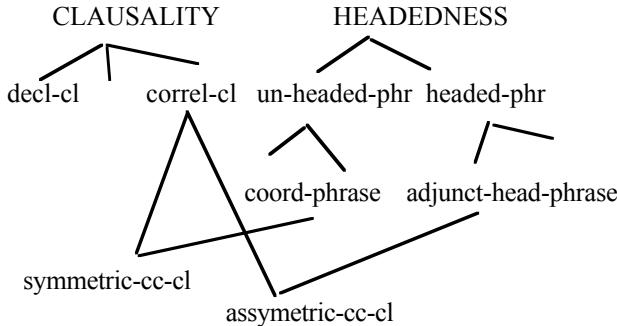
3.2. CC constructions

We follow Borsley 2004 in assuming that CC belong to a family of specific correlative constructions (they are binary clauses, each clause starting with a correlative word) which inherit from more general constructions of the language. French and Spanish data show that CC have two subtypes:

- symmetric CC, which inherits from coordinate phrases (Spanish and French A speakers)
- asymmetric CC, which inherits from head-adjunct phrases (Spanish and French B speakers)

We thus define the following clause hierarchy:

(32)



We define a general (binary) correlative-clause type, that is suitable for CC and also for other correlative constructions, such as *as-so* constructions in English (cf Borsley 2004):⁶

(33)	correl-cl \Rightarrow	$\boxed{\text{SYNSEM}} \boxed{\text{LEFT}} \boxed{\text{CORREL}} \boxed{\text{nil}}$ $\boxed{\text{CONT}} \boxed{\text{proposition}}$ $\boxed{\text{DTRS}} \boxed{[\text{LEFT} \text{CORREL} \boxed{\$} \text{one}], \boxed{\text{LEFT} \text{CORREL} \boxed{\$} \text{one}}} >$
------	-------------------------	--

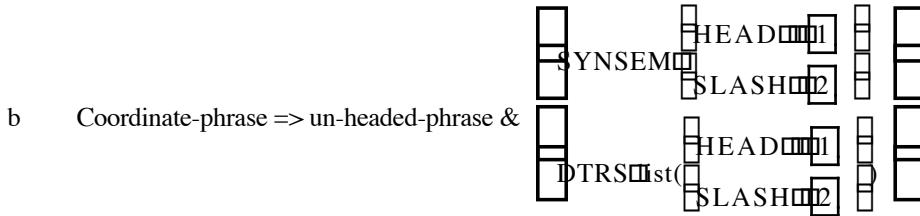
This type of clause has a Content of type proposition, two daughters with non *nil* LEFT CORREL features, and does not inherit the LEFT CORREL value of its left Daughter.

We now look at the two subtypes of correl-clauses. The symmetric subtype inherits from coordinate phrases. We assume that coordinate phrases are n-ary un-headed phrases with a (optional) conjunction inside one (or more) conjunct(s), and shared features between mother and daughters:

(34)a	Coordinate-phrase \Rightarrow	$\boxed{\text{SYNSEM}} \boxed{\text{CONJ}} \boxed{\text{nil}}$ $\boxed{\text{DTRS}} \boxed{\text{list}([\text{CONJ} \boxed{\$} \text{nil}]) \oplus \text{list}([\text{CONJ} \boxed{0 \neq \text{nil}}])}$
-------	---------------------------------	--

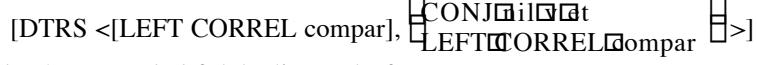
⁵ We follow Abeillé 2003, 2005 in analysing coordinate conjunctions as weak syntactic heads with a CONJ feature.

⁶. We follow Ginzburg and Sag (2000)'s analysis of verbless clauses as finite clauses (inheriting from hd-fragment-phrase).

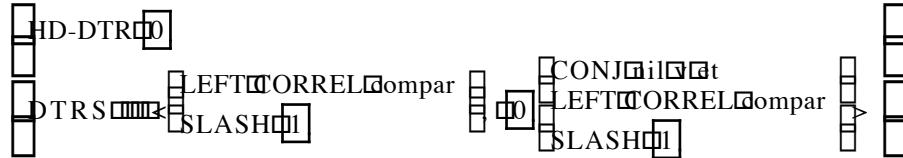


CCs inherit from correl-clauses and are defined as follows for French (35) and Spanish (36) :

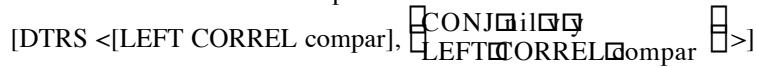
(35)a symmetric-CC-cl (A speakers) => correl-cl & coord-phr &



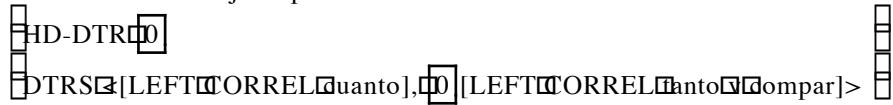
b asymmetric-CC-cl (B speakers) => correl-cl & hd-adjunct-phr &



(36)a symmetric-CC-cl => correl-cl & coord-phr &



b asymmetric-CC-cl => correl-cl & hd-adjunct-phr &



Conclusions

Comparative correlatives (CC) inherit from other properties in each language but require specific constructions. Two syntactic patterns are clearly available for Spanish, a symmetric one (with the conjunction 'y') which can be analysed as a particular case of coordinate construction, and an asymmetric one (with the specifier *cuanto*) which can be analysed as a particular case of subordinate construction (like English CC). French only has one CC construction, which behaves as a symmetric construction (with the conjunction 'et'), but with, for some speakers, a few asymmetric properties.

We conclude that two different syntactic patterns are needed for CC constructions crosslinguistically (contra Den Dikken 2005). Their semantics remains to be investigated.

References

- Abeillé A., 2003. A lexicalist and construction-based approach to coordination, S Muller (ed), Proceedings HPSG Conference, CSLI on-line publications.
- Abeillé A., Godard D. 2003. The syntactic flexibility of French degree adverbs, S Muller (ed), Proceedings HPSG Conference, CSLI on-line publications.
- Abeillé A., Bonami O., Godard D., Tseng J. 2005. Les syntagmes nominaux français de la forme de-N, Travaux de Linguistique, 50, 79-98
- Abeillé, A., Borsley R. 2006, La syntaxe des comparatives corrélatives en français et en anglais, Faits de langue, 38.
- Abeillé, A. 2005, Les syntagmes conjoints et leur fonction syntaxique, Langages, 160, p.42-66.
- Beck S. 1997. On the semantics of comparative conditionals, Linguistics and Philosophy, 20, 229-232.
- Bonami O., G. Boyé, J. Tseng. 2004. An Integrated Analysis of French Liaison. G. Jaeger (éd), Formal Grammar Proceedings HPSG Conference, CSLI on-line publications.
- Borsley R. 2004. An approach to English comparative correlatives, HPSG Conference, CSLI on-line publications.
- Culicover, P., R Jackendoff. 1999. The View from the Periphery: The English Comparative Correlative. *Linguistic Inquiry* 30.543-571.
- Den Dikken, M. 2005. Comparative correlatives comparatively, *Linguistic Inquiry* 36, 497-532.
- Ginzburg J., I. Sag 2000. Interrogative Investigations, CSLI Publications.
- Obenauer, H-G. 1983. Une quantification non-canonical : la quantification à distance, *Langue française*, 58, 66-88
- Sánchez López C. 2005 Correlaciones comparativas en español, Congreso Coreano de Hispanistas, Seúl
- Savelli M-J. 1995, Autant le dire, quelques éléments comparatifs sur la macro-syntaxe de plus/moins/autant, *Recherches sur le français parlé*, 13, 67-90.