A Computational Model of the Spanish Clitic System

Luis A. Pineda and Ivan V. Meza

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS)
Universidad Nacional Autónoma de México (UNAM)
{luis, ivanvladimir}@leibniz.iimas.unam.mx

Abstract. In this paper a computational model of the Spanish clitic system is presented. In this model clitic pronouns receive a dual analysis in which enclitics are considered inflexions while most proclitics are considered independent lexical units, hence proper clitics. The model covers the analysis of simple periphrases that, in addition to auxiliary and modals, have a single content verb (e.g. *puede comérselo*, *se lo ha querido comer*) and also the analysis of complex periphrases with more than one content verb (e.g. *le hubiera visto comérsela*, *se la hubiera visto comer*). The model introduces three operations on clictis: cancellation, composition and subsumption, and is formalized in Head-driven Phrase Structured Grammar; the standard machinery of this theory is extended with one combination scheme, the head-proclitic rule, and one principle, the clitic principle, that is satisfied by Spanish clitic sentences. A computational implementation of the theory with the Linguistic Knowledge Building (LKB) tool is also reported.

1 Introduction

Intuitively, a clitic is an unstressed particle that is attracted to a stressed word, its phonological host, and the resulting object is perceived as lexical unit[1]; unlike inflexions and derivations, that are assembled with their stems at the morpho-lexical level of linguistic representation, clitics are combined with their host at the syntactic level. According to Zwicky and Pullum (1983, pp- 503):

> "…word-clitic combinality is largely governed by SYNTACTIC considerations. The conditions governing the combinability of stems with affixes are of quite a different sort: they are MORPHOLOGICAL and/or LEXICAL in character, being concerned with the substructure of a finite set of words"

However, it is not always clear what is the linguistic level of representation for a given particle; in order to make this distinction Zwicky and Pullum (*ibid.*) advanced a number of criteria that we summarize as follows: (1) inflexions attach to words of specific syntactic categories while clitics do not exhibit this restriction, so clitics can attach to words of different categories and they often do so, (2) the combination host-clitic is very regular while inflexions show exceptions, (3) the meaning of clitic-host

[1] See, for instance, the introduction of Nevis (1991).

A. Gelbukh (Ed.): CICLing 2005, LNCS 3406, pp. 73 – 82, 2005.

combinations is the same as the meaning of expressions that show no such reduction (e.g. *she is gone* means the same as *she´s gone*) but inflexions do show idiosyncrasies, (4) cliticizised forms cannot be affected by syntactic operations, while affixed words can (e.g. no syntactic rule treats *I´ve* as a constituent[2]) and (5) clitics can attach to combinations already cliticisized, but inflexions cannot attach to already inflected words. Following these criteria Miller and Sag (1995) and also Abeillé *et al.* (1996) have classified French clitic pronouns as inflexions (*pronominal affixes* in Miller and Sag´s terminology) and Monachesi (1999) has adopted a similar criteria for Italian; however, the case for Spanish is not that clear: according to (1), and perhaps (2), clitic pronouns behave more like inflexions; according to (3) clitics present a dual behaviour, and according to the other three they behave more like clitics[3]. These criteria reflect a further implicit intuition about the architecture of the grammar and assume that the morpho-lexical and syntactic levels of representation are independent, and that the internal structure of units assembled in the former level (i.e. words) cannot be altered or broken down by syntactic operations. Consequently, if the combination takes place at the syntactic level, the resulting unit is a pseudo-word, or rather a clitic-host combination.

From this consideration, a common test to distinguish clitics from affixes is whether the particle can have a wider scope over coordination (point (4) in the list above): if the pronouns are inflexions assembled with the verb by a morphological operation, they cannot be factored out in coordination operations. However, in Spanish, *lo llevó y lo puso sobre la mesa* (he/she took it and put it on the table), for instance, can also be expressed as *lo llevó y puso sobre la mesa*, which is grammatical and has the same meaning. In other cases the grammaticality of the second form is marginal, as in *le gusta y quiere* (she likes him and loves him) and in others the construction is clearly ungrammatical as shown by *te vas o te quedas* (you go or you stay) versus **te vas o quedas*. The rule seems to be that when the pronoun substitutes the direct or indirect complement of a transitive verb, it can appear either next to their verbal host within a coordination or move out from this construction as a single realization; if the pronoun appears next to an intransitive verb, on the other hand, it cannot be moved out and has to be realized attached to its phonological host. In this latter case it behaves like an inflexion.

Further evidence about the realization of some proclitics as words is provided by interruptions and repairs in spontaneous speech; in our corpus, forms like *me…muéstrame otra vez los muebles* (to-me … show-me again the furniture) appear often (Villaseñor *et al.*, 2001; Pineda *et al.*, 2002); despite that words can be interrupted in inter-syllable positions, we have observed no cases in which the interruption splits off a stem from its inflexion. Accordingly, if the proclitic were an inflexion it could not be split off after lexical realization.

On the basis of these considerations, we propose a dual analysis for clitic constructions: on the one hand enclitics are considered inflexions, but proclitics that represent normal complements of verbs are considered independent lexical units, which combine with their phonological host in the syntax and are proper clitics; on

[2] Although this cannot be ruled out altogether if surface structure and intonation receive an incremental integrated analysis, as in Categorial Grammar (Steedmann, 1991).

[3] See also Klavans (1985).

the other hand, clitic pronouns that substitute complements with an idiosyncratic character (e.g complements to intransitive verbs, reflexive and pseudo-reflexive verbs, some ethical datives, and the adjectival phrases in attributives), either proclitics or enclitics, are considered inflexions.

2 The Basic Model

In the basic form of the phenomenon clitic pronouns substitute the direct and indirect object of verbs by accusative and dative pronouns that appear next to verb by its right or left side, forming the enclitic and proclitic constructions respectively. In simple clitic sentences there is only one verb of content, and the clitic pronouns substitute its arguments. Also, in non-periphrastic constructions the verb is both the cliticisized object and the phonological host. For instance, in *el padrino le sirve una copa al muchacho, y* ***éste se la da a la novia***[4] (the best man pours the glass to the boy, and he gives it to the bride[5]), the pronouns *se* and *la* substitute the direct and indirect objects of the verb *da*/gives (i.e. *una copa* (the glass) and *la novia* (the bride) respectively); also, the clitic *se* is a duplication of the explicit realization of the complement. The examples (1) illustrate the "standard" sentence of the previous example and a set of possible variations including clitic pronouns.

(1) a. *El padrino da* [*la copa*]$_i$ [*a la novia*]$_j$
The best man gives the glass$_i$ to the bride$_j$
b. *dala*$_i$ [*a la novia*]$_j$
c. *dale*$_j$ [*la copa*]$_i$
d. *dase*$_j$*la*$_i$
e. *dase*$_j$*la*$_i$ [*a la novia*]$_j$
f. *la*$_i$ *da* [*a la novia*]$_j$
g. *le*$_j$ *da* [*la copa*]$_i$
h. *se*$_j$ *la*$_i$ *da*
i. *se*$_j$ *la*$_i$ *da* [*a la novia*]$_i$

However, when clitics occur in periphrases, the phonological host can be an auxiliary or modal verb[6] different from the cliticisized one as in ***el post no lo he podido escribir por la mañana***[7] (I have not been able to write the post in the morning); we give two alternative realizations of this sentence in (2); although in (2.b) the cliticisized verb *escribir* (to write) is also the phonological host, in (2.c) the cliticisized verb and the phonological host (i.e. *haber*[8]) are different.

[4] The main examples in this paper were extracted from the internet, which we consider our corpus for the present paper. Other sentences sequences (1) to (4) are variants of the reference one that are acceptable for native speakers.

[5] http://omega.ilce.edu.mx:3000/sites/litinf/huasteca/html/sec_45.htm

[6] We adopt Gili Gaya's terminology and call modal verbs to intentional verbs appearing in periphrasis.

[7] http://blogs.ya.com/vivirsintabaco/

[8] In our model, auxiliary verbs are subject raising as they are not agentive, and their syntactic subject is the same as the subject of its complement, which is a verbal phrase; similarly, modals, like *querer* are subject control, as they also share their subject with their verbal phrase complements, although these latter forms are agentive (Pineda and Meza, 2004).

(2) a. *No he podido escribir [el post]*$_i$
Not have been-able to-write the post
I have not been able to write the post
b. *No he podido escribirlo*$_i$
c. *No lo he podido escribir*$_i$

For this reason we distinguish between the clitic host, the cliticisized verb, from the phonological host, and we say that in a well-formed clitic sentence the pronouns attached to the phonological host cancels the corresponding arguments of the clitic host. Following Miller and Sag (*ibid.*) and Monacheci (*ibid.*), we consider cliticizised verbs as valence reduced realizations of their basic forms, which require overt complements. We define cliticization as a lexical operation on the basic form of verb; this operation removes the cliticisized arguments from its complements list, and places them in a *clitic-lists* attribute which, in conjunction with the subject and complement attributes, defines the valence of verbs. Our approach has a lexical orientation and we postulate no movement, traces or empty categories, and non-local dependencies are captured through structure sharing, as commonly done in categorical and unification formal approaches to grammar. The model is framed in HPSG (Pollard and Sag, 1994; Sag and Wasow, 1999), and cancellation operations are defined through the standard combination principles of this theory (e.g. head-complement rule, head-specified rule, the GAP principle, etc.). For clictic cancellation to take place, the clitic host must be within the scope of the phonological host (e.g. *pudo verlo comersela* versus **la pudo verlo comerse*) as will be illustrated below.

Clitic pronouns sequences present a rigid and idiosyncratic order that poses a challenge to the analysis of the phenomenon. In our model we postulate that there is a clitic lexicon which codifies all clitic sequences that occur in a dialect, with the corresponding order and case information, and there is an entry in the clitic lexicon for each sequences of one, two or possible three pronouns; clitic pronouns have a default case (e.g. *lo* and *la* are accusative and *le* and *se* dative) but they can be used with a different case (e.g. *le* and *se* can be accusative given rise to the so-called *leísmo*) and we define an entry in the clitic lexicon for each sequence of pronouns with a different case assignment. This approach permits to analyze simple clitic sentences in terms of a single cancellation operation. We distinguish three cases: (a) simple lexical cancellation, (b) composite lexical cancellation and (c) syntactic cancellation. Simple lexical cancellation is defined in terms of a lexical rule that implements cliticization and performs the insertion of the pronouns in a single operation, permitting the analysis of (1b-1e) and (2b), for instance. Composite lexical cancellation is defined in terms of two lexical rules: one implements the cliticization operation on the clitic host, and the other performs lexical insertion on the phonological host if structure sharing between the clitic lists of both the clitic and phonological hosts is permitted (i.e. through the head-complement rule), as in (*la reina pudo haberlo visto y escuchado/*the Queen could have seen it and listened it). Finally, syntactic cancellation is analyzed in terms of the lexical rule that cliticisizes the host, and the head pro-clitic rule that combines an entry in the clitic lexicon with a verbal phrase if the structure of the clitic list attribute of the predicate corresponds with the structure of the sequence in the clitic lexicon (e.g 1f-1i and 2c); this rule captures the intuition that proclitics are proper clitics.

3 Complex Periphrasis

The model presented so far follows closely Monachesi´s analysis for Italian, with the exception of the use of the head-proclitic rule whose corresponding effect in Monachesi´s affixial approach is achieved through lexical rules; however, the analysis of the Spanish complex periphrases with more than one content verb motivates further our dual analysis. In *se lo oi decir en varios reportajes*[9] (I hear him to say it in several interviews) the subjects of the two content verbs are different (the speaker is the one who listens but a third party is the one who says it); in addition, the syntactic object of *oí* (hear) is shared with the subject of *decir* (to say) and the composite verbal phrase *oi decir* has a composite direct object "*se lo*". Examples (3) presents the "standard" non-cliticisized sentence and some of its clitized variations:

(3) a. *Oí* [*a el*]$_i$ *decir* *[el comentario]*$_j$
hear to him$_i$ to-say the comment$_j$
I hear him to say the comment
b. **Oilo*$_i$ *decirlo*$_j$
c. *Oyélo*$_i$ *decirlo*$_j$
d. **Oyélo*$_i$*lo*$_j$ *decir*
e. *Oyése*$_i$*lo*$_j$ *decir*
f. *Le*$_i$ *oi decirlo*$_j$
g. *Se*$_i$ *lo*$_j$ *oí decir*

In this sequence, the clitict *se lo* occurs as an enclitic in (3e) but as a proclitic in (3g). In this case, both of the pronouns are in the accusative (i.e. substitute direct objects) and *se* is used instead of *le* (with *leísmo*) or *lo*, as no sequence of two *l*´s pronouns is allowed in Spanish (e.g. 3.d). The sequence shows that two clitic hosts can compose their accusative clitizations if they are next to each other (i.e. accessible), and the result of this operation is composite clitic argument. We refer to this operation as clitic composition. This operation is implemented through lexical rules and structure sharing, and clitic sentences of this form are also analyzed in terms of single cancellation. The ungrammaticality of (3b) is due to an idiosyncratic lexical restriction of Spanish for the phonological host, as participles and finite forms (but imperatives) cannot have enclitics, while infinitive, imperatives and gerunds require enclitics always.

The composition operation illustrated in (3) "builds" a clitic word in which all constituting pronouns have a different referent; however, this is not always the case. In ***la vi comiéndose la mesa fría*** *con los ojos*[10] (I saw you/her eating the cold table with the eyes) the verb *comer* (to eat) has an idiosyncratic dative complement that co-refers with its subject, forming an ethical dative that marks that the subject of this action is also its beneficiary. We present some variations of this sentence in (4):

[9] http://www.carp.org.ar/eng/idolos.php3

[10] http://www.mundomatero.com/chistes/junio2000.html

(4) a. *Vi* [*a usted*]$_i$ *comiendo* [*la cena*]$_j$ [*por/para usted*]$_i$
see to you$_i$ eating the dinner$_j$ for you$_i$
I see you eating the dinner for you own sake
b. *Vi* [*a usted*]$_i$ *comiendose*$_i$ [*la cena*]$_j$
c. *Vi* [*a usted*]$_i$ *comiendose*$_i$*la*$_j$
d. **Víla*$_i$ *comiendose*$_i$*la*$_j$
e. *Vela*$_i$ *comiendose*$_i$*la*$_j$
f. **Vela*$_i$*+se*$_i$*la*$_j$ *comiendo*
g. *Vese*$_i$*la*$_j$ *comiendo* (i.e. *se*$_i$ = *la*$_i$*+se*$_i$)
h. *Se*$_i$ *la*$_j$ *vi comiendo*
i. *La*$_i$ *vi comiendose*$_i$*la*$_j$

Sentences (4a) does not really occur in the language and it is used only as an aid to illustrate the meaning of (4b) in which *comer* has already the dative reflexive *se* as enclitic; the clicitization of the direct objects of *vi* and *comiendo* gives rise to the composite predicate *vi comiendo,* with a composite direct object represented by *la*$_i$*+se*$_i$*la*$_j$. However, in this composition the object of *visto* co-refers with the dative *se* of *comiendo*, and the redundant form *la*$_i$*+se*$_i$ is reduced as *se*$_i$, with the dative case prevailing, and the remaining *se*$_i$*la*$_j$ form represents the whole of the composite clitic argument as shown (4g) and (4h) in the enclitic and proclitic forms respectively. We refer to the reduction of this argument, in which an accusative pronoun is subsumed by a co-indexed dative form, as clitic subsumption. If the co-indexed arguments have the same case, they can also be subsumed in a composition. The analysis of sentences with clitic subsumption is carried out with a single cancellation operation, and the ungrammaticality of (4d) is due to the lexical restriction on participles and finite forms for enclitics. (4f) shows, in addition, that two co-indexed pronouns cannot occur next to each other, and subsumption is obligatory, as shown in (4.g).

Next we illustrate the analysis of (4.h). The lexical entry of the word "*se la*" in the clitic lexicon is shown in Figure 1. This entry has a local *synsem* attribute with the attributes of category CAT and the restriction of the semantic content *CONT/RESTR*. Also the head value of this entry is *clitic*.

$$\begin{bmatrix} se\quad la & \\ SS\,|\,LOC & \begin{bmatrix} CAT & \begin{bmatrix} HEAD & clitic \\ VAL\,|\,CL-LIST & \langle NP-dat_i,\quad NP-acc_j \rangle \end{bmatrix} \\ CONT\,|\,RESTR & \{[person-rel]_i,[female-rel]_j\} \end{bmatrix} \end{bmatrix}$$

Fig. 1. Clitic word

Now, we come to the cliticization of *comiendo* (the gerund of *comer*/to eat). The basic lexical entry for the verb *comer* is illustrated in Figure 2. The lexical rule that cliticisizes the verb is shown in Figure 3; in addition to including the direct object in the *CL-LIST*, this rule also adds an idiosyncratic extra complement in the *CL-LIST*, with a dative case (i.e. *se*), which is co-indexed with its subject, producing the reflexive connotation of the ethical dative.

$$\begin{bmatrix} tr-vrb & \\ SS \mid LOC & \begin{bmatrix} CAT & \begin{bmatrix} ORTH & comer \\ HEAD & verb \\ VAL \mid CL-LIST & \langle\ \rangle \\ ARG-ST & \langle NP_i, NP_j \rangle \end{bmatrix} \\ CONT \mid RESTR & \left\{ \begin{bmatrix} show-rel & \\ ACT & i \\ UND & j \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$

Fig. 2. Lexical entry for *comer*

$$\begin{bmatrix} ORTH & \#1 \\ HEAD & \#2 \& \begin{bmatrix} verb & \\ FORM & fin \end{bmatrix} \\ VAL \mid CL-LIST & \langle\ \rangle \\ ARG-ST & \langle \#3 \rangle \oplus \#a \end{bmatrix} \mapsto \begin{bmatrix} ORTH & \#1 \\ HEAD & \#2 \\ VAL \mid CL-LIST & \#a \\ ARG-ST & \langle \#3 \rangle \end{bmatrix}$$

Fig. 3. Cliticization rule for content verbs

Now we turn to the production of cliticisized *vi*. The basic form of object-control verbs is shown in Figure 4 and its cliticization rule in Figure 5. This rule removes the direct object from the complement list of the verb and includes it in its *CL-LIST* attribute; this argument is added on to the clitic list of its complement verb (e.g. *comiendo*), defining in this was a clitic composition. However, this clitic argument is co-indexed with the dative cliticisized argument of the second verb, and these two complements (of *vi* and *comer*) represent the same object and are subsumed into one.

$$\begin{bmatrix} ocv-lxm & \\ SS \mid LOC \mid CAT \mid VAL & \begin{bmatrix} SUBJ & \langle [\ ] \rangle \\ COMPS & \left\langle \#1 \& NP-acc_i, \begin{bmatrix} SUBJ & \#1_i \\ COMPS & \langle\ \rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Fig. 4. Lexical entry for *ocv-lxm*

The analysis of the final sentence is shown in Figure 6. The lexical entries for the verbs are produced by the lexical rules in Figure 3 and 5 out of the lexical entries in Figures 2 and 4 respectively; these combine to form the clitic composition *vi comiendo*, which in turn is combined with the clitic word "se_{DAT} la_{ACC}" through the Head-Proclitic rule that implements the syntactic cancellation scheme.

$$\begin{bmatrix} ocv-lxm & \\ ORTH & \#1 \\ HEAD & \#2 \\ ARG-ST & \langle \#3, []_i, \#4 \rangle \\ & \end{bmatrix} \mapsto \begin{bmatrix} ORTH & \#1 \\ HEAD & \#2 \\ VAL \mid CL-LIST & \langle NP-acc_i \rangle \oplus \# a \\ ARG-ST & \langle \#3, \#4 \,\&\, [CL-LIST \quad \# a] \rangle \end{bmatrix}$$

Fig. 5. Cliticization lexical rule for object-control verbs

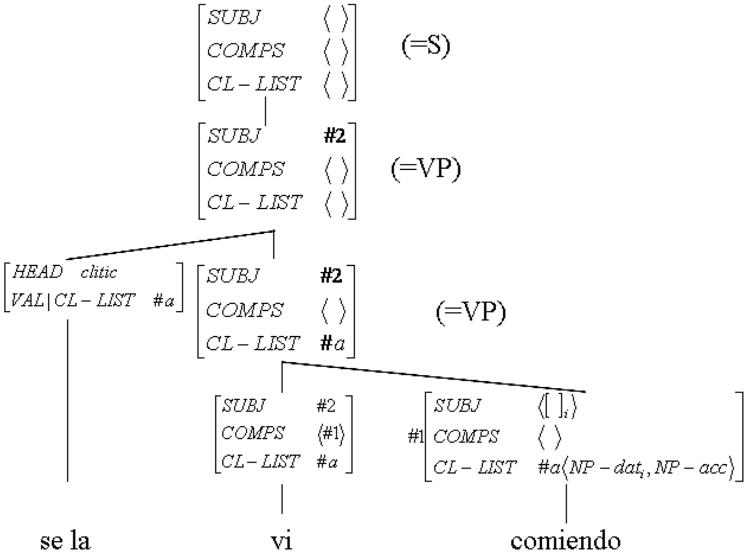

Fig. 6. Analysis of sentences with clitic subsumption

The reflexive connotation of the co-indexed pronouns can be better appreciated in (4i) *La*$_i$ *vi comiendose*$_i$*la*$_j$ where the cliticizations of both of the clitic hosts is not composed, and the direct object of *vi* appears as proclitic but the two complements of *comer* appear as enclitics; here, the two direct objects can be realized with the accusative *la* despite that they have different referents: the proclitic refers to the woman and the enclitic to the dinner; nevertheless, the proclitic is still co-indexed with the indirect object of *comer* represented by *se*, hence the reflexive interpretation. The analysis of this construction requires two cancellations: simple lexical cancellation by the right and syntactic cancellation by the left, but in both of these cases the clitic host is within the scope of its corresponding phonological host, and the two scopes do not overlap. We say that this kind of constructions has two independent clitic domains, and the sentence is analyzed in terms of one cancellation per independent clitic domain. More generally, the clitic host is within the scope of the phonological host if the former is within the clitic domain of the latter, and there is a binding path allowing the co-referring relation.

The composition and subsumption operations have an additional consequence: in coordinated structures, like *lo llevó y puso sobre la mesa*, the co-indexed cliticizations of both of the verbs are composed, and one argument is reduced by clitic subsumption too, resulting in an composite clitic argument which is factored out as a proclitic to the whole coordination, and the analysis requires the head pro-clitic rule, as illustrated in Figure 7.

On the basis of these observations we propose the following clitic principle: Spanish clitic sentences can be analyzed in terms of one cancellation operation per independent clitic domain, and the clitic composition and subsumption operations. Or more simply: a cliticization either basic or produced through composition or subsumption must be within the scope of its phonological host.

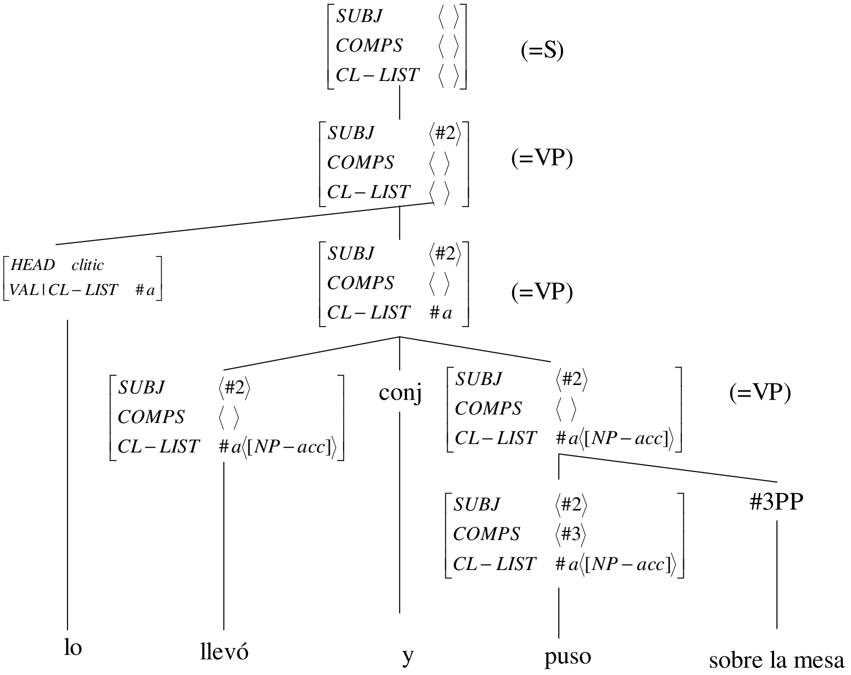

Fig. 7. Analysis of clitic coordinated sentence

4 Conclusions and Implementation

In this paper we have presented a theory for the analysis of Spanish clitic system with a dual character: proclitics that represent “normal complements” like direct and indirect objects of transitive verbs are independent lexical units and hence proper clitics, while enclitics are inflexions; other proclitics, representing extra complements (i.e. arguments extending the basic argument structure of the verb), whether these are proclitics or enclitics, are inflexions (e.g. *me voy*, *comerse*), and these attach to their hosts as lexical idiosyncrasies. In this theory the arguments of the cliticisized verb must be within the scope of the phonological host, and there is a single cancellation per independent clitic domain. Composite predicates with two content verbs can be formed by the clitic composition operation, and co-indexed arguments in compositions can be subsumed, producing composite predicates, as in complex periphrases and coordination. The theory postulates that the structure of clitic sentences involves an underlying phenomenon of argument reduction which is a natural way to account for sentences involving complex predicates, built out of independent verbs. In reflexive sentences the subject is co-indexed with the direct or indirect object, and the reflexive relation holds as long as the second co-indexed argument is within the scope of the first; however, if the argument appears twice, due to structure sharing between constituents of composite predicates, the extra argument needs not to appear explicitly and it is reduced. The fact that this phenomena appears in unrelated constructions like complex periphrases and coordination provides further support and motivation for our analysis and theoretical machinery.

The theory has been formally developed in HPSG (Pineda and Meza, 2004) and the results are backed by its implementation in LKB (Copestake, 2002).

References

1. Abeillé, A, Godard, D., Miller, P. and Sag, Ivan. 1998. 'French Bounded Dependencies' in Luca Dini and Sergio Balari (eds.), Romance in HPSG, Standford: CSLI Publications.
2. Copestake, Ann. 2002. The LKB System, Stanford University, http://www-csli.stanford.edu/\symbol/~aac/lkb.html
3. Gili Gaya, Samuel. 1991. Curso Superior de Sintaxis Española. Biblograf, S. A., Barcelona.
4. Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61, 95-120.
5. Miller P. H. and Sag, Ivan. 1995. 'French Clitic Movement Without Clitics or Movement', *Natural Language and Linguistic Theory* 15, 573–639.
6. Monachesi, Paula. 1999. 'A Lexical Approach to Italian Clitization', Lecture Notes series No. 84, CSLI, Stanford, Cambridge University Press.
7. Nevis, J. A., Joseph, B. D., Wanner, D. and Zwicky, A. M. 1994. 'Clitics, A Comprehensive Bibliography 1892-1991'. Library and Information Sources in Linguistics, 22. John Benjamins Pub. Co., Amsterdam/Philadelphia.
8. Pineda, Luis, Massé, Antonio, Meza, Ivan, Salas, Miguel, Schwarz, Erik, Uraga, Esmeralda and Villaseñor, Luis. 2002. 'The Dime project', Proceedings of MICAI-2002, Lectures Notes in Artificial Intelligence 2313, pp.166–175.
9. Pineda, L. & Meza, I. Un modelo para la perífrasis española y el sistema de pronombres clíticos en *HSPG*, *Estudios de Lingüística Aplicada*, Num. 38, pp. 45-67, 2003.
10. Pineda, L. & Meza, I. The Spanish pronominal clitic system, Internal report, Department of Computer Science, IIMAS, UNAM, México, 2004 (48 pp).
11. Pollard, Carl and Sag, Ivan. 1994. Head-Driven Phrase Structure Grammar, CSLI, Stanford. The University of Chicago Press, Chicago & London.
12. Sag, Ivan and Wasow, Thomas. 1999. Syntactic Theory: A Formal Introduction, CSLI Publications, Stanford.
13. Steedman, Mark. 1991. Structure and Intonation, *Language* 67, 260–298.
14. Villaseñor, L., Massé, A. & Pineda, L. A. 2001. 'The DIME Corpus', Memorias 3º. Encuentro Internacional de Ciencias de la Computación ENC01, Tomo II, C. Zozaya, M. Mejía, P. Noriega y A. Sánchez (eds.), SMCC, Aguascalientes, Ags. México, Septiembre, 2001.
15. Zwicky, A, & Pullum, G. 1983. Clitization vs. Inflection: English N'T', *Language* 59, 502–513.