

The Grammaticalization of Auxiliaries: Spanish Clitic Climbing

Author(s): John Myhill

Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society (1988), pp. 352-363

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

The grammaticalization of auxiliaries: Spanish clitic climbing

John Myhill

University of Michigan

0.0. Introduction: The synchronic study of grammaticalization. Studies of grammaticalization to date (e.g. Givón 1971, Heine and Reh 1984, Bybee and Pagliuca 1987) have focused on documenting and classifying types of grammaticalization processes using diachronic data (actual or reconstructed) of grammaticalization processes which have already taken place*. Typically, a grammatical morpheme is shown to be related to a word from which it has diachronically developed. Studies of this type have given us an idea of what types of words are likely to develop into what types of grammatical morphemes. However, the details of this process and the factors motivating it are still fairly obscure; this is because studies have focused on the beginning and endpoint of this process, when the morpheme in question may clearly be categorized as (respectively) lexical and grammatical. In between, the morpheme may be either lexical or grammatical, and the alternation between the lexical and grammatical uses represents a synchronic process of grammaticalization. Grammaticalization as a synchronic process has not been subjected to the sort of systematic study that grammaticalization as a diachronic process has, and so we know relatively little about it; however, the two processes represent two sides of the same coin, and we cannot understand grammaticalization without understanding both of them.

The present paper is intended to work toward an understanding of grammaticalization as a synchronic process. The specific phenomenon under study is the alternation between the following two types of constructions in Spanish:

- 1) Voy a verlo.
I-go to see-him
'I'm going to see him.'
- 2) Lo voy a ver.
him I-go to see
'I'm going to see him.'

In the construction exemplified in (1), the clitic pronoun *lo* is suffixed to the non-finite verb *ver* (in this case the non-finite verb is an infinitive but the situation is the same when it is a present participle). In the construction exemplified in (2), the clitic pronoun 'climbs' into the position before the finite verb; the process moving the pronoun to preverbal position has been referred to as 'clitic climbing'¹ (Napoli 1981) (it will be referred to as 'CC' in the remainder of this paper).

Rizzi 1976 argues that constructions like 1 involve a finite verb taking a non-finite complement in a lower clause while CC constructions like 2 involve a restructuring operation which turns the finite verb into an auxiliary and turns the infinitive or participle into the main verb². The entire structure *voy*

a *ver* is then under the V node, and the clitic *lo* is prefixed to this complex as it would be prefixed to a simple finite verb (cf. *lo vi*='him I-saw'='I saw him'). Rizzi presents a number of formal arguments for this analysis. On this view, the clitic does not really climb; the appearance of climbing is caused by the restructuring rule, although the term 'clitic climbing' has continued to be used in reference to this structure. In the non-climbing structure, on the other hand, there is no restructuring and the non-finite verb serves as a complement to the finite verb; the clitic accordingly attaches directly to the verb of which it is an argument (the non-finite one).

'Restructuring' involves a synchronic change from a verb to an auxiliary. The change from verb to auxiliary has been shown to be one stage in a diachronic grammaticalization process; in this process, the auxiliary eventually develops into a marker of tense, aspect, or modality. Diachronic changes of this type are well documented in a variety of languages (Heine and Reh 1984, Traugott 1972), and there have been a number of studies examining typical changes of this type; for example, the words forming the AUX class in English have all developed diachronically from main verbs (Traugott 1972), and they are now fairly far along the path to developing into markers of tense, aspect, and modality. Restructuring in Spanish is a synchronic process which is one step in a diachronic development of this type; studying it can therefore shed light on how this development takes place.

There have been quite a few studies of the formal properties of CC (e.g. Rizzi 1976, 1978, Aissen and Perlmutter 1976, Luján 1977, Suñer 1980, Napoli 1981), establishing what environments CC is at least possible in. However, these studies are only helpful to the study of grammaticalization in that they tell us when grammaticalization does *not* take place. Even when it is grammatically possible, CC does not always apply, and no previous study has examined the factors controlling whether or not CC actually applies in a situation where it is possible. In order to understand CC as a synchronic grammaticalization process, I have done a quantitative study of the factors affecting the actual use of CC in Spanish texts; the results of this study will be reported in section 2. First, however, I will turn in the next section to a brief review of relevant work done on CC and related topics.

1.0. Research of related topics. In this section, I will discuss research which is of direct relevance to my own study of CC.

1.1. The association between grammaticalization and certain meanings. Cross-linguistic studies of grammaticalized meaning have shown that certain meanings are much more likely than others to be expressed grammatically. The most extensive study of this type is Bybee 1985. Bybee looked at a sample of 50 languages selected as to be random and geographically and genetically diverse and determined the likelihood of various tense, aspect, and modality meanings being expressed grammatically through verbal inflections in these languages; she found that certain meanings are considerably more likely than others to be represented inflectionally.

A reasonable initial hypothesis to suggest is that the meanings most likely to be represented grammatically through inflections are also the meanings most likely to trigger a synchronic grammaticalization process such as the restructuring process associated with CC; meanings rarely associated with inflections should trigger restructuring less frequently, while meanings never associated with inflections should not allow restructuring at all³.

1.2. Earlier studies of the semantics of CC. It is clear that some finite verbs allow CC while others do not. For example, 1 and 2 show that *ir* allows CC, while 3 and 4 show that *insistir* 'insist' does not:

3) Insistí en comerlas.

I-insist on eating-them

'I insist on eating them.'

4) *Las insistí en comer.

Rizzi 1978 notes that the verbs allowing CC in Italian may be divided into three semantic categories, modals, aspectuals, and motion verbs, and Suñer 1980 makes a similar observation about Spanish; *insistir* (among many other verbs) does not fall into any of these categories and so does not allow CC. Lists of verbs allowing and not allowing CC may be found in Aissen and Perlmutter 1976 and Napoli 1981.

These lists are undeniably helpful in understanding CC; however, they only tell part of the story. For one thing, as pointed out by Suñer 1980 and Napoli 1981, there are a number of verbs for which speaker's grammaticality judgments vary considerably. Additionally (and more importantly for the present study), we will see that the three classes of verbs which uncontroversially allow CC, modals, aspectuals, and motion verbs, differ in terms of their frequency of use in CC as opposed to non-CC constructions, and there are even differences in CC frequency between different verbs in each of these classes; this has important consequences for the theory of grammaticalization. In order to appreciate the differences between these classes in terms of usage, it is necessary to do text counts, which will be the topic of the next section.

2.0. The quantitative study of CC in Spanish. A Spanish data base of several hundred pages was subjected to a quantitative analysis; the writings used in this study are listed in the appendix. Two factors were found to significantly affect the likelihood of CC, namely the semantic properties of the finite verb and the relative topicality of the subject and the clitic. 750 tokens were coded for the semantic type of the finite verb, and the results of this count will be reported in section 2.1; 543 tokens were coded for the relative topicality of the subject and the clitic, and the results of this count will be reported in section 2.2.

2.1. Effect of semantic type of finite verb on likelihood of CC. It seems reasonable to hypothesize that CC should be more likely when the finite verb represents a meaning which is commonly represented grammatically in the

languages of the world and less likely when the meaning of the finite verb is less likely to be represented grammatically. Looking first at verbs which do not allow CC in Spanish (e.g. *insistir*), we find that Bybee does not report any instances of the meanings associated with these verbs being represented inflectionally, which is what we would expect. Turning now to verbs which do allow CC, I chose 11 verbs which had meanings which Bybee found could be represented inflectionally and coded a total of 750 occurrences of these verbs when they took a non-finite complement with a clitic argument, in environments where CC is possible⁴; I then counted how often CC actually took place in these environments. The basic findings confirmed the hypothesis suggested in section 1.1; verbs with meanings which Bybee found to be likely to be expressed inflectionally favored CC, while verbs with meanings which were only rarely expressed inflectionally in Bybee's sample allowed CC, but only rarely. In this section, I will present and discuss these findings.

Spanish verbs which allow CC may be divided up into a number of semantic classes, and within each class the frequency of CC from one verb to another is fairly consistent. I will list these classes from the class most likely to take CC down to the class least likely to take CC. The constructions most strongly favoring CC have progressive meaning, using the verbs *estar* 'be', *ir* 'go', and *andar* 'walk', and *venir* 'come' plus the present participle (in this construction, the three motion verbs are bleached of much or all of their semantic content). Data illustrating this are given in table 1:

Table 1-Likelihood of CC in progressive constructions

	CC	non-CC	CC%
<i>estar</i> + pres. part.	75	9	89
<i>ir</i> + pres. part.	42	7	86
<i>venir</i> + pres. part.	5	1	83
<i>andar</i> + pres. part.	5	2	71

Each of these verbs may be used in non-progressive constructions, but in that case CC is not so strongly favored.

Ir 'go' may have future meaning when it is used followed by the preposition *a* and the infinitive; in this case it loses its meaning of motion and CC is quite common, as shown in table 2:

Table 2-Likelihood of CC in periphrastic future constructions

	CC	non-CC	CC%
<i>Ir</i> (<i>a</i>) (future meaning)	77	23	77

The next set of verb types all expressed the degree of commitment the speaker has to the proposition and may be loosely grouped together as 'epistemic modalities'. These include *ir* in its use as an irrealis marker (e.g. *¿Como te voy a olvidar?* = 'How you I-go to forget' = 'How would it be possible for me to forget you?'), *haber* followed by the preposition *de* and the infinitive (cor-

3.0. Conclusion. The findings of the present study support the formal analysis of CC in Rizzi 1976 in that the verbs characteristically used in CC constructions have meanings of the type which we would expect to be likely to be represented with more grammaticalized morphemes like auxiliaries. However, Rizzi's 1978 characterization of the semantic characteristics of the verbs allowing CC (modals, aspectuals, and motion verbs) has been shown to represent an oversimplification of the situation. Certain types of 'aspectuals' (e.g. progressives) strongly favor CC while others (inceptives) disfavor it. Among modals, epistemics favor CC while deontics disfavor it. These findings are what we would expect in the light of studies of the relationship of form and meaning such as Bybee 1985.

It is clear that the effect of the relative topicality of the subject and the clitic is secondary in importance to the semantic characteristics of the finite verb; table 9 shows that when the verb is of the type favoring CC (progressives, futures, and epistemics), CC is favored (although less so) even when the subject is higher than the clitic on the AH, while when the verb is of the type disfavoring CC (deontic modals, motion verbs, and inceptives), CC is disfavored (although less so) even when the clitic is higher than the subject on the AH.

Table 9-Effect of semantic type and AH

	Verbs favoring CC		
	CC	non-CC	CC%
Clitic outranks subject on AH	68	7	91%
Subject outranks clitic on AH	28	19	60%
	Verbs disfavoring CC		
	CC	non-CC	CC%
Clitic outranks subject on AH	23	54	30%
Subject outranks clitic on AH	12	113	10%

Nevertheless, the relative topicality of the subject and clitic has a fairly strong and statistically significant effect regardless of the semantic characteristics of the verb.

There is to my knowledge nothing that we know about the theory of grammaticalization which would predict the effect of the AH on the likelihood of CC; all of the research on grammaticalization which I know of focuses on the effect of the semantic characteristics of the grammaticalized element rather than the effect of the pragmatic characteristics of its arguments. The most plausible-sounding principle which I can suggest which might be incorporated into the theory of grammaticalization to explain this phenomenon is 'Grammaticalization of a verb is favored when the verb lacks grammatical relations with the topic of the sentence'. Thus when the clitic is more topical than the subject, the finite verb lacks grammatical relations with the most topical NP in the sentence; intuitively, it seems reasonable to suggest that this makes it a less prototypical verb, so that it is more likely to be syntactically demoted.

It is of course still unclear whether this principle, or anything similar to it, is indeed relevant to grammaticalization. Regardless of how this turns out, the effect of relative topicality on grammaticalization is clear. Additionally, there are a number of cases where neither relative topicality nor the semantics of the verbs can be responsible for the application or non-application of CC; even when the verb favors CC and the clitic is higher than the subject on the AH, CC still does not apply 9% of the time, and this suggests that there are still more factors affecting the likelihood of grammaticalization.

Diachronic studies of grammaticalization have uncovered quite a few patterns relating the semantics of a verb to its likelihood of being grammaticalized and its resulting function if it is grammaticalized. The present work suggests that synchronic study, because it looks at grammaticalization from a different perspective, will be able to uncover other factors affecting this process which have not thus far been discovered through diachronic research of this phenomenon.

Footnotes

*I thank Bill Croft, Peter Hook, and George Lakoff for their helpful comments on earlier drafts of this paper. This research for this paper was funded by a grant from Language Learning.

¹I am not including in this study constructions where the finite verb and the finite verb have different subjects (e.g. *Se la oí cantar* = 'him it I-heard sing' = 'I heard him sing').

²CC also occurs in Italian, where the relevant formal properties vis a vis arguments for the restructuring analysis are the same as in Spanish, and in fact Rizzi 1976 developed the 'restructuring' analysis of CC on the basis of data from Italian. Aissen and Perlmutter 1976 present arguments supporting a similar analysis for CC in Spanish; they refer to this phenomenon as 'clause reduction', but do not specify what part of speech the matrix verb becomes after this operation. I refer in the text here to Rizzi's analysis because viewing the finite verb in a CC construction as an auxiliary is crucial to the relevance of this construction to the theory of grammaticalization.

³It would of course be preferable if there existed a study of the likelihood of various meanings being expressed through auxiliaries in the languages of the world, as this would be more directly comparable to restructuring in Spanish, which creates an auxiliary rather than a verbal inflection. Unfortunately, I do not know of any studies of this type, and so a comparison with a study of the likelihood of different meanings being expressed inflectionally is the best possible alternative at the moment.

⁴Aissen and Perlmutter 1976 claim that epistemic *deber* (*de*) does not allow CC, while *deber* with the meaning 'ought, should' does allow it. If true, this would go against the theory of grammaticalization argued for here. However, it does not seem to be true. My data base included only two examples of *deber de* in an environment allowing CC, and CC occurred once, in the following

example:

...el río se debía de haber llevado, quién sabe desde cuándo,
 the river refl must have carried who knows since when
 el tamarindo que estaba en el solar de mi tía Jacinta...
 the tamarind tree that was in the grounds of my aunt J.

'The river must have carried off, who knows how long ago, the tamarind tree in my aunt Jacinta's yard.'

(Rulfo 1953:30)

⁵I excluded from my count instances where the non-finite verb had two clitic arguments, as in for example *Te lo voy a decir* = 'you it I-go to say' = 'I'm going to say it to you'. This was because it was unclear how to code such examples for the relative topicality of the subject and clitic (see section 2.2), since there are two clitics. I also excluded cases where lexical material intervened between the finite and non-finite verbs (aside from prepositions subcategorized by the finite verb) as this environment strongly disfavors CC.

⁶The following table gives the number of tokens for each of the cells in table 8:

		SUBJECT			
		2	1	3HS	3O
C	2	X	58	12	23
	1	29	X	58	52
L	3HS	9	53	X	68
	3O	23	82	66	X

Appendix

The following are the data sources used for the quantitative study:

- Benedetti, Mario. 1953. *Quién de nosotros*. Montevideo: Editorial Nueva Imagen.
- Fuentes, Carlos. 1962. *La muerte de Artemio Cruz*. Bogota: Fondo de Cultura Económica.
- García Marquez, Gabriel. 1974. *La mala hora*. Barcelona: Plaza and Janes.
- . 1981. *Crónica de una muerte anunciada*. Barcelona: Editorial Bruguera.
- Rulfo, Juan. 1953. *El llano en llamas*. Mexico City: Fondo de Cultura Económica.
- . 1955. *Pedro Páramo*. Mexico City: Fondo de Cultura Económica.
- Quiroga, Horacio. 1980. *Cuentos*. Mexico City: Editores Mexicanos Unidos.

References

- Aissen, Judith, and David M. Perlmutter. 1976. Clause reduction in Spanish. BLS 2.1-30.
- Andrews, J. Richard. 1975. *Introduction to Classical Nahuatl*. Privately

printed.

- Aronson, Howard I. 1982. *Georgian: A reading grammar*. Columbus, Ohio: Slavica.
- Bybee, Joan. 1985. *Morphology*. Amsterdam: John Benjamins.
- and William Pagliuca. 1987. *Cross-linguistic comparison and the development of grammatical meaning*. SUNY at Buffalo ms.
- Givón, Talmy. 1971. *Historical syntax and diachronic morphology: An archaeologist's field trip*. CLS 7.349-415.
- Gordon, Lynn. 1986. *Maricopa morphology and syntax*. Berkeley: University of California Press.
- Heath, Jeffrey. 1976. Substantival hierarchies. In *Grammatical categories in Australian languages*, ed. by R.M.W. Dixon, pp. 172-90. Canberra: Australian Institute of Aboriginal Studies.
- Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.
- Kuno, Susumu and Etsuko Kaburaki. 1977. Empathy and syntax. *Linguistic Inquiry* 8.627-72.
- Langacker, Ronald. 1974. Movement rules in functional perspective. *Language* 50.630-64.
- Luján, Marta. 1977. Clitic promotion and mood in Spanish verbal complements. Mimeo.
- Napoli, Donna Jo. 1981. Semantic interpretation vs. lexical governance: Clitic climbing in Italian. *Language* 57.841-87.
- Rizzi, Luigi. 1976. *Ristrutturazione*. *Rivista di Grammatica Generativa* 1.1-54.
- . 1978. A restructuring rule in Italian syntax. In *Recent transformational studies in European languages*, ed. by Samuel J. Keyser, pp. 113-58. Cambridge, Mass.: MIT Press.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical categories in Australian languages*, ed. by R.M.W. Dixon, pp. 112-71. Canberra: Australian Institute of Aboriginal Studies.
- Suárez, Jorge A. 1983. *The Mesoamerican Indian languages*. Cambridge: Cambridge University Press.
- Suñer, Margarita. 1980. Clitic promotion in Spanish revisited. In *Contemporary studies in Romance Languages*, ed. by Frank H. Nuessel Jr., pp. 300-30. Bloomington: Indiana University Linguistics Club.
- Traugott, Elizabeth. 1972. *The history of English syntax*. New York: Holt, Reinhart and Winston.
- Watkins, Laurel J. 1984. *A grammar of Kiowa*. Lincoln, Neb.: University of Nebraska Press.