# The Spanish Impersonal Construction: A Lexical Rule-Free Analysis of the Cliticization Process

*Natalia Brines*

This paper constitutes a contribution to the ongoing linguistic research concerning the status of lexical alternations in constraint-based linguistic frameworks. The treatment of such theoretical issue will be based on the analysis of the Spanish Impersonal Constructions, which will be analyzed in relation to their personal transitive counterparts, in which there is an explicit volitional agent carrying out the event denoted by the verb. There will two main issues to be considered: firstly, the fact that clitics are special words with special properties: their position in the sentence is fixed, they usually appear attached to a word, and generally, they cannot be stressed. Thus, a study of the nature of clitics will be carried out which, in accordance to recent works on this topic ((Miller and Sag 1997), (Monachesi 1995), (Zwicky and Pullum 1983)), will conclude that they are lexically generated units and, therefore, cannot be dealt with by means of syntactic rules. Secondly, and due to the well-known problematic nature of lexical rules in constraint-based frameworks, I will propose a lexical rule-free analysis of the lexical alternation exemplified above based on the use of underspecification with Minimal Recursion Semantics (MRS). The framework to be employed is Head-Driven Phrase Structure Grammar (HPSG) as presented in (Pollard and Sag 1994).

## 1 ON THE NATURE OF THE SPANISH CLITIC "SE"

Thus, (Monachesi 1995) and (Miller and Sag 1997), when applying a number of criteria to distinguish postlexical clitics and affixes as proposed in (Zwicky and Pullum 1983), reach the conclusion that these clitics are affixal in nature. When applying these criteria to Spanish clitics, the same conclusions are reached:

**Degree of Selection with Respect to the Host**: Like other Romance clitics, Spanish object clitics can only attach to verbs, even when the verb is not in VP initial position:

(1) Esta tarde le compro un regalo.
This evening I buy him/her a present.

(2) Se lo compro en el mercado.
I buy it for him/her in the market.

**Arbitrary Gaps in the Sets of Combinations**: Some clitic combinations are not possible in Spanish. In general, the me(acc-1st-sg)-te(dat-2nd-sg) pronoun combination is not acceptable:

(3) *Me(dat) te(acc) ha presentado.
me you has introduced.
He introduced you to me.

**Morphophonological Idiosyncrasies**: The third person dative pronoun changes its surface form when an accusative pronoun is also present:

(4) Le di un libro a Juan.
cl-dat-3rd-sg gave a book to Juan.
I gave a book to Juan.

(5) * Le lo di.
cl-dat-3rd-sg it(cl.acc-3rd-sg) gave.
I gave it to him/her.

(6) Se lo di.
cl-dat-3rd-sg cl-acc-3rd-sg gave.
I gave it to him/her.

**Rigid Ordering of Clitics**: The free order of overt NP objects contrasts with the rigid ordering of clitics:

(7) Le doy un libro a Juan.
I give a book to Juan.

(8) Le doy a Juan un libro.
I give to Juan a book.

(9) A Juan le doy un libro.
To Juan I give a book.

(10) Se lo doy.
cl-dat-3rd-sg cl-acc-3rd-sg I give.
I give it to him/her.

(11) $*$ Lo se doy.
I give it to him/her.

(12) $*$ Doy se lo.
I give it to him/her.

**Coordination**: Spanish clitics cannot generally have wide scope over coordination:

(13) $*$Le he encargado y he comprado un libro.
I have booked and bought a book for him.

(14) He encargado y he comprado un libro para Juan.
I have booked and bought a book for Juan.

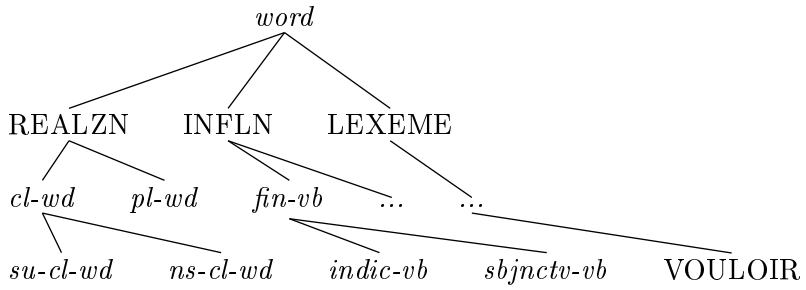
Thus, Spanish clitics behave very much in the same way as French and Italian clitics: they appear to behave as affixes when Zwicky & Pullum criteria are applied, which argues in favor of an inflectional approach to Spanish clitic formation.

## 2 A LEXICAL RULE-FREE ANALYSIS OF THE SPANISH IMPERSONAL ALTERNATION

The treatment of French pronominal affixes proposed by Miller & Sag derives from the assumption that they are morphological units and thus are not subject to syntactic rules.

By using monotonic inheritance, they treat cliticization as an inflectional process in the lexical hierarchy, without having to resort to the use of lexical rules. Their approach makes use of the ARG-ST list as a filtering device stopping certain verbal arguments from surfacing syntactically, making them instead morphologically attached clitics. This is achieved by ensuring that only ARG-ST elements which are not of type *affix* have a syntactic equivalent. Thus, the lexical hierarchy will develop entries from a number of cliticized types (*cl-wd* for which certain arguments in ARG-ST will be marked as affixes. Arguments of this type will not be allowed in either SUBJ or COMPS.

The hierarchy (15) of elements of type *word*:




(15)

is responsible for lexically generating object clitics through constraints associated with the types in the hierarchy, ensuring that affixes which appear in certain verb types as elements in ARG-ST are not coindexed with elements in SUBJ and COMPS:

(16)
$$_{pl\text{-}wd}\left[\text{SYNSEM}\left[\text{LOC} \mid \text{CAT}\left[\begin{array}{l}\text{VALENCE}\left[\begin{array}{l}\text{SUBJ } \boxed{2}\\ \text{COMPS } \boxed{3}\end{array}\right]\\ \text{ARG-ST } \left\langle \boxed{2},\boxed{3}\right\rangle\end{array}\right]\right]\right]$$

(17)
$$_{cl\text{-}wd}\left[\text{SYNSEM}\left[\text{LOC} \mid \text{CAT}\left[\begin{array}{l}\text{HEAD verb}\\ \text{VALENCE}\left[\begin{array}{l}\text{SUBJ } \left\langle \boxed{2}\right\rangle\\ \text{COMPS } \boxed{3}\end{array}\right]\\ \text{ARG-ST } \left(\boxed{2}\,\boxed{3}\right)\text{O nelist}\left(\textit{aff}\right)\end{array}\right]\right]\right]$$

(18)
$$_{su\text{-}cl\text{-}wd}\left[\text{SYNSEM}\left[\text{LOC} \mid \text{CAT}\left[\begin{array}{l}\text{VALENCE}\left[\text{SUBJ } \langle\rangle\right]\\ \text{ARG-ST } \left\langle\left[\text{AFF,NOM}\right],\ldots\right\rangle\end{array}\right]\right]\right]$$

(19)
$$_{ns\text{-}cl\text{-}wd}\left[\text{SYNSEM}\left[\text{LOC} \mid \text{CAT}\left[\text{VALENCE}\left[\begin{array}{l}\text{SUBJ } \left\langle \boxed{1}\right\rangle\\ \text{ARG-ST } \left\langle \boxed{1},\ldots\right\rangle\end{array}\right]\right]\right]\right]$$

From the hierarchy in (15) above we can see that there are two types of verbal realization: *pl-wd* (plain word), ensuring that every member in the ARG-ST list is mapped into an overt phrase, and *cl-wd* (cliticized word), for which at least one of the elements in their argument structure will be realized affixally, and therefore will not be mapped into the syntax.[1].

As mentioned, this analysis is limited to cases in which the clitic is an argument of the verb, and consequently an element in its ARG-ST. Setting constraints on the linking between ARG-ST and valence features ensures that clitics are treated as lexically generated elements which appear in ARG-ST at a "presyntactic" level.

There are, however, a number of constructions in Spanish and other Romance languages in which the clitic is not a verbal argument but a grammatical marker triggering a semantic change on the verb and an argument reduction process. The Spanish Impersonal Construction falls under this category. A sentence such as:

(20) Se vende libros.
Books are sold.

expresses a change of state undergone by the verbal argument, as opposed to the transitive variant, which expresses a causative change of state carried out by a volitional agent (Pedro vende libros - Pedro sells books).

Extending Miller & Sag's analysis to cases in which the clitic is merely a grammatical marker reducing the verb's subcategorization by altering its semantic structure cannot be done by simply extending their treatment of object clitics. The reason for this is the fact that these markers do not mirror a verbal argument and therefore are not members of the ARG-ST list, thus making it impossible to deal with them by using constraints on the type of the elements of ARG-ST. Furthermore, the semantic change triggered by the presence of the clitic cannot be dealt with by means of traditional monotonic inheritance mechanisms.

[1] We will not comment here on the morphological form and the realization of pronominal affixes. In Miller & Sag(1997) this is taken care of by means of two features: FORM and I-FORM, and a function FPRAF from the I-FORM, the HEAD value and the ARG-ST list into the FORM value. See Miller & Sag for details.

## 2.1 Minimal Recursion Semantics (MRS)

Minimal Recursion Semantics (MRS) is a meta-language for describing semantic structures which provides us with a very productive use of underspecification and a structured description of partial semantic information.

(Copestake, Flickinger, Pollard, and Sag 1998) introduce the basis for the theory and apply it to scope phenomena in such a way so that an ambiguous structure can get different scope readings depending on the instantiation of certain variables. In the same way, and since we need a more flexible semantic ontology in order to deal with the alternation processes under study, we will see how an MRS structure is adequate for an analysis of diathesis phenomena based on inheritance mechanisms.

They state the basis of the theory by proposing a mechanism for dealing with both real scope and its possible ambiguities. They introduce a number of variables in a flat semantic representation[2] which "can be thought of as handles which enable us to "grab" particular propositions in the flat list"[3] and which will carry the necessary scope information so that all and only the correct readings are accounted for.

The main advantage of this approach is the possibility of using variable underspecification in order to account for scope ambiguity. This property of MRS is the one we will exploit below in order to analyse the Spanish Impersonal Alternation.

Let us exemplify the basics of the theory. The sentence "Every dog chased some cat" can have two different readings depending on the quantifier that is given wide scope. The use of variable underspecification will reflect this ambiguity and will allow the two possible scope readings to be derived:

(21) 1:every(x,3,n), 3:dog(x), 7:cat(y), 5:some(y,7,m), 4:chase(e,x,y).

[2]We will not discuss here the motivations in favor of a flat semantic representation. For a description of these see (Copestake, Flickinger, Pollard, and Sag 1998)

[3](Copestake, Flickinger, Pollard, and Sag 1998)

Within HPSG, MRS semantic structures will be defined in terms of feature structures. An MRS representation within HPSG "consists of a structure of type *mrs-str*, with appropriate features HANDEL and LISZT, which take values of type *handle* and *list* respectively. (...) The value of LISZT is defined to be a flat list of *rels* (relations) which all have HANDELS and other features depending on their type"[4].

The structure in (21) above can therefore be converted into the following HPSG structure, where the CONTENT value of "every dog chased some cat" would look as follows:

(22)
$$\begin{bmatrix} \text{HANDEL } \textit{handle} \\ \text{INDEX } \boxed{8} \\ \text{LISZT } \left\langle \begin{bmatrix} \textit{every\_rel} \\ \text{HANDEL} \quad \textit{handle} \\ \text{BV } \boxed{2} \\ \text{RESTR } \boxed{3} \\ \text{BODY} \quad \textit{handle} \end{bmatrix}, \begin{bmatrix} \textit{dog\_rel} \\ \text{HANDEL } \boxed{3} \\ \text{INST } \boxed{2} \end{bmatrix}, \begin{bmatrix} \textit{some\_rel} \\ \text{HANDEL} \quad \textit{handle} \\ \text{BV } \boxed{6} \\ \text{RESTR } \boxed{7} \\ \text{BODY} \quad \textit{handle} \end{bmatrix}, \right. \\ \left. \begin{bmatrix} \textit{cat\_rel} \\ \text{HANDEL } \boxed{7} \\ \text{INST } \boxed{6} \end{bmatrix}, \begin{bmatrix} \textit{chase\_rel} \\ \text{HANDEL} \quad \textit{handle} \\ \text{EVENT } \boxed{8} \\ \text{ACT } \boxed{2} \\ \text{UND } \boxed{6} \end{bmatrix} \right\rangle \end{bmatrix}$$

The underspecification of the value of the HANDEL feature in the outermost structure and the feature structures associated with the quantifiers (*every_rel* and *some_rel*) will make sure that two different readings develop with different quantifier scope. These two readings will result from the two possible bindings of the outermost HANDEL in the CONTENT structure: it can be bound both to the HANDEL value of the feature structure of type *every_rel* or to the one of type *some_rel*.

## 2.2 The Spanish impersonal alternation

The constraints on verbs who can enter the Impersonal construction is that they must be dyadic verbs whose agent is or can be[5] volitional.

The example below is an instance of a verb which can appear in impersonal form:

[4] Copestake et al. (1995)

[5] Verbs underspecified for volitionality, such as "abrir" (to open), can also appear in impersonal structures

(23) Nosotros vendimos libros en el parque.
We sold books in the park.

since it satisfies the requirements of being a dyadic verb with a volitional agent. Therefore, the equivalent impersonal sentence can be constructed:

(24) Los libros se venden en el parque.
The books are sold in the park.

Let us consider the entry for the verb "vender" (to sell), whose MRS semantic structure has an unbound variable:

(25)

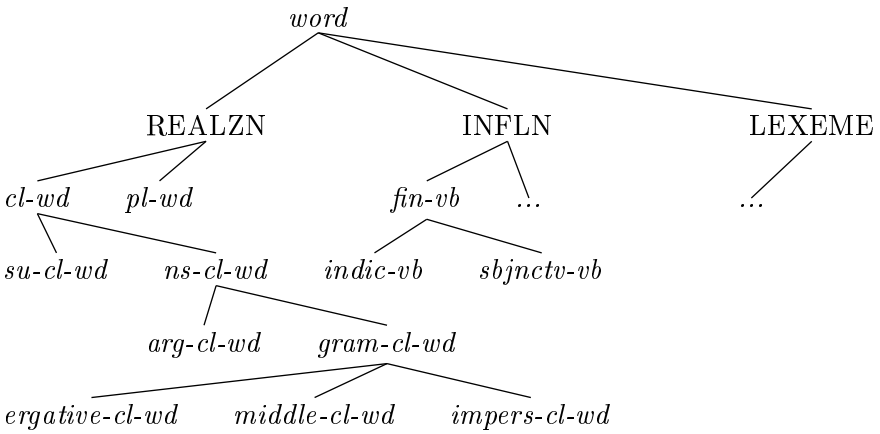
$$\begin{bmatrix} \textit{vender} \\ \text{SS}\,|\,\text{LOC}\,|\,\text{CONT} \begin{bmatrix} \textit{handeld-str} \\ \text{HANDEL}\ \_ \\ \text{INDEX}\ \boxed{10} \\ \text{LISZT} \left\langle \begin{bmatrix} \textit{cause} \\ \text{HANDEL}\ \boxed{1} \\ \text{BODY}\ \boxed{5} \\ \text{ACTOR}\ \boxed{3}\ \textit{volitional} \\ \text{CAUSE}\ \boxed{4} \end{bmatrix}, \begin{bmatrix} \textit{cause} \\ \text{HANDEL}\ \boxed{5} \\ \text{BODY}\ \boxed{9} \\ \text{UND}\ \boxed{6} \\ \text{BECOME}\ \boxed{7} \end{bmatrix}, \begin{bmatrix} \textit{state} \\ \text{HANDEL}\ \boxed{9} \\ \text{BODY}\ \boxed{8} \\ \text{STATE}\ \boxed{10}\ \textit{sold} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The constraint on *impers-cl-wd* ensures that impersonal verbs have the outermost HANDEL variable bound to the HANDEL in the second element in LISZT:

(26)

$$\begin{bmatrix} \textit{impers-cl-wd} \\ \text{SS}\,|\,\text{LOC}\,|\,\text{CONT} \begin{bmatrix} \textit{impersonal} \\ \text{HANDEL}\ \boxed{5} \\ \text{LISZT} \left\langle \begin{bmatrix} \text{HANDEL}\ \boxed{1} \\ \text{BODY}\ \boxed{5} \\ \text{ACTOR}\ \boxed{3}\ \textit{volitional} \end{bmatrix}, \begin{bmatrix} \text{HANDEL}\ \boxed{5} \\ \text{BODY}\ \boxed{9} \end{bmatrix}, \ldots \right\rangle \ldots \end{bmatrix} \end{bmatrix}$$

This constraint will be part of the lexical hierarchy. Thus, in order to account for cliticized verbs with non-argument clitics, we will extend the hierarchy for *word* proposed by Miller & Sag so that cliticized verbs with grammatical markers are allowed:




(27)

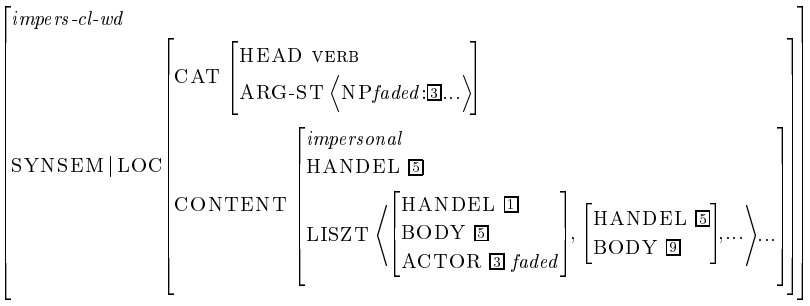
Thus, impersonal constructions are developed as a subtype of cliticized words inheriting the constraints from the type *gram-cl-wd*. This type will be defined below.

In order to stop semantic arguments which are outside the scope of the outermost HANDEL from surfacing syntactically, we will propose a subtype of *synsem* called *faded* with the following constraint associated with it:

(28) $$_{faded}\left[\text{LOC} \mid \text{CONTENT}\left[\text{INDEX } faded\right]\right]$$

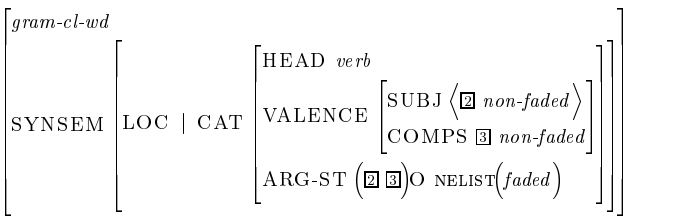
Any structure whose index is marked as faded will have a synsem of type *faded*. Impersonal constructions will always have the semantic argument outside the scope of the outermost HANDEL marked as faded. Thus, we can rewrite (26) in the following way:

(29)




By means of unification processes, these arguments will appear to have a synsem of type *faded* in their corresponding ARG-ST list variants. We must now ensure that they do not surface into the syntax. By rewriting the constraint on the type *gram-cl-wd*, of which *impers-cl-wd* is a subtype, we will ensure that the COMP and SUBJ lists are made up of non-faded elements, thus stopping *faded* elements in ARG-ST from manifesting in the valence features[6]:

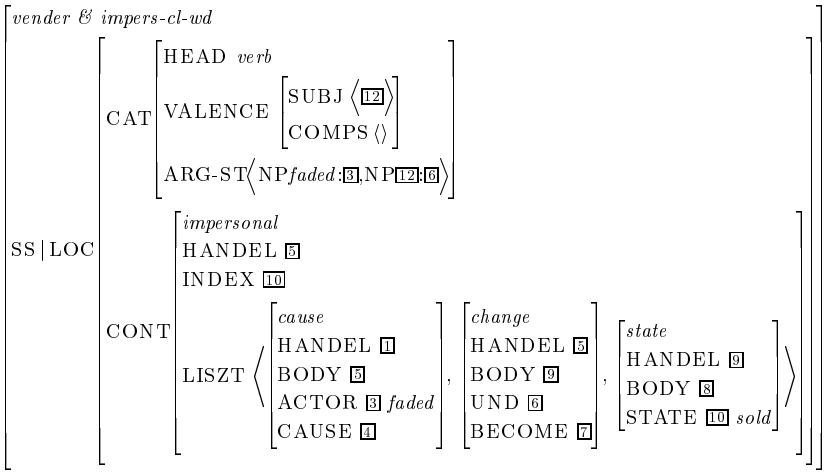
(30)




[6] *non-faded* is, like *faded*, a subtype of *synsem*, and it has the *canonical* synsem as its subtype

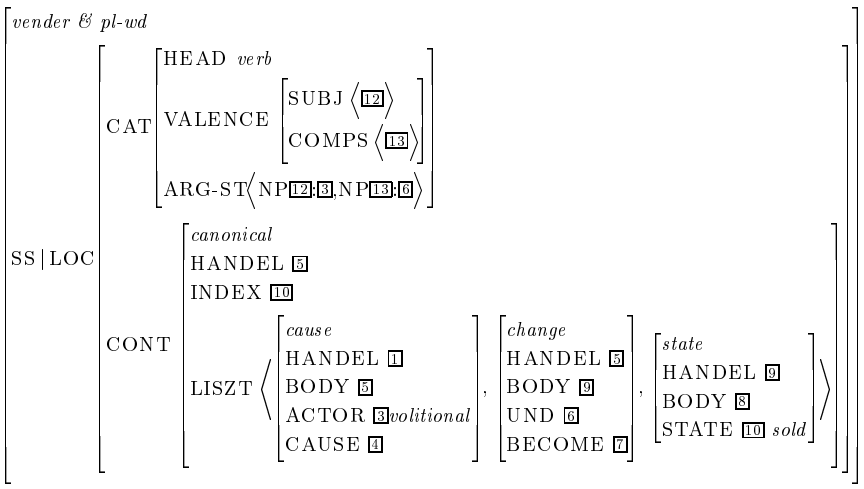
The Impersonal form of “vender” is thus constructed as follows:

(31)




The transitive variant will develop from the meet of the lexeme and the type *pl-wd*, which has a canonical CONTENT binding the outermost HANDEL variable to the first feature structure in LISZT:

(32)



I can conclude then that horizontal relations within the type hierarchy can be analysed using monotonic inheritance mechanisms when an MRS approach to semantics is adopted. The Spanish Impersonal Construction is an example of the problematic nature of cliticization processes in which the clitic is not an argument of the verb, but instead it triggers a semantic change that has to be encoded in the different constructions a verbal entry can enter. By mens of underspecification processes implemented in HPSG through MRS semantics, lexical rules are no longer needed for a correct treatment of these constructions.

## References


Copestake, A., D. Flickinger, C. Pollard, and I. Sag (1998). Minimal recursion semantics: An introduction.

Miller, P. and I. Sag (1997). French clitic movement without clitics or movement. *Natural Language and Linguistic Theory*.

Monachesi, P. (1995). *A Grammar of Italian Clitics*. TILDIL Dissertation Series.

Pollard, C. and I. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Zwicky, A. and G. Pullum (1983). Cliticization vs. inflection: English n't. *Language 59*(3), 502–513.



*Natalia Brines* is a member of the CL/MT Group at The University of Essex (UK). Her main research topic is the status of lexical processes within constraint-based formalisms, focusing on HPSG. Current Address: Department of Language and Linguistics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. e-mail: *nbrine@essex.ac.uk*.

A variation of this paper concentrating on the causative/inchoative alternation was presented at ESSLLI 99.