A ROLE AND REFERENCE GRAMMAR ACCOUNT OF THIRD-PERSON CLITIC CLUSTERS IN SPANISH

by

Valeria A. Belloro

A project submitted to the
Faculty of the Graduate School of
The State University of New York at Buffalo
in partial fulfillment of the requirements for the
degree of

Master of Arts
Department of Linguistics

Buffalo, September 2004

ACKNOWLEDGMENTS

ABSTRACT

This paper analyses the morphosyntactic properties of Spanish clitic constructions, in terms of their ambiguous nature between pronominal arguments and agreement markers. It focuses on the Spanish spoken in Argentina, and especially on the characteristics of third-person clitic cluster constructions. The analysis is inscribed within the framework proposed by the model of Role and Reference Grammar (RRG). I suggest that RRG's PSA hierarchy is relevant for capturing both the likelihood that that a lexical PSA and the dative and accusative arguments will be syntactically expressed across different dialects, and the Argentinean Spanish preference for coding the plural of the dative argument into a position that otherwise encodes the number features of the accusative. With regards to the expression of the independent NPs, I argue that it is governed by the activation levels of discourse referents, whereas its position depends on the possible focus types of Spanish three-place verb constructions.

Table of Contents

List of Figures

1. INTRODUCTION

Spanish clitics look like the loose thread that forces the entire sweater unravel. They seem to interact with all the components of grammar, and the accounts that treat them in the privacy of one of them have to face sooner rather than later the series of counterexamples flagged by the supporters of another. Spanish dialectal varieties add some chaotic flavor to the mix, and often enough one founds that descriptive adequacy has been sacrificed in the way of accommodating to the principles of the grammatical model supported in each case.

In particular, there has been considerable controversy regarding the best way to capture the complex patterns of co-occurrence of clitics and their correspondent lexical counterparts (i.e. cases of "doubling"). Some researchers have considered clitics to be pronominal heads, whereas others have favored their analysis as (object) agreement markers. In turn, this issue has affected (or depended upon) the word-class to which they were ascribed, and whether their pronominal or affixal characteristics were highlighted.

This paper does not attempt to criticize all previous accounts nor to advance the definitive solution for the clitics phenomenon, but simply to review some of the most paradigmatic proposals that had been put forth (mainly within the generative tradition) and to signal the advantages of approaching the challenges posed by Spanish clitics from the perspective of a grammatical theory that explicitly supports the view that syntactic structures are influenced by semantic and pragmatic factors, as it is the case of Role and Reference Grammar (RRG, Van Valin and LaPolla, 1997; Van Valin, in press).

This proposal will focus specifically on the analysis of third-person clitic clusters in Argentinean Spanish, the particular "doubling" restrictions that they generate, and the idiosyncrasies they reflect in terms of agreement. I suggest that a consistent account of

these seemingly diverse characteristics is only possible if one allows the incorporation of the functional forces that drive the use of language into the grammatical description.

The organization of this paper is as follows. Section 2 reviews the morphosyntactic characteristics of clitic constructions, the arguments that have been advanced to consider clitics either as pronominal heads or as agreement markers, and the challenges raised by cases of "doubling". Section 3 revises the study of "doubling" constructions as a manifestation of the distribution of information in the sentence (i.e. in terms of topic and focus), and describes the possible patterns of information structure in Spanish. Section 4 addresses an idiosyncratic morphological feature of clitic clusters in some varieties of American Spanish: the marking of the plurality of the dative argument into the accusative clitic. It shows that this tendency, despite the scarcity of studies devoted to it, is far from marginal, constituting the preferred strategy in more than half of the cases for speakers of Argentinean Spanish. Section 5 presents the model of RRG, with particular attention to three-place verbs constructions. Section 6 proposes a representation of Spanish clitic cluster constructions in RRG, incorporating an Agreement Index node that assigns the agreement features coded by the clitics in accordance to a realizational approach to morphology. Likewise, the patterns of co-occurrence of clitic clusters and NPs are naturally explained in terms of RRG's formalization of focus types and activation levels of discourse referents, and examples of semantics to syntax linking are made explicit to illustrate these issues. In Section 7 I summarize the phenomena discussed in the paper and some of the consequences that derive from them.

2. THE MORPHOSYNTACTIC PROPERTIES OF CLITIC CONSTRUCTIONS

Most of the work to date on the properties of Romance clitics has been developed within the generative tradition, and can be grouped along two basic proposals: the "movement hypothesis" (Kayne 1975, among others) and the "base-generation hypothesis" (Jaeggli 1981, among others).

In the analyses that follow Kayne's seminal work, clitics are considered pronominal heads generated in canonical argument position. It is assumed that due to their weak phonological nature the clitics "move" in order to attach to the host (i.e. the verb), leaving a "trace" in the argument's structural position. Since the argument's structural position is occupied by a trace, this analysis further predicts that the occurrence of a lexical phrase in the same slot would yield ungrammatical results. This prediction is born out in the analysis of French data, in which clitics and lexical phrases occur in strict complementary distribution. However, it was promptly noted that such strict complementarity does not hold for other Romance languages (for instance, Spanish) in which the co-occurrence of both pronominal clitics and their correspondent lexical phrases occurs quite extensively.

In an attempt to resolve this problem, it was proposed that the lexical phrases do not occupy an argument position but are "dislocated" elements, thus avoiding the violation that the "movement hypothesis" would impose on Theta-Criterion in the cases of clitic "doubling". However, this proposal was also proved unviable since (i) clitics can serve as antecedents of anaphors (which must be bound from argument positions); (ii) there does not need to be a pause between this supposedly "dislocated" element and the rest of the clause; (iii) it would imply treating identical objects occupying similar linear positions sometimes as arguments and other as adjuncts, depending on whether they

are "doubled" by a clitic or not (cf. Franco 2000 and the references herein for a more thorough revision of these hypothesis).

As an alternative to the "movement hypothesis" it was proposed that the clitics are directly generated on their surface position, thus leaving the canonical argument position free to be occupied, in the cases of "doubling", by the corresponding lexical phrase. This approach, known as the "base-generation hypothesis" was originally developed by Rivas (1977) and Jaeggli (1981). However, it also faces some problems. If the constituent formed by the clitic and the verb is assumed to form a syntactic unit, the Lexical Integrity Hypothesis (LIH, as proposed by Chomsky 1970) needs to be abandoned, since clitics enter into syntactic operations (such as "climbing" and functioning as binders). On the other hand, if the LIH is preserved, one cannot account for the clitics' affixal nature. This latter alternative is nevertheless the one that Jaeggli explicitly favors, presenting three reasons why clitics should be considered as words, separated from the verbs to which they attach: clitics can "climb" in the syntactic structure, they do not affect the stress pattern of the verb, and they show nominal inflectional morphology. Still, these reasons do not always hold. Whereas it is apparent that the phenomenon of clitic "climbing" does not have any correlate among affixes[1], it is less clear that clitics never affect the stress pattern of the verb to which they attach. On the contrary, some enclitic cluster constructions cause the stress to shift to the ultimate syllable at least in some varieties of Spanish (cf. Argentinean "poné" vs. "poneteló" 'put' vs. 'put it on you'[2]). Regarding the fact that clitics retain some inflectional similarities with

---

[1] At least in Spanish. However Franco (2000:182) refers to Laka (1993) as presenting "evidence from languages with full-fledged verbal agreement [in which] agreement markers can be displaced within the inflectional amalgam."

[2] Fernandez Soriano (1993) mentions some other phonological process generated by the presence of the clitic: for instance, Standard Peninsular Spanish eliminates in the imperative the second person plural –d and the first person plural –s from the verb: poned – poneos; vamos,

4

other words, such as the –*s* marking of plurality and the contrast –*a/-o* to mark feminine and masculine gender, one must note that number inflection in –*s* only applies to third person clitics (*le/les*, *lo/los*) and gender inflection only to its subset marking accusative case (*lo(s)/la(s)*). The rest of the paradigm presents neither number nor gender inflection. Jaeggli's final argument is that whereas the agreement marker is obligatory, clitic pronouns are optional. We will see below that, on the one hand, certain constructions seem to require the presence of the clitics and, on the other hand, the cases analyzed as optional can be explained if one considers clitics, as other agreement phenomena, influenced by semantic and pragmatic factors such as the status of each element along an Animacy Hierarchy[3] and their role in phenomena of topic continuity in discourse structure.

Furthermore, there seems to be more straightforward evidence supporting the view of clitics as affixes also from a purely grammatical perspective. Following the tests proposed by Zwicky and Pullum (1983), Monachesi (ms.) examines the characteristics that relate clitics to affixes in several Romance languages. All of these characteristics are applicable to Spanish:

1) *Degree of selection with respect to the host*: Similar to affixes, Romance clitics (save Romanian) are constrained in terms of their combinatory possibilities (they can only combine with a verb). As inflectional affixes, they do not alter the lexical category of the host.

---

vámonos. Likewise, she acknowledges the tendency, in colloquial speech, to attach the plural subject agreement morpheme to the cluster formed by the verb and the clitic: denle → delen.

[3] Advancing arguments in favor of the "agreement hypothesis", Franco (2000:169) notes that "Comrie (1989), among many others, shows that many agreement relations are driven by an Animacy Hierarchy which has an overall cross-linguistic validity. Thus, arguments that occupy a high position in the Animacy Hierarchy hold stronger or more uniform agreement relations than those occupying a lower position. In broad terms, first, second and third person pronouns in this order occupy the highest positions in this hierarchy followed by definite human nouns, definite nouns, etc, whereas inanimate generic nouns occupy the lowest ones" .

2) *Rigid order*: Like in the case of affixes, clitic clusters are arranged into an idiosyncratic rigid order (cf. Perlmutter 1971).

3) *Coordination*: Clitics cannot have wide scope over coordinated verbs. The only exception is with verbs closely related semantically (like "wash and dry") and only if the clitics precede the verbs: "Lo lavamos y secamos en cinco minutos" ('We wash and dry it in five minutes'); *"Para lavar y secarlo en cinco minutos" (*'To wash and dry it in five minutes').

4) *Arbitrary gaps*: As in inflectional paradigms, there are arbitrary gaps affecting the clitic's combinatory possibilities. For instance, in Spanish there cannot be a sequence formed by a dative third person clitic preceded by a first or second person accusative one (*"Me le entregaron", *'They handed me to him').

5) *Morphophonological idiosyncrasies*: As it is the case with inflectional formations, the phonological shape of clitics may be affected by other clitics with which they combine. One Spanish example of this phenomenon is the case of the "spurious *se*", which is an allomorph of the dative "le(s)" when combined with an accusative clitic.

6) *Verb left-detachment*: As has been argued for Italian (Beninca 1988) and Catalan (Vallduvi 2001), Spanish bare infinitives and cliticized infinitives can be left detached: "Dormir, duermo barbaro" (To sleep, I sleep fantastically); "Saludarla, la saludo todos los dias" (To greet her, I greet her every day). The same construction is ungrammatical if the bare infinitive is followed by a complement (*"Saludar a María, la saludo todos los dias" (*To greet María, I greet her every day). Monachesi argues for Italian, and the same is applicable to Spanish, that the fact that bare and cliticized infinitives have the same status offers another argument in favor of the affixal view of clitics in these languages[4].

---

[4] There are still reasons to distinguish clitics from regular affixes: To the syntactic processes mentioned by Jaeggli, one must add the fact that there are the morpho-syntactic properties of the

Further, another problem faced by the "base-generation" hypothesis is that inasmuch as it considers clitics to be arguments of the verb, it needs to create some "absorption" rules (either in terms of thematic roles or in terms of case) that account for the role of the NP in the cases of "doubling". The idea of analyzing the clitics as absorbing a thematic role, proposed by the advocates of the "dislocation" version of the "movement hypothesis", was argued against above. Franco (2000) also argues against the option of considering that clitics absorb case assigned to the lexical phrase, in turn, by the preposition *a* (Jaeggli 1981), offering examples in which the preposition does not necessarily occur (idem:156):

(1)   La   comí   la   torta
        ACC.CL ate.1s.   the cake
        'I ate the cake'

The proposal put forth by Franco shares with the "base-generation hypothesis" the assumption that clitics are generated *in situ*. However, he favors the idea of treating them as object-verb agreement morphemes, on a par with subject-verb agreement. As Franco himself notes, one of the challenges faced by this "agreement hypothesis" is to explain why there are apparent restrictions for direct object agreement that do not hold in the cases of agreement with the subject or indirect object.

Let us remember that the presence of the dative clitic in what are traditionally called "ditransitive" constructions is allowed across all varieties of Spanish. Moreover, it has been argued that there are certain constructions in which the dative clitic is obligatory. Fernandez Soriano (1993) mentions the cases of inalienable possession constructions (2), and predicates introducing a "benefactive" (3) or an "experiencer" (4), as demanding the occurrence of dative clitic:

---

verb what determine whether the clitic will precede (finite verbs) or follow the head (infinitives and imperatives). This alternation is not typical of affixes which, moreover, across Romance languages usually occur after the head (Monachesi, p.44).

(2)   (*)le   duele   la   cabeza   a Juan
        DAT.CL   hurt.3s. the   head   to Juan
        'Juan has a headache'

(3)   (*)le   preparé   una   tarta a mi amigo
        DAT.CL   prepared.1s. a   cake   to my friend
        'I prepared my friend a cake'

(4)   (*)le   gusta   el   cine   a Juan
        DAT.CL   like.3s. the   movies   to Juan
        'Juan likes the movies'

With respect to accusative clitics, its presence is likewise obligatory if there is a strong pronoun in direct object position (5). Otherwise, certain restrictions apply. For instance, it has been argued (Silva-Corvalán 1984, Suñer 1988) that the accusative clitic is optional if the direct object is animate (6) or, at least, specific (7). These last two alternatives, in turn, are restricted to some dialectal varieties; in particular, to the Spanish of the Rio de la Plata region. Even this dialect, however, seems to avoid the "doubling" of (unspecific) indefinite objects (8):

(5)   (*)Lo vi a él   I saw him

(6)   (Lo) vi a Juan   I saw Juan

(7)   (Lo) compré el regalo   I bought it the present

(8)   (*Lo) compré un regalo *I bought it a present

Regarding the animate-specific restrictions, Franco claims that the ungrammaticality of the non-specific examples can be reinterpreted as a side effect of the fact that non-specificity tends to correlate with a low position in the Animacy Hierarchy (see footnote 2), which he posits to be a more relevant criterion. In support of this assumption he presents the following examples (idem:176):

(9) a. Juan lo   invitaba   a uno y luego se   olvidaba.
       Juan ACC.CL. used to invite to one and then DAT.CL. forgot
       'Juan used to invite people and then forget all about it'

8

b. En ese departamento lo   admiten   a   cualquiera.
   In   that department   ACC.CL admit.3pl to   anyone
   'In that department they admit anyone'

Franco argues that both "uno" and "cualquiera" allow clitic "doubling" because, despite their non-specific status, both rank high in the Animacy Hierarchy. However, Franco's account is dependent on Chomskyan's generative models, in which the Animacy Hierarchy has no formal status. In order to solve this problem, he needs to correlate this feature with particular syntactic positions, mapped onto structural configurations. Roughly, for the purpose of explaining examples of accusative doubling his hypothesis depends on assuming an AgrDO node and positing that the lexical phrase raises to its Spec position in other to check the feature [presuppositional]. On this account, doubled accusatives must always be presuppositional, either unspecific (as in the examples (9) a. and b. above), or specific, as in cases like the following (adapted from idem:183), which is only acceptable if referring to one particular student already mentioned or otherwise recoverable:

(10)   Lo   he   visto   a   un estudiante.
        ACC.CL have   seen to   a   student
        'I have seen a student'

Thus, according to Franco, object-verb agreement is a strategy to scope objects out of the VP in order to guarantee an unambiguous presuppositional reading of the direct object. Further, since this is claimed to be an "overt" movement, it is more costly than the assumed "covert" movement of the dative NP to Spec of AgrIO in dative "doubling" constructions. Within this perspective, this distinction further explains the fact that the latter is more productive.

Of course, when faced with the term "doubling", one must ask which one the "doubling" element is, whether the clitic or the correferential NP. As is apparent,

proposals that consider the clitic as a pronominal argument will tend to assume that the "doubling" element is the lexical NP, whereas supporters of the interpretation of clitics as agreement morphemes will tend to posit that the "doubling" element is the clitic. Under the latter view, it is claimed that the argument position is structurally realized by either the NP or the phonologically null *pro*. This perspective, granted the relevant conditions summarized above, would allow one to account for the cases in which only the clitic appears in surface structure (agreeing with *pro*), for the cases in which the only "overt" manifestation are the lexical arguments (i.e. when there is no "object agreement" on the verb), and for the cases in which either of the arguments is "doubled" by a clitic.

Further, by ordering the agreement nodes in the relevant order (with AgrIO dominating AgrDO) this account may also accommodate another idiosyncratic feature of clitic constructions with ditransitive predicates: Whereas the dative clitic may or may not co-occur with the accusative clitic (11, 12); the accusative clitic, on the contrary, *requires* a dative clitic to occur as well (13):

(11)   Juan le   dio   un ramo de flores   a María.
        Juan DAT.CL gave.3s. a   bunch of   flowers to María.
        'Juan gave María a bunch of flowers'

(12)   Juan se   lo   dio   (a María).
        Juan DAT.CL ACC.CL gave.3s (to María)
        'Juan gave it to her'

(13)   *Juan lo   dio   a María.
         Juan ACC.CL gave.3s. to María
        'Juan gave it to María'

One may assume that the ungrammaticality of (13) derives from the relative status of the verb complements along the Animacy Hierarchy, since it represents a case in which a lower-ranking argument (the bunch of flowers) would be agreed with over a higher-ranking one (María).

However, ditransitive constructions generate other difficulties which seem to require an explanation based on pragmatic factors, such as the distribution of information in the sentence. Consider that, under an account such as Franco's, it isn't clear what would be the restrictions with respect to whether one *or both* of the complement positions are occupied by *pro* or by the correspondent lexical forms. Sentences in which both arguments are "doubled" are nontheless, at least under a prosodically unmarked reading, ungrammatical:

(14) *Se   los   dejé   los chicos   a los abuelos
        DAT.CL ACC.CL left.1s. the children to the grandparents
      'I left the grandparents the children'

The sentence becomes acceptable, however, when one of the arguments, or both, occurs in pre-verbal position:

(15)   Los chicos se los dejé a los abuelos.

(16)   A los abuelos se los dejé, los chicos.

(17)   Los chicos a los abuelos, se los dejé.

(18)   A los abuelos, los chicos se los dejé.

These examples seem to address not the principles that govern the occurrence of the clitics, but of their correspondent lexical forms, something that is not fully explained under the previous analyses and that seems to call for an account in terms of discourse structure.

3. CLITIC "DOUBLING" AND DISCOURSE STRUCTURE

Approaches to clitic doubling phenomena in terms of the distribution of information in discourse aren't new. Wiessenrieder (1995) mentions Silva-Corvalán's (1981) pioneering study of accusative doubling in Argentinean Spanish, which showed not only that animate/definite referents are more likely to be doubled, but also they are more likely to be topics of discourse. Likewise, Wiessenrieder correlates dative verb agreement with the reference to entities that are highly identifiable and topical in the discourse segment, using as corpus the novel "El beso de la mujer araña", by the Argentinean author Manuel Puig.

In a recent article, Colantoni (2002) confirms the relevance of the animacy scale for clitic doubling, based on quantitative data obtained from a corpus of oral interviews with Spanish speakers from Corrientes, Argentina. She finds that the "doubling" cases serve either to render the referent of the pronoun more identifiable (as in the cases of elements newly introduced in the discourse context), or to assign contrastive focus to one constituent when there is a set of candidates competing for the reference (p. 326-327). As is apparent, her findings surprisingly seem to correlate the occurrence of doubling with focal, instead of topical, elements.

But again, this disagreement with regards to whether "doubling" is preferred in topical vs. focal contexts seems to arise from the interpretation of which the "doubling" element is. Wiessenrieder treats the dative object agreement as "doubling" the lexical phrase, whereas Colantoni focuses on the cases in which the lexical phrase "doubles" the clitic[5]. At bottom, they aren't saying contradictory things: one claims that the clitic

---

[5] A similar proposal is put forth by García-Miguel (1991): "Between *lo vi* and *lo vi a él* [ACC.CL saw.1.s.; ACC.CL saw.1.s. him] there is only a difference in terms of information structure, which consists in the fact that *a él* is the contrastive focus, whereas *lo* cannot be it. This difference is the same that exists in the subject between *lo vio* and *lo vio él* [ACC.CL saw.3.s.; ACC.CL saw.3.s. he]. The 'objective conjugation' does not differ at all in this regard from the 'subjective conjugation' (p. 400, my translation).

marks topical elements, and the other that the lexical phrase serves to mark focus. Both interpretations are in fact consistent with Topic Continuity Hierarchies such as the one proposed by Givón (1983) and Levinson (1987), among others. These hierarchies represent the "markedness" of occurrence as Topic along the following (increasing) continuum (from Van Valin, in press): Zero < Clitic/Bound pronoun < Unstressed pronoun < Stressed pronoun < Definite NP < Indefinite NP. Therefore, it should be natural to find a correlation between topical elements and clitics, as one between lexical expressions and focal elements. However, since both Colantoni and Wiessenrieder's studies are analyzing "doubling" constructions, their descriptions end up being incompatible, unless one would want to assume that the same referent could be coded as topical *and* focal at the same time.

A related problem is that, since neither the lexical complements nor the clitics form a homogeneous class with regard to its privileges of occurrence (cf. the restrictions on the presence of the accusative clitic discussed above), it isn't at all clear that it is possible to lump together, in a construction-independent basis, the functions that are served by the dative and accusative clitics on the one hand vs. the functions served by direct and indirect lexical complements on the other. Those differences seem particularly relevant in cases in which *both* clitics appear in the construction (something that, to my knowledge, has received little attention). If anything, one would predict that in these cases the most allowed lexical phrase would be the accusative one, since it is the one more likely to be FOCAL (let us remember that it is consistently assumed that at the level of discourse organization there is a correlation between dative NPs, animate referents, and discourse TOPICS). However the evidence points in the opposite direction, with sentences like (19) being more acceptable than the ones like (20):

(19) Se   lo   compré   a María.
      DAT.CL ACC.CL bought.1s. to María.
      'I bought it (for) María'

(20) ? Se   lo   compré   el regalo.
       DAT.CL ACC.CL bought.1s. the present
      'I bought her the present'[6]

Further, one may need to account for the acceptability differences depending on whether (and which one of) the complements are fronted (cf. (15) to (18) above). In order to attempt this, we may need a little excursus that allow us to review what has been said about Spanish grammatical means to express the distribution of information in the sentence.

3.1. INFORMATION STRUCTURE IN SPANISH

Along the typological-based analysis proposed in Van Valin (1999) within the model of Role and Reference Grammar (RRG, Van Valin and La Polla 1997), Spanish can be categorized as a token of those languages that have both flexible word-order and a rather flexible focus structure. Following Lambrecht (1994), RRG assumes three possible "focus types", which correspond to cross-linguistic recurring patterns in the distribution of information in the sentence. These types are "predicate focus", "sentence focus" and "narrow focus". Examples of each type in Spanish are given below:

---

[6] To the objection that these examples show complements that differ strikingly in terms of "animacy", if we abide by the pragmatic awkwardness, we can see that the result is the same in cases in which this difference is neutralized. Imagine (a) as the discourse context:

(a)   Le di a Juan el hijo que quería (I gave Juan the son he wanted)
(b)   Se lo di a Juan.
(c)   ? Se lo di el hijo.

(21)
Q: What happened to your car?
A: Se me descompuso.   Predicate focus
   'It broke down'.

(22)
Q: What happened?
A: (i) Se me descompuso el auto.   Sentence focus
   (ii) El auto se me descompuso.
       'The car broke down'.

(23)
Q: I heard that your motorcycle broke down.   Narrow Focus
A: (i) EL AUTO se me descompuso.
   (ii) Se me descompuso EL AUTO.
        'THE CAR broke down'.
   (iii) Es EL AUTO el que se me descompuso.
        'It is THE CAR that broke down'.

In Spanish, predicate focus normally correlates with a predicate phrase whose last constituent receives neutral focal accent. Sentence focus is realized by constructions in which the subject NP is explicit and typically occurs post-verbally[7]. Narrow focus can target any constituent in the sentence, which will receive emphatic accent and will be interpreted contrastively. In (23), the narrow focus concerns the subject NP. Spanish can mark narrow focus on the subject by assigning it emphatic accent in its canonical position (23.i), by means of an inverted construction (23.ii), or by a cleft (23.iii).

The basic SVO word-order coincides with an unmarked distribution of information in the sentence in which the topic (normally the subject) precedes the focus (typically the verb and its complements). As we just saw, however, the correlation is not systematic. Zubizarreta (1999) recognizes two possible kind of topicalized constructions in Spanish ("hanging topic" and "left-dislocation"), and one kind of focalized structure ("pre-posed focus construction"). Let us remember that in Spanish the unmarked focus position

---

[7] Differently from Italian, for instance, Spanish can allow the default SVO order in Sentence focus constructions (Zubizarreta 1999:4225), as shown in the example (22.ii).

coincides with the last syllable of the melodic group of the clause. Consider the following example:

(24)   Compré el periódico.
         '(I) bought the newspaper'.

In (24) there is a neutral focal accent falling within the domain of the last constituent (i.e. "el periódico"), and therefore this sentence can be uttered as an answer to the question "What did you buy?" "You bought something" is considered as pragmatically presupposed information (Lambrecht 1994), and "el periódico" as part of the pragmatic assertion (i.e. as focal)[8]. On the other hand, it is possible to have an emphatic accent assigned to any stressed morpheme. When this accent falls into a morpheme different than the one selected by the neutral accent and assigns a (different) value to a variable introduced in the presupposed proposition, the interpretation is contrastive, as in the narrow focus constructions illustrated above. Another example is given in (25).

(25)   PEDRO compró el periódico (no Juan).
         'PEDRO bought the newspaper (not Juan)'.

Here the emphatic accent (in capital letters) falls on "Pedro", and therefore this constituent is interpreted as carrying contrastive focus. In this case the pragmatic presupposition is that "Someone-other-than-Pedro bought the newspaper". We can refer to it as the "topical"[9] segment. The "pre-posed focus" derives its contrastive

---

[8] Lambrecht (1994:52) defines "pragmatic assertion" as "the proposition expressed by a sentence which the hearer is expected to know or believe or take for granted as a result of hearing the sentence uttered". Accordingly, the "focus" of an utterance is "the part that is asserted in a declarative utterance or questioned in an interrogative utterance" (Van Valin, in press).

[9] I will use, for the sake of simplicity, the term "topical" to refer to what in Lambrecht's terms corresponds to the "pragmatic presupposition" ("The set of propositions lexicogrammatically evoked in an utterance which the speaker assumes the hearer already knows or believes or is

interpretation from the necessarily emphatic nature of the focal accent in a position other than the one assigned by default. When a pre-posed focus targets an object, it triggers a subject inversion resulting in the order OVS. Furthermore, in these cases the presence of a correferential accusative object clitic is ungrammatical. The relevant examples are given below.

(26)   EL PERIODICO (*lo) compró Pedro

(27)   *EL PERIODICO Pedro compró.

As it is possible to have focal subjects (as illustrated in (25) above), it is also possible to have topical elements other than the subject, either by means of a "hanging topic" construction (28) or by "left-dislocation" (29):

(28)   Con respecto al periódico, Pedro sólo compra Clarín.
        About the newspaper, Pedro only buys Clarín.

(29)   El periódico lo compró Pedro.
        The newspaper, Pedro bought it.

There are certain asymmetries that differentiate these constructions. Both the "hanging topic" ("Con respecto al periódico") and the "left-dislocated" element ("El periódico") are claimed to occur in the "left periphery" of the sentence, a position reserved for topical elements. However, the "hanging topic" constructions serves to introduce a change in the discourse topic, and it can be preceded by phrases such as "en cuanto a", or "con respecto a" ('about X', 'with respect to X'). Its correlate in the clause can be any constituent, and it does not need to be any kind of grammatical dependency between them. "Left-dislocation", on the other hand, is more restricted.

---

ready to take for granted at the time of speech" 1994:52). On the other hand, I will use the term "topicalized" when the assignment of "topical" status to a given constituent depends on a grammatically marked construction.

There is a grammatical dependency between the "left-dislocated" element and its counterpart in the clause, which furthermore cannot be a position within a relative clause, an adverbial clause, or the subject. When it relates with an object, the presence of a correferential clitic is obligatory. As (29) above shows, the boundary of the "left-dislocated" constituent does not need to be marked by an intonation break, nor does it require being adjacent to the verb.

With regards to focal elements, given the default patterns of focal accent mentioned before, it will be the object that will be in focus in a basic SVO linearization. However, Spanish also admits VOS order, as in (30)

(30)   Compró el periódico Pedro

In this case, the subject of the sentence ("Pedro") gets focal interpretation, since its position coincides with the one that receives neutral focal accent. However, it is also possible to have the subject in a "right-dislocated" position, as in (31)

(31)   Compró el periódico, Pedro.

As it is represented by the comma, in this case there is a pause that separates this constituent from the rest of the clause. Since the subject appears in a peripheral position, it is outside the scope of the neutral focal accent, which falls in the object (i.e. "el periódico")[10].

I mentioned before that one of the differences between a "left-dislocation" construction and a "pre-posed focus" construction is that the former requires (and the latter prevents) the occurrence of a correferential clitic. Another difference is that whereas in a "left-dislocation" construction it is possible to have more than one

---

[10] Zubizarreta notes, however, that in these cases it is possible to have two interpretations: one with broad focus on both the object and the subject, and one with narrow focus on the object, in which the subject is "topicalized".

constituent preceding the verb (33), in the case of a "pre-posed focus" construction this yields ungrammatical results, since the focalized segment needs to be adjacent to the verb (32). Zubizarreta provides the following examples:

(32)   *Estoy segura de que la MANZANA, a EVA le dio Adán (y la PERA a MARÍA)
        'I am sure that the APPLE, Adam gave to EVE (and the PEAR to MARÍA)

(33)   Estoy segura de que la manzana, a Eva se la dio ADAN
        'I am sure that the apple, ADAN gave to Eve'

With these distinctions in mind, we can now try to account for the grammaticality differences in "cliticized" three-place verbs constructions. The relevant distinctions seem to be in terms of their Potential Focus Domains (PFD, Van Valin and La Polla 1997, Van Valin, in press). In constructions involving a three-place verb triggering "object agreement", the PFD excludes the DO, which cannot be focal:

(34)   *Se lo compré el regalo.

In this sentence the neutral focal accent falls in "el regalo", and the construction is ungrammatical. However, it is possible to have the object in a "right-dislocated" position (along the lines of the constructions illustrated in (30) and (31) above). In this case there will be a pause marking the boundary:

(35)   Se lo compré, el regalo.

This restriction does not apply to dative objects, which are within the ADF in these constructions and, therefore, can be focal:

(36)   Se lo compré a María.

We need to note that one cannot claim that accusative arguments are necessarily topical across the board, since there are examples of accusative "clitic

doubling" in monotransitive clauses in which the neutral focal accent falls in the accusative NP (37.a). It is even possible to have contrastive narrow focus in the accusative argument (37.b):

(37)   'Whom did you see?'
    a. Lo vi a Juan.
       'I saw Juan'

       'I heard you saw Martín'.
    b. No, lo vi A JUAN.
       'No, I saw JUAN'.

However, I will claim that what defines these cases is the precise level of activation of the accusative argument in its representation in RRG's Logical Structure (cf. below p.52-53), which is consistent with the fact that these cases of "doubling" typically involve proper nouns, strong pronouns, or otherwise specific referents.

We also noted that the acceptability judgments varied if one (or both) of the arguments were fronted. Now we can explain these cases in terms of the relevant focus structures in Spanish. When the accusative argument is fronted, it occurs in a "left-dislocated" position outside the clause (but within the sentence) and is interpreted as topical:

(38)   Los chicos, se los dejé a los abuelos

When the fronted argument is the dative, it occurs in a "pre-posed focus construction", and it is interpreted contrastively:

(39)   A LOS ABUELOS se los dejé los chicos

We will see below (sections 5 and 6) that we can account naturally for the position and function of these "dislocated" and "pre-posed" elements in terms of the

Layered Structure of the Clause (LSC) and the information coded in the Logical Structure proposed by RRG.

In the preceding two sections we reviewed both the syntactic and the pragmatic approaches to clitic doubling, paying special attention to three-place verbs constructions, and in particular to the complex interaction of clitics and NPs that they exemplify. In the following section, I will present another "idiosyncratic" characteristic of third person clitic clusters in Spanish, before we attempt to give a consistent account of all these phenomena in terms of Role and Reference Grammar.

4. THE TRANSLATION OF PLURALITY IN SPANISH CLITIC CLUSTERS

As illustrated in many examples above, accusative clitics agree in number and gender with their referents, resulting in a set of four possible forms: 'lo, la, los, las', for masculine, feminine and their plurals, respectively. On the other hand, dative clitics neutralize gender agreement, coding only number: 'le, les'. Moreover, it is traditionally assumed that a process of dissimilation of the sequence of lateral sonorants that originated in Old Spanish gave rise to the invariable form 'se' when the DAT is followed by an ACC clitic ('le lo', 'les lo' became 'se lo', etc.) Consider the following examples (adapted from Company, 1998):

(40)   a.   Juan compró UN DEPARTAMENTO$_i$ PARA SU HIJO$_j$ >
                 > Juan LE$_j$ compró un departamento$_i$.
                 > *Juan LE$_j$ LO$_i$ compró > **Juan SE$_j$ LO$_i$ compró**.
                 'Juan bought an apartment for his son.'

        b.   Juan compró UN DEPARTAMENTO$_i$ PARA SUS HIJOS$_j$ >
                 > Juan LES$_j$ compró un departamento$_i$
                 > *Juan LES$_j$ LO$_i$ compró > **Juan SE$_j$ LO$_i$ compró**.
                 'Juan bought an apartment for his sons.'

c.   Juan compró UNA CASA$_i$ PARA SU HIJO$_j$ >
         > Juan LE$_j$ compró una casa$_i$
         > *Juan LE$_j$ LA$_i$ compró > **Juan SE$_j$ LA$_i$ compró**.
         'Juan bought a house for his son.'

d.   Juan compró UNA CASA$_i$ PARA SUS HIJOS$_j$ >
         > Juan LES$_j$ compró una casa$_i$
         > *Juan LES$_j$ LA$_i$ compró > **Juan SE$_j$ LA$_i$ compró**.
         'Juan bought a house for his sons.'

Whereas this is the "normative" cliticization, speakers of many dialects of American Spanish tend to add a plural morpheme into the ACC pronoun when faced with a referentially plural dative, thus yielding:

b(i).   Juan compró UN DEPARTAMENTO$_i$ PARA SUS HIJOS$_j$ >
         > **Juan SE$_j$ LOS$_i$ compró.**
         'Juan bought an apartment for his sons.'

d(i).   Juan compró UNA CASA$_i$ PARA SUS HIJOS$_j$ >
         > **Juan SE$_j$ LAS$_i$ compró.**[11]
         'Juan bought a house for his sons.'

The most widespread interpretation of the phenomenon is to consider it an "error" ascribed to "popular speech". The Grammar of the Spanish Royal Academy mentions one example as an instance of "dialectal variation"; Alonso and Henríquez Ureña (1951) include a reference to it in the section "Error Correction" of their grammar; Kany (1945:141) labels it a "syntactic error"; and Flórez (1977:141) states that it is "apenas pasable en el habla familiar" (barely acceptable in informal style).

In a sense, the very profusion of grammarians recognizing this "marginal error" suggests that the phenomenon may be more widespread they may want to recognize. This is in fact what the extensive corpora gathered during the 70's throughout major Latin American cities seem to suggest. These corpora of spontaneous oral speech were

---

[11] Actually, I will argue below that the cliticized form needs to correlate with a dative argument marked by "a", not "para", a distinction we are here disregarding for the sake of the clarity of the relevant context for the examples.

collected following precise pre-established norms that render them quite homogeneous, and were aimed to reflect the production of *cultured* speakers[12]. The following table representing the occurrence of the innovative vs. orthodox cliticization in different dialects was taken from De Mello (1992):

"SE LOS" VS. "SE LO" WITH SINGULAR DIRECT OBJECT REFERENT					
CITY	SE LOS			SE LO	
BOGOTA	6	(75%)		2	(25%)
BUENOS AIRES	10	(67%)		5	(33%)
CARACAS	6	(25%)		18	(75%)
HAVANA	4	(57%)		3	(43%)
LA PAZ	0	(0%)		0	(0%)
LIMA	0	(0%)		2	(100%)
MADRID	0	(0%)		6	(100%)
MEXICO CITY	13	(76%)		4	(24%)
SAN JUAN	0	(0%)		6	(100%)
SANTIAGO	9	(53%)		8	(47%)
SEVILLA	0	(0%)		2	(100%)
TOTAL	48	(46%)		56	(54%)

Figure 1: Global incidence of innovative cliticization.

The table in Figure 1 shows that, despite the total numbers of fully pronominalized three-place predicates is small, the innovative cliticization represents the most common use in several dialects of American Spanish. These percentages, therefore, shed doubts on its interpretation as a marginal error and its adscription to "uncultured" speakers. Note also that there is a skew with regards to the American vs. Peninsular varieties, in which it has been unattested. Among American cities, the innovative marking seems particularly prominent in the dialects of Mexico City, Bogotá, and Buenos Aires. The scarce studies that mention this phenomenon assume that it

---

[12] The corpora to which I am referring are the result of the "Proyecto de estudio combinado de la norma lingüística culta de las principales ciudades de Iberoamérica y de la Península Ibérica", whose objective was to collect a corpus of spoken Spanish to serve as the basis for language study (cf. De Mello, 1992).

occurs as a way of solving the ambiguity that the lack of number inflection in the suppletive form of the dative clitic generates. Consider the following excerpts:

> "En cuanto a las combinaciones *se los*, *se las*, por *se lo*, *se la*, en frases como *se los advierto*, para evitar la ambigüedad que encierra la forma singular, el uso chileno sigue la corriente general del habla hispanoamericana: 'Los niños pidieron pan y no había quién se los partiese.'" (Oroz, Rodolfo. 1966. La lengua castellana en Chile. Santiago: Universidad de Chile. p. 377).

> Regarding the combination *se los*, *se las*, instead of *se lo*, *se la* in phrases such as *se los advierto* [I warn you all], *in order to avoid the ambiguity of the singular form*, the Chilean use follows the general trend of Hispano-American speech: 'Los niños pidieron pan y no había quién se los partiese' [The boys asked for bread and there was no one who would cut it-plural for them] (In this and the following quotations, translation and emphasis mine).

> Esta neutralización motiva los usos anómalos a los que se refiere la presente nota; en ellos, el hablante, presionado por la necesidad de indicar la pluralidad del tercer actante, transfiere a la forma pronominal del segundo, que es paradigmáticamente pluralizable, el gramema de número que correspondería al tercer actante y que no puede ser indicado allí por no ser la forma que lo designa pasible de una pluralización (Rivarola, 1985).

> *This* [number] *neutralization* [of the dative clitic] *motivates* the anomalous uses which are the topic of the present article. In them the speaker, *pressured by the necessity of indicating the plurality of the third participant*, *transfers the number morpheme* that would correspond to the third participant -and which cannot be indicated there because the form is not able to be pluralized-, to the pronominal form of the second, which is paradigmatically pluralizable.

However, there are no reasons offered as to why the speakers are in fact "pressured by the necessity of indicating the plurality of the third participant". One must also note that the "ambiguity solving" hypothesis, as presented, is flawed in that these constructions *do not* resolve the potential for ambiguity in the interpretation of the referential meaning. If "se lo" leaves the hearer to infer from the context the number of the dative, "se los" leaves the hearer to infer which participant is referred to by the plural

morpheme (Compare 40.b(i) above, repeated below for convenience as (41.a) with (41.b):

(41) a. Juan compró UN DEPARTAMENTO$_i$ PARA SUS HIJOS$_j$ >
         > **Juan SE$_j$ LOS$_i$ compró.**
         'Juan bought an apartment for his sons.'

      b. Juan compró UNOS DEPARTAMENTOS$_i$ PARA SU HIJO$_j$
         > **Juan SE$_j$ LOS$_i$ compró.**
         'Juan bought some appartments for his son.'

The framework traditional grammarians draw on leaves unexplained why one inference would be more burdensome than the other. The reason I will propose is that this "innovative" marking tendency derives from general patterns of agreement that accommodate along the Animacy Hierarchy, with direct dative arguments (those introduced by *a*) taking prevalence over accusative ones.

Another challenge for the "ambiguity resolving" analysis derives from the fact that, as noted by Company (1998: 536), it is often the case that the dative NP to which *se* is referring occurs in the same sentence or in the sentence immediately preceding the clitic cluster, or that its referents are the hearers themselves, all contexts which render the number clarification seemingly "unnecessary". This can be illustrated with an example taken from an interview in the corpus of "El Habla Culta de la Ciudad de Buenos Aires" (Barrenechea 1987; interview 11):

(42) Es <u>la anécdota</u> más pintoresca que yo tengo porque jamás me pasó una cosa así en mi vida. Tengo muchas, ¿no?, pero esa me parece que es más divertida para contár<u>selas</u> <u>a ustedes.</u>

      It is the most colorful anecdote that I have, because it never happened to me a thing like this in my life. I have a lot, you see?, but that one it seems to me that it is funnier to tell you all about.

The conventional approaches also depend on the assumption that the case roles represented by the clitics remain unchanged. Under this view, *se* is still dative, and *lo*

still accusative, even if "enriched" by the presence of the plural morpheme coding the number of the dative NP. In general, these researchers' conjecture is that we are dealing with a case of "*number* attraction":

> "En el habla general (...) en la combinación del pronombre *se* con el pronombre *lo*, es frecuente que el valor de pluralidad de *se* pase a *lo*" (Vidal de Battini, Elena. 1964. El español de la Argentina. 2nd Edition. Buenos Aires: Consejo Nacional de Educación. p. 183).

> In colloquial speech (...) in the combination of the pronoun *se* with the pronoun *lo*, it is frequent that *the value of plurality* of *se* gets transferred to *lo*".

> "Se trata insistentemente de indicar la pluralidad del complemento indirecto *se* añadiendo una *s* al complemento directo que sigue inmediatamente, *lo* o *la* (...) Semejante atracción de número por parte del complemento (...) es corriente en el español de América" (Kany, 1945:141).

> Speakers insistently try to indicate the plurality of the indirect complement *se* adding an *s* into the direct complement that immediately follows *lo* or *la* (...) Such a *number attraction* on the part of the complement (...) is common in American Spanish.

Likewise, the grammar of the Spanish Royal Academy (RAE, 1999:1571): "el plural que se observa en el complemento directo es en realidad el plural del complemento indirecto" (the plural observed *in the direct complement* is actually the plural of the indirect complement).

An alternative perspective arises if one considers this phenomenon not as an instance of "number attraction" but of "*case* attraction", in which *los* appears in the plural not because *se* cannot take it, but because in these constructions *lo* is behaving as (or showing some properties of) a *dative* pronoun. Number agreement would be, under this view, just an epiphenomenal manifestation of an alteration in the *case* properties of the clitic cluster. Let us consider now, therefore, the alternative analysis that has been proposed along these lines.

4.1. THE CASE FOR "CANNIBALISTIC" DATIVES

Company (1998) links three innovative processes in Spanish, all of which have in common a reinforcement of a dative-marking tendency. One of them, traditionally known as "leísmo", consists of using the dative *le* instead of *lo* as the "direct object" of transitive verbs (p. 532):

(43)   Este coche está sucísimo. –Pues ya lávaLE
         'This car is very dirty. –Well, wash it (DAT) right now!'

The second phenomenon concerns the loss of prepositional case-marking (*a*) in personal accusative NPs when a dative-indirect object NP occurs with it in ditransitive sentences. The result resembles a (morphological) Primary-Object pattern, with the preposition *a* marking the patient of monotransitive clauses in one case and the recipient of ditransitive clauses in the other. Company illustrates with the following examples:

(44)   El emperador entregó *A* SU HIJA
         A SU HIJA = ACC-DO
         'The emperor gave (away) his daughter'.

(45)   El emperador entregó Ø SU HIJA *A* LOS CONQUISTADORES
         SU HIJA = ACC-DO + A LOS CONQUISTADORES = DAT-IO
         'The emperor gave the conquerors his daughter'.

The third phenomenon presented by Company is the most relevant to the issue under discussion in this paper; the plural marking of *lo* when following a referentially plural *se*. The motivation offered by Company's account is threefold. First, *se* is the only unstressed pronoun not transparent as to the number of the referent (unlike *lo, los, la, las, le, les*), yielding an unbalanced paradigm. Second, this dative clitic is homonymous with the reflexive pronoun, which generates cases of ambiguity between dative and reflexive readings in certain constructions:

(46)   Pedro debe un viaje a Europa a su esposa y a su hija → Pedro SE LO debe.
            'Peter owes his wife and daughter a trip to Europe'.

(47)   Pedro se debe [a sí mismo] un viaje a Europa → Pedro SE LO debe.
            'Peter owes himself a trip to Europe'.

Finally, by using the innovative cliticization speakers are, according to this hypothesis, indicating the higher cognitive prominence of dative referents (prototypically humans, specific animates, willing recipients, beneficiaries or experiencers) over accusative ones (typically inanimate beings, objects or abstract concepts), and thus "manifesting their own evaluation of linguistic forms" (idem:539). As is apparent, Company's idea is that in these constructions the "plural" morpheme is not merely coding the number properties of the dative. On the contrary, she proposes that it has been reinterpreted, and that it "[has] added the value of animacy-humanness typical of datives to its plurality" (idem:544). The three innovative phenomena that she reviews can be seen, therefore, as "a struggle of two different entities for the same semantic space, that of Prominent Object". The term that she coins to refer to this process derives from this tendency of Spanish datives to "usurp, 'eat' the form of accusatives"; hence the term "cannibalistic". Company proposes that the phenomena she reviews have the outcome of restructuring the Spanish case-marking system in terms of Primary-object – Secondary-object distinctions, at least in these constructions[13].

Company's hypothesis involve the idea of regarding this phenomenon as a change in progress, in which the dative advances projecting first the plural marker, then its gender, and finally gives way to the total displacement of the accusative clitic for the dative *les* (DAT+NUMBER > DAT+NUMBER+GENDER > COMPLETE DAT). However,

---

[13] Regarding Spanish as a primary-object language, there have been some proposals that relate the indirect object "doubling" construction in Spanish with the "dative shift" constructions in English (cf. Bleam (2001 and references therein).

we need to point out that in the corpus of "El Habla Culta de la Ciudad de Buenos Aires" (HCBA) no instances of gender agreement with the dative argument were attested, not even in the cases in which the dative referred to a human, which adds to the referential value of grammatical gender.

Company (2001:15) makes a stronger claim regarding the case features of *se lo* (emphasis mine):

> "The new cliticization behaves as a lexicalized, single, basically unanalyzable form: *selos*, *selas*, *seles*; in other words, *selos*, *selas*, *seles*, constitute a simplified structure, having only one object pronoun, only one argument, the Dat, which is the only pronoun that emerges morphologically, while the Acc remains inert in this grammatical area. **Data from Rivarola (1985) and De Mello (1992: 171, 174) show that most Spanish speakers do not recognize two objects in this clitic sequence**, but they interpret the sequence *selos*, *selas*, *seles* as having only one object, the Dat recipient."

We have already raised one caveat to this claim. In the HCBA corpus the dative does not seem to be the only form that emerges morphologically, since *lo* continues to behave as accusative regarding gender agreement[14]. More intriguingly, the alleged "data" from Rivarola and De Mello showing that the sequence has become lexicalized as a dative pronoun does not seem to be particularly strong: Rivarola's is a four-page comment on the phenomenon, in which in fact no explicit mention of this "lexicalization" appears. The "lexicalization" hypothesis is an interpretation done by De Mello (idem:171), based on Rivarola's article. In it, however, no position seems to be taken with regards to the notion that speakers interpret the lexicalized unit as "having only one object, the Dat recipient":

---

[14] If one accepts the hypothesis that these results show just the first stages of this change in progress, it would be interesting to investigate why it is in this order that the "cannibalization" of grammatical features occurs.

"...'Quizás sea más plausible pensar en que el fenómeno es tan general y está tan estabilizado que se pluraliza el *lo* a pesar de que el *se* sería unívoco gracias a la construcción preposicional' ['It may be more plausible to think that the phenomenon is so general and stabilized that *lo* is pluralized despite the fact that *se* would be unambiguous thanks to the prepositional construction']. Rivarola's comment suggests that the *se lo* construction has become lexicalized, so that the Spanish speaker considers *se lo* as a singular lexical item"[15].

My impression is, on the contrary, that *los* is acting as a Janus' face, pointing in two directions at once (coding properties of both the dative *and* the accusative) and that it is precisely this ambiguity what presents an interesting challenge to grammatical formalizations.

Another problem comes from the implications generated by a discussion of "number" agreement, in general. Consider the following table (Figure 2), stating the four possible alternatives in terms of number combinations between dative and accusative referents, based on my own analysis of the data from the HCBA corpus:

		ACCUSATIVE	
		singular	plural
DATIVE	singular	lo (100%) *los (0%)	*lo (0%) los (100%)
	plural	lo (44%) *los (56%)	*lo (0%) los (100%)

**Figure 2: Percentage of orthodox and innovative (*) cliticization
in clitic sequences according to the number of the referent.**

When the two objects are singular, we attest, not surprisingly, the form *lo* in 100% of the cases. When both objects are plural, there is also a single form, *los*. Both instances are trivial, since there is no way of identifying what NP the number feature is agreeing with. The third box (highlighted in bold) in represents the instances in which the

---

[15] This is also the idea presented in the Grammar of the RAE (p. 1258), but again, no evidence is provided. The RAE, however, refrains from claiming that the sequence refers only to the dative.

accusative clitic takes the plural feature of the dative (56%), and the ones in which it does not (44%). However, if as Company suggests, speakers were interpreting "the sequence *selos*, *selas*, *seles* as having only one object, the Dat recipient" we could expect that the remaining case (plural accusatives and singular datives) would show also some divergence in terms of number marking, with at least some cases in which the number features of the dative would prevail. In these cases, however, the markedness of plurals over singulars seems to be interacting with the "innovative" cliticization, and speakers do not replace the marked (plural) form with the unmarked (singular) one. In all of these instances the clitic agrees in number with the accusative NP exclusively.

An alternative account of the sequence *se lo(s)* is that proposed by García (1975), which has the advantage of providing an independent motivation for the use of *se* instead of *le(s)* in these constructions, and which can also be interpreted as presenting implicit counter-arguments for the "lexicalized" hypothesis. García claims that the "reflexive" *se* and the "pseudo-*le*" *se* should not be analyzed as homonymous, inasmuch this generates a situation in which the similarity between two semantically unrelated forms is greater that the similarity between two allomorphs. This "resemblance", she argues, does not refer to "the coincidence of sound –which must be granted in any homonymous analysis- but [to] the much more interesting and significant fact [that this form] is invariable in number" (idem:410).

As an alternative, García posits a common abstract meaning for the pronoun; that of "OTHER, LOW DEIXIS"[16], which is "ideally suited to play down or defocus an entity that would normally deserve or claim focus, such as a human agent" (idem:214). Thus, in a situation as complex as one involving three different participants (one coded

---

[16] Her proposal implies a radical redefinition of the Spanish pronominal system in terms of the categories *speaker, hearer, other; high, low deixis; focus, non-focus; and most, less and least active*. I will only address here the aspects that are directly relevant for the data under analysis, inviting the reader to resort to García (1975) for a thorough presentation of her proposal.

by the verb ending and the other two by clitic pronouns), it is claimed that the form *se* has the advantage of diminishing the inferential complexity presented by the construction, since it downplays one of the participants; something that *le* could not do.

Moreover, García argues that it would be only in *le lo that two explicit case meanings are used, since in every other clitic sequence in the language we find at most one case-explicit clitic, the other clitic being left to find its case role by inference. Under this analysis *le* is full of unhelpful precision, "a precision that is not only unnecessary but in fact stands in the way of a prompt and easy inference of the intended message". *Se*, on the contrary, appears as the ideal substitute, "in that it means third person and lacks the precise case information that makes *le* at the same time so unnecessary and so troublesome. In fact, if *se* is substituted for *le*, exactly the same routine can be followed in the third person as is always followed with *me* and *te* [all inferentially assigned Dat in clitic sequences]" (idem:416-8)[17].

The possibility that the "pseudo-*le*" had been synchronically reinterpreted as the so-called "reflexive" is a suggestive one. However, it faces one problem. If that were the case, we would expect the innovative cliticization to occur in both reflexive and ditransitive constructions. However, when "se" is reflexive, the translation of plurality to the accusative clitic does not occur:

(48)   Ellos se compraron un libro → *Ellos se loS compraron.
        'They bought themselves a book'.

These instances show that "reflexive" and "pseudo-*le*" cannot be analyzed as a single form with exclusively pragmatic meaning. Again, the phenomenon of this "innovative" cliticization in Spanish can only be accounted for by frameworks that

---

[17] This account not only has the advantage of providing an explanation for the avoidance of *les lo*; it is in turn consistent with RRG's assumption that "the speaker will choose a form for the sentence that will allow the hearer to create the proper (i.e. most relevant) context of interpretation with the least amount of processing effort" (Van Valin and LaPolla 1997: 201).

encompass cognitive and pragmatic factors into the examination of syntactic structure. The model provided by Role and Reference Grammar, which assumes a *communication-and-cognition* perspective, presents itself as a good candidate for capturing the formal properties of these "deviant" constructions without disregarding the communicative principles that make them arise. In the following section, therefore, I will summarize the RRG model, to finally propose a tentative analysis of Spanish ditransitive clauses with innovative cliticization and explain the complex interaction of clitics and NPs in three-place verbs constructions to which I referred at the beginning of the paper.

5. THREE-PLACES PREDICATES IN ROLE AND REFERENCE GRAMMAR

5.1. PRESENTATION OF THE MODEL

RRG proposes a direct mapping between the syntactic and the semantic representations of the sentence, and only one level of syntactic representation. The CLAUSE is conceived as a layered structure of grammatical units ("Layered Structure of the Clause", LSC), with a CORE consisting of the verb or other predicating elements (the NUCLEUS), as well as its direct and oblique arguments (ARG), and a PERIPHERY, linked to the CORE, incorporating the adjuncts. Consider the following illustration of a canonical monotransitive sentence:
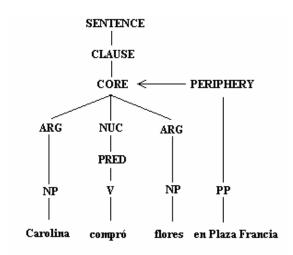
```
                          SENTENCE
                             |
                           CLAUSE
                             |
                   CORE  ←————— PERIPHERY
                  ╱   |    ╲        |
               ARG   NUC   ARG      |
                      |     |       |
                     PRED   |       |
                      |     |       |
               NP     V     NP      PP
                |     |     |       |
            Carolina compró flores en Plaza Francia
```

Figure 3: The Layered Structure of the Clause.
('Carolina bought flowers in Plaza Francia')

Whereas these elements are claimed to be semantically motivated and universal, there may be pragmatically motivated, language specific additional elements beyond the ones represented in Figure 3. Question words in Spanish appear in a clause-initial position which is distinct from the CORE-internal slot occupied by the subject. This position is called "Pre-core Slot" (PrCS), and it is also the one that is occupied by "pre-posed" focal or "left-dislocated" topical elements as the ones reviewed in section 2.1. The phrase "flores" is in the PrCS in "FLORES compró Carolina en Plaza Francia" ('It is FLOWERS what Carolina bought in Plaza Francia'). The PrCS is inside the CLAUSE but outside the CORE, and as we mentioned before, there is no intonation break or pause separating its elements from the rest of the sentence, nor there is a resumptive pronoun within the clause referring to these elements when they function as semantic arguments of the clause.

Further, there is also a "Left-detached Position" (LDP), inside the SENTENCE but outside the CLAUSE in which "topicalized" elements occur. Both the "hanging topic" and the "left-detached" constructions mentioned in section 2.1. occur in it. Contrary to elements in the PrCS, elements in the LDP are separated from the clause by a pause or

34

intonation break, and there must be a resumptive pronoun in the clause when they function as its semantic arguments. "Las flores" is in the LDP in the sentence "Las flores las compró Carolina en Plaza Francia" ('The flowers, Carolina bought them in Plaza Francia'). Likewise, there are languages in which these elements may also occur in sentence final position. In this case, we will talk about a "Post-core Slot" (PoCS) a "Right-detached Position", respectively. As we will see below, in Spanish both linear orders are possible. An abstract representation of a sentence containing the pre- and post-core slots and the detached positions is given below:
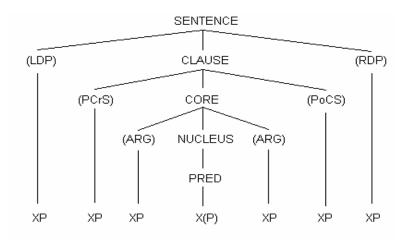

Figure 4: The Layered Structure of the Clause (revised).

With respect to the semantic structure of the clause, it is captured by a model of lexical decomposition of verbs based on Vendler's *Aktionsart*, and distinguishes five basic types: states, activities, achievements, accomplishments and active accomplishments, and a causative version of each one of them, with semantic and syntactic tests determining the appropriate *Aktionsart* for each clause.

The following chart with the lexical representation of each type, or "Logical Structure" (LS) is taken from Van Valin and LaPolla (1997:109):

Verb Class	Logical Structure
State	**pred´** (x) or (x, y)
Activity	**do´** (x, [**pred´** (x) or (x, y)])
Achievement	INGR **pred´** (x) or (x, y), or INGR **do´** (x, [**pred´** (x) or (x, y)])
Accomplishment	BECOME **pred´** (x) or (x, y), or BECOME **do´** (x, [**pred´** (x) or (x, y)])
Active Accomplishment	**do´** (x, [**pred$_1$´** (x, (y))]) & BECOME **pred$_2$´** (z, x) or (y)
Causative	α CAUSE β, where α, β are LSs of any type

The semantic interpretation of an argument will be a function of its position in these LSs. Since the system contemplates predicates with a maximum of two arguments, three-place verbs are represented by complex LSs of the type: [do' (x, Ø)] CAUSE [BECOME pred' (y, z)].

The argument positions will be linked to generalized AGENT-like and PATIENT-like thematic roles, captured by the notions of ACTOR and UNDERGOER, respectively. In order to determine the semantic macroroles of particular argument positions, the theory proposes the following Actor-Undergoer Hierarchy (AUH, Van Valin and LaPolla 1997:146):
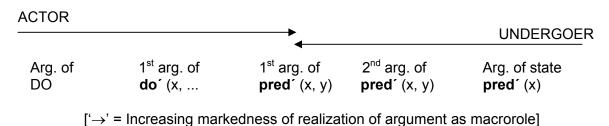
ACTOR
————————————————————————→   UNDERGOER
                        ←————————————————————————

Arg. of DO	1st arg. of **do´** (x, ...	1st arg. of **pred´** (x, y)	2nd arg. of **pred´** (x, y)	Arg. of state **pred´** (x)

['→' = Increasing markedness of realization of argument as macrorole]

Figure 5: The Actor-Undergoer Hierarchy [AUH].

This hierarchy states that 'argument of DO' (AGENT) is the unmarked choice for ACTOR, and 'argument of pred' (x)' (PATIENT) is the unmarked choice for UNDERGOER. A marked choice for ACTOR is possible only if the higher-ranking arguments are not present in the clause. Conversely, a marked choice for UNDERGOER will account for phenomena such as English's "dative shift", or the few Spanish verbs that present locative alternation (e.g. "Cargó las cajas en el camión" vs. "Cargó el camión con cajas"; '(He) loaded the boxes in the truck' vs. '(He) loaded the truck with boxes'). In any logical structure with two arguments, the leftmost will be the ACTOR and the rightmost the UNDERGOER. When macroroles occur as core arguments, they are always direct. The only instance of an oblique macrorole is the (optional) actor of passive constructions.

RRG characterizes transitivity in semantic terms; i.e. it refers to the number of MACROROLES a verb takes. According to the Default Macrorole Assignment Principles, the number of macroroles a verb has is less or equal to the number of arguments in its LS. Since there are only two macroroles, there is no notion of "ditransitivity" in the theory.

Moreover, grammatical relations such as subject, direct object and indirect object have no theoretical status in RRG. For the traditional notion of subject, RRG uses the concept of Privileged Syntactic Argument (PSA). In turn, the notions of DIRECT and OBLIQUE CORE ARGUMENTS account for traditional DO and IO, respectively. For accusative languages such as Spanish, the PSA in active constructions corresponds to the highest-ranking direct core argument, according to the following Selection Hierarchy:

Arg.ofDO>1^{st}arg.of do'(x, …)>1st arg. of pred'(x,y)>2^{nd}arg of pred'(x, y)>arg of pred'(x)

Figure 6: PSA Selection Hierarchy.

The linking between syntax and semantics is captured by Linking Algorithms that specify the appropriate mapping between the syntactic and semantic representations of the sentence (idem:326, 340-1). These linking algorithms are governed by the Completeness Constraint, which states that all the specific arguments in the semantic representation of a sentence must be realized in the syntax, and that all the referring expressions in the syntax must be linked to a position in the semantic representation. Case and preposition assignment rules determine the marking on each NP. Below preliminary Spanish case-marking and preposition assignment rules are provided:

Case marking rules for Spanish:
a. Highest ranking core macrorole takes nominative case.
b. Other core macrorole takes accusative case.
c. Non-macrorole direct core arguments take dative as their default case[18].

Preposition assignment rules for Spanish three-place predicates:
a. Assign *para* to non-MR y argument in LS segment: …BECOME **have**' (y, z)
b. Assign *en* to non-MR y argument in LS segment: …BECOME **be-LOC**' (y, z)

Figure 7: Case and preposition assignment rules.

The last piece of information that we need for our present purpose relates to the different morphosyntactic means for expressing the discourse-pragmatic status of the elements in a sentence. As we mentioned in section 3.1, this can be captured in terms of Focus structure, which affects the type of referring expression that will be selected to fill an argument position in LS according to the following markedness hierarchy (Van Valin 2001):

---

[18] Regardless of their macrorole status, I am considering *a* as coding direct core arguments: either the second [animate] argument position of pred' (x,y) in a monotransitive clause (whose verb is optionally marked by the clitic *lo*), or the first argument position pred' (x,y) in a ditransitive clause, (whose verb is in turn marked by the clitic *le*).

Zero   Clitic/Bound Pro   Pro [-Stress]   Pro [+Stress]   Definite NP   Indefinite NP

Markedness of Occurrence as Topic

Figure 8: Coding referents in terms of possible functions.

Therefore, if the ACTOR and the UNDERGOER are topical, the unmarked realization of these arguments will be the verb-ending for the first and the accusative clitic for the second. In these cases, only a bundle of pronominal features will occur in the arguments positions of the LS, which will be linked to the syntactic representation via an "Agreement Index" node, as explained in section 6. In that section I will also argue that the type of expressions selected is further affected by the activation levels of the referents in discourse representation.

5.2. SPANISH THREE-PLACE PREDICATES

As mentioned above, the system of lexical decomposition employed in RRG imposes a complex LS for the representation of three-place verbs. Some examples with Spanish verbs are given below:

dar	[**do**' (x, Ø)] CAUSE [BECOME **have**' (y, z)]
mostrar	[**do**' (x, Ø)] CAUSE [BECOME **see**' (y, z)]
enseñar	[**do**' (x, Ø)] CAUSE [BECOME **know**' (y, z)]
poner	[**do**' (x, Ø)] CAUSE [BECOME **be-LOC**' (y, z)]

Spanish does not have, unlike English, variable UNDERGOER selection for transfer verbs: the UNDERGOER is always the rightmost argument in the LS. What seems to be more interesting in three-place verbs are the alternations in the coding of the third argument, which can occur either as oblique, marked by *para* or *en*, or as direct

39

argument, marked by *a*, and coindexed with the co-occurring dative clitic *le(s).* The examples in (49) illustrate the grammatical and ungrammatical possibilities for "Juan bought the bicycle for his sons":

(49)   a. Juan compró la bicicleta *para* sus hijos.

              b. Juan *les* compró la bicicleta *a* sus hijos.

              c. *Juan compró la bicicleta a sus hijos.

              d. *Juan les compró la bicicleta para sus hijos.

The example in (b) could be thought of as a type of applicative construction, in which a non-argument, the beneficiary NP *para sus hijos* (in (a)), is treated as a direct core argument of the verb (and therefore marked by *a*). When treated as a direct core argument, the presence of the clitic is obligatory (c)[19]. The example in (d) shows that *les* cannot co-occur with obliques, which suggests that the syntactic valence of the verb has increased from 2 to 3. However, it does not mean that the semantic macrorole assignment has changed. Crucially, the only passive forms of (49) are the ones presented in (50), with *la bicicleta* (the bicycle) as passive-PSA, (note the same complementary occurrence of *les* and *para*):

(50)   a. La bicicleta fue comprada (por Juan) para sus hijos.

              b. La bicicleta les fue comprada (por Juan) a sus hijos.

              b'.*Sus hijos fueron comprados la bicicleta (por Juan).

---

[19] Jaeggli, among others, claims that the presence of the dative clitic is "obligatory" in all cases except when the "dative" argument codes a Goal, as in the following sentence:

      A las doce en punto, la dirección (les) entregó las notas a los estudiantes.
      'At 12 o'clock sharp, the administration gave the grades to the students'.

In these cases, he says that the presence of the dative clitic is "optional, but highly preferred" in the Rio de la Plata dialect (1981:12). In this paper I will lump together the categories "obligatory" and "optional but highly preferred"; leaving for a later analysis the precise quantitative account of this difference.

40

A similar process affects LOC verbs, such as *poner* (put). Consider the two versions of 'Paula put the adornments on the Christmas tree':

(51)   a. Paula puso los adornos *en* el árbol de navidad.
           b. Paula *le* puso los adornos *a*l árbol de navidad.

When these predicates take a non-macrorole direct core argument marked by *a* (51.b), they are subject to innovative cliticization. Below is an example of this tendency taken from the HCBA corpus:

(52)   ...cualquier cosa que se trate de sacar adelante hay que llevarla totalmente cocinada y ponér<u>selas</u>, [al tesorero y al contador] es decir: "Aprieten este botón, pongan la coma acá y hagan esto." (15.17)

'...anything that one tries to put forth one has to bring it totally cooked up and present it to them (lit. put-them-it) [the treasurer and the accountant] I mean: 'Push this button, put the semicolon here and do this'.'

6. A PRELIMINARY RRG ACCOUNT OF CLITICIZATION IN SPANISH

Spanish is a dependent-marking language with, arguably, certain head-marking properties, exemplified by the fact that, as Lenz put it, "el verbo encierra en sí todo el régimen de la oración" (1920:54-55); that is, the verb stem, plus its subject and object agreement markers, can constitute a sentence on its own. The proposals we reviewed touched on this issue, inasmuch they took different positions regarding whether to treat object clitics as inflectional affixes on a par with verb-subject agreement (preserving the basic dependent-marking nature of the language), or as pronominal arguments (likening Spanish to head-marking languages). The latter alternative fares well when one has to account for the fact that the clitics can be the only representation of the verbs' complements that appear in the sentence, but it needs to resort to abstract levels of representation ("movement" and "theta-role/case absorption" rules) to accommodate the

cases of "doubling". The former has the advantage of explaining more naturally the "doubling" cases (and, moreover, to potentially allow for a unified account of subject and object agreement), but still needs to posit phonologically null categories (such as "*pro*") in order to explain how the valency of the verb is fulfilled when the only overt manifestation of the verb's complements in the sentence are these "agreement affixes".

If on the other hand one wants to avoid "abstract" syntax, this ambiguous nature of Spanish clitics, and similar cases in other "double-marking" type languages, may be accounted for in a construction-specific basis. This is the alternative favored for instance in Bresnan and Mchombo's (1987) analysis of Chicheŵa (a Bantu language spoken in East Central Africa). Their analysis presents evidence in support of Givón's claim that the grammatical agreement between a verb and its subject and object cannot be distinguished from the anaphoric relation between a morphologically bound pronoun and a discourse topic, and they propose therefore that both phenomena should be accounted for by means of a unique mechanism, assigning different functions to these "agreement markers" depending on the construction in which they occur. Within the framework of Lexical-Functional Grammar, Bresnan and Mchombo claim that Chicheŵa's verbal affixes may mark either grammatical or anaphoric agreement. In the first case, the NP bears an argument relation to the verb, whereas in the second it is the affix, as an incorporated pronoun, which functions as the verb's argument, with the correferential NP marked as topic. In particular, they propose that Chicheŵa's obligatory "subject marker" receives a pronominal interpretation when there is no subject NP in the phrase structure, and therefore it is ambiguously used either for grammatical or anaphoric agreement. Chicheŵa's optional "object marker", on the contrary, is used for anaphoric agreement exclusively, and when the correspondent "object" NP co-occurs with it, it does not function as an object but as a topic, bound by the incorporated "object marker" pronoun on the verb. In a similar vein, Van Valin and LaPolla (1997:331-2)

suggest that, in "double-marking" type languages, the independent NP counts as the core argument if present, with the bound marker on the verb functioning as the verb's argument otherwise.

I will propose a modification to this view, and argue that, regardless of whether they co-occur with independent NPs or not, Spanish clitics (as well as the "PSA agreement" on the verb) should be linked to an "agreement index" node (AGX). The AGX is a dependent of the NUCLEUS, and it receives the agreement specifications of all core argument positions present in the Logical Structure. By positing this new node, we can capture the fact that the verb-governed dependents can be omitted without affecting the grammaticality of the phrasal unit, while at the same time taking into account the relevance of the lexical NPs for all sort of grammatical phenomena that cannot be accounted for without explicit reference to their behavior.

We mentioned before that Completeness Constraint (CC) states that all the specified arguments in the semantic representation of a sentence must be realized in mophosyntactically, and all the referring expressions in the syntax must be linked to a position in the semantic representation. However, I am suggesting that in cases in which the only expression of the verb's arguments are the PSA agreement on the verb and the object clitics, the positions in the semantic representation are linked not to an argument node but to the AGX. Therefore, it is necessary to modify the CC in order to allow elements in the AGX to satisfy it as well. In exchange, once done this we may be able to capture the characteristics that liken Spanish to head-marking languages without violating its basically dependent-marking nature.

Following Everett's proposal (personal communication), and consistent with an "inferential-realizational" approach to morphology (Stump, 2001), I will assume that, in cliticized constructions, the arguments positions in LS are represented by a bundle of agreement features assigned to the AGX node and further interpreted by a morpho-

phonological rule that spells out their correct realization. Stump distinguishes between "lexical" and "inferential" theories of inflection. The first assumes that inflectional morphemes have their own entry in the lexicon, associated with the morphosyntactic features they express. The second posits that there is a rule associating a particular affix with the presence of certain morphosyntactic properties in the lexical root. Further, whereas "incremental" theories propose that it is the addition of the inflectional affixes that allows the root to acquire their morphosyntactic properties, "realizational" views presuppose that it is the word's association with this set of morphosyntactic properties that "licences" the attachment of the corresponding affixes. For instance, from an "inferential-realizational" view, the inflected English word "likes" results from the application a realization rule appending –*s* to any verb stem associated with the properties '3sg. subject agreement', 'present tense', and 'indicative mood' (2001:1).

In our case, the rules that generate the corresponding agreement markers on the verb depend on the information linked to the AGX: the PSA agreement depends on a rule that interprets the information coded by the features marked for nominative, and spells them out in the verb-ending. Undergoer agreement depends on the agreement features marked for accusative, and non-macrorole agreement, on the ones marked for dative. For instance, the presence of a feature bundle expressing '3 person', 'masculine' and 'plural' in an argument position marked for accusative case in the Logical Structure licenses a realization rule that generates "los" in the first position before the verb's stem. A feature bundle expressing '3 person', 'plural' in an argument position marked for dative yield "les" (or "se" if accusative features are also present), in the preceding verbal slot, and so on.

Further, this approach allows us to capture the "innovative" marking reviewed in section 4 in terms of the rules realized by the AGX. The idea is that the rules "read" the information in the AGX in a fixed sequence, first the features from the actor, then the

ones from the undergoer, and finally the ones from the non-macrorole: NOM:person:number> ACC:person:gender > DAT:person> NUMBER. If the ACC information does not fill the plural slot (i.e. if it is marked "-plural") the plural features from the DAT may be spelled out in this available position, yielding the sequence "se lo/a-s" preferred by "innovative" speakers.

Below is a preliminary version of the Linking Algorithms for Spanish, including a step that links the agreement specifications in Logical Structure and the AGX in the Syntactic Template.

Preliminary Spanish Linking algorithm: Semantics → Syntax

1. Construct the semantic representation of the sentence, based on the LS of the predicator.
2. Determine the actor and undergoer assignment following the AUH (Figure 5).
3. Determine the morphosyntactic coding of the arguments
   a. Select the PSA, based on the PSA selection hierarchy (Figure 6).
   b. Assign the XPs the appropriate case markers and/or adpositions (Figure 7).
4. Select the appropriate syntactic template(s) for the sentence.
5. Assign the arguments to positions in the syntactic representation of the sentence.
   a. Assign the pronominal features of the arguments to the AGX.
   b. Assign the [-WH] strong pronouns or full NPs, if any, to the appropriate positions in the clause.
   c. If there is a [+WH] XP,
      1. Assign it to the normal position of a non-WH-XP with the same function, or
      2. Assign it to the precore slot, or
      3. Assign it to a position within the potential focus domain of the clause (default = the unmarked focus position).
   d. A non-WH XP may be assigned to the precore or postcore slot, or left or right detached positions, subject to focus structure restrictions.
   e. Assign the XP(s) of LS(s) other than that of the predicator in the nucleus to
      1. the periphery (default), or

2. the precore or postcore slots, or

3. the left or right-detached positions.

Preliminary Spanish Linking algorithm: Syntax → Semantics

1. Determine the macrorole(s) and other core argument(s) in the clause.

    a. If the verb is intransitive, assign the PSA macrorole status.

    b. Determine the voice of a transitive verb:

        1. If it is the unmarked voice, the PSA is actor.

        2. If it is passive, the PSA is undergoer.

            a. The actor may appear in the periphery marked by *por*.

            b. If there is no actor in the periphery, replace the variable representing the highest ranking argument in the LS with Ø.

    c. Assign macrorole status to the other direct core argument, if it is not dative.

2. Retrieve from the lexicon the LS of the predicate in the nucleus of the clause and with respect to it determine the actor and undergoer assignments, following AUH (Figure 5), subject to the following proviso:

    a. Determine the linking of the non-macrorole core argument:

        1. If there is a two-place state predicate in a complex LS, and if the non-macrorole core argument is dative, link it with the first argument position in the state predicate.

        2. If there is a two-place state predicate in a complex LS, and if the non-macrorole core argument is marked by a locative or purposive adposition, link it with the first argument position in the state predicate.

3. Determine the linking of the agreement features of the predicate in the nucleus:

    a. Link the nominative agreement features to the actor.

    b. Link the accusative agreement features to the undergoer.

    c. Link the dative agreement features to the non-macrorole.

4. Link the arguments determined in step 1 with the arguments determined in step 2 and the agreement features determined in 3 until all core arguments are linked.

5. If there is a predicative adpositional adjunct, then retrieve its LS from the lexicon, insert the LS of the core as the second argument in the LS and the object of the adposition in the periphery as the first argument.

6. If there is an element in the pre- or postcore slot,

    a. Assign it to the remaining unlinked argument position in the semantic

representation of the sentence.

b. If there are no unlinked argument positions in the sentence, then treat the WH-
word like a predicative preposition and follow the procedure in step 4, linking the
WH- word to the first argument position in the LS.

The following example illustrates the role of the AGX in the semantics to syntax
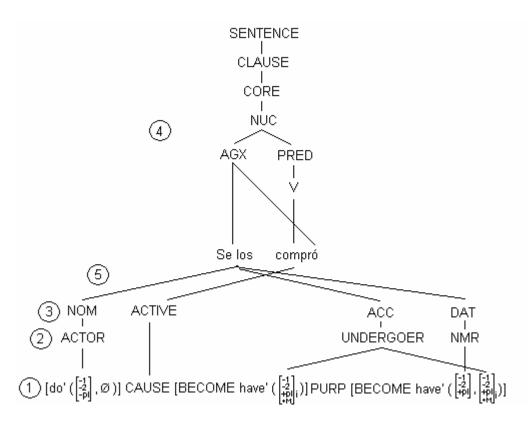linking of a fully pronominalized three-place verb construction:

Figure 9: Representation of the AGX node in the LSC and the Semantics to Syntax Linking.
('He bought them for them')

This sentence may be uttered in response to the question "What did Juan do
about the children's presents?" It is a case of a narrow focus structure, with the referents
in discourse representation highly activated, and therefore coded merely by features
bundles in the LS. The speaker wants to express that the presents were bought, and
therefore selects the semantic representation of the verb "comprar". In this example,

"comprar" shows a complex lexical representation in which the first component, do' (x, $\varnothing$), refers to the activity part, the second component, BECOME have' (z), to the accomplishment, and the third, BECOME have' (z, y), to the incorporated benefactive (step 1). The next step is actor and undergoer assignment. The leftmost argument position in the LS will correspond to the actor. The undergoer corresponds to the rightmost argument positions according to the AUH (the only position of the first BECOME have' (y) and the second position of the second BECOME have' (z,y)). This leaves the remaining argument position, *z*, as a non-macrorole argument. The third step is to determine the morphosyntactic coding of the arguments. Since the actor is represented exclusively by feature bundles, there is no lexical PSA (3.a.). Regarding the case assignment, the actor receives nominative case, the undergoer receives accusative case, and the non-macrorole receives dative. The fourth step is the selection of the syntactic template which crucially, according to this proposal, involves an AGX alongside the PRED as a constituent of the NUCLEUS. The fifth and final step is the linking of the arguments in the LS to the positions in the LSC. The verb is linked to the V node, and the feature bundles to the AGX. That is, the AGX will receive the features NOM,-1,-2,-pl. from the actor; the features ACC,-1,-2,+masc,+pl. from the undergoer; and the features DAT,-1,-2,+masc,+pl. from the non-macrorole. A morpho-phonological realization rule reads this information, spells out their realization, and places the markers in their corresponding slots, according to the description above.

We can now turn to the cases of "doubling" referred to in sections 2 and 3, and likewise explain them without having to resort to "abstract syntax". Thus, let us now attempt a semantics-to-syntax representation of a sentence in which the clitics cluster occurs with correreferential NPs:
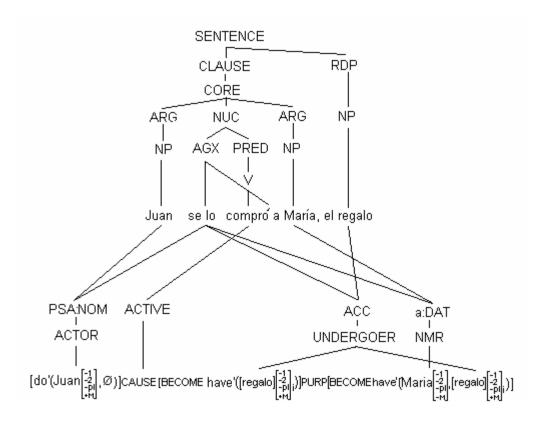
Figure 10: Clitic cluster co-occurring with independent NPs.
('Juan bought it for Maria, the present')

As we saw in 3.1., three-place verbs constructions impose a pragmatic constraint that disallows the accusative clitic being coindexed with an independent NP in the PFD. In other words, in a three-place verb construction, the only non-actor argument allowed in the core is the non-macrorole. Therefore the ACC phrase, if occurs, has to occupy a core-external position. In this case, it occurs in the RDP, since there is a pause separating the ACC phrase from the rest of the clause, and the resumptive accusative clitic is obligatory. As we also mentioned, in Spanish the neutral focal accent coincides with the last syllable of the melodic group. Therefore the preceding example may be uttered as an answer for "For whom did Juan buy the present?" with the constituent "a María" receiving neutral focal accent.

The linking is similar to the one described for Figure 9. One difference is that in this case the LS consists of both the features bundles and the lexical items that will appear in the syntactic structure, arguably due to a lower level of activation of the referents in discourse representation. In accordance with the AUH, "Juan" is selected as the ACTOR, and "regalo" as the UNDERGOER. Following the PSA selection hierarchy, "Juan" is selected as PSA, and marked for nominative case. The undergoer receives accusative case, and the non-macrorole, which will be realized as a direct core argument, receives the preposition *a* in accordance with the case assignment rules stated in Figure 7. Then the appropriate syntactic template is selected. In this case, it includes a right detached position as one of the constituents of the sentence, and two arguments positions alongside the nucleus, in order to accommodate the independent XPs: "Juan" is assigned to the first argument position in the syntactic template, "a Maria" to the second argument position, and "el regalo" to the RDP due to PFD constraints. The feature bundles in LS are assigned to the AGX and interpreted in the same fashion as the one illustrated for the example in Figure 9. The result is a representation in which the clitics co-occur with the independent NPs, linked in the same way as in a fully pronominalized sentence. The only constraints observed are due to pragmatic factors in terms of the possible Spanish focus types, and there is no need to posit abstract levels of representation nor to vary the functional interpretation of the agreement morphemes depending on the construction in which they occur.

A final issue that we need to address is whether we can advance an independent motivation for the occurrence of the clitics and/or the independent NPs. As we saw repeatedly in the previous discussion, the facts have been described quite accurately by researchers belonging to several different schools. However, there still seems to be room to recapitulate them, to focus on the less discussed characteristics of these facts in relation to three-place verbs, and to propose a more homogeneous account for their

motivations. As we saw in Figure 9, all the lexical arguments of the three-place verb can be omitted. Constructions such as the exemplified in Figure 9 correspond to instances in which the arguments are not only topical (a negative definition derived from the fact all of the syntactic structure is focal) but more importantly highly activated in the discourse representation. The presence of the PSA in its canonical preverbal position (unless receiving contrastive focal accent) expresses a topical element and, therefore, if one is constrained exclusively to the distinction between topic and focus in sentence structure, it is difficult to account for the fact that the topical element is nevertheless not so topical as to be omitted, which would be the expected result in a pro-drop language such as Spanish. Likewise, a distinction merely in terms of sentential focus structure cannot explain why accusative strong pronouns are obligatorily doubled by the accusative clitic, whereas this "doubling" is optional if the accusative phrase is expressed by a lexical item. Focus structure alone also fails to explain why is it that when the accusative clitic "doubles" a lexical phrase, such phrase is necessarily interpreted as presupposed, despite falling within the domain of the neutral focal accent in the syntactic structure.

In order to explain these phenomena, the relevant distinctions seem to be in terms of the distribution of activation levels in discourse representation, which is orthogonal to focus structure. RRG incorporates the discourse referents' activation levels as part of the information expressed in each of the argument positions in the LS. The distinction of different levels of activation is a fine grained one. Following studies by Prince (1981b) and Chafe (1987), among others, Van Valin proposes that there are, at least, five categories: *active*, i.e. actively under consideration in the discourse by means of direct mention; *accessible*, i.e. not actively under consideration by readily recognized by the addressee due either to knowledge of the world or occurrence in the immediate environment of the speech situation; *inactive*, i.e. previously mentioned but not actively under consideration and not assumed by the speaker to be recognized by the

addressee; *brand new anchored*, i.e. not previously mentioned but related to something already mentioned or accessible; and *brand new unanchored*, i.e. not previously mentioned or related to anything previously mentioned (in press:70).

A detailed study of actual corpora is needed to examine how these categories interact with the representation of the referents in the syntactic structure in the kind of constructions we have been analyzing. We may advance, however, the prediction that the argument positions will be expressed in the syntactic structure exclusively by means of verb agreement when the referents are *active*, which correlates with their representation in LS merely by feature bundles further expressed by the AGX. In the case of the actor and NMR, verb agreement will co-occur with a pronominal phrase if their argument positions are *accessible*, and with lexical phrases if *inactive* or *brand new*. With regards to the undergoer, if it is *accessible* the agreement on the verb will co-occur with a co-indexed strong pronoun, whereas if it is *inactive* it will co-occur with a lexical NP. Finally, if the argument is *brand new*, it will be represented in the syntax by the lexical phrase exclusively.

The following table summarizes the possible cases, representing whether the AGX will spell out the agreement features of each of the arguments and whether there will be lexical or pronominal phrases co-occurring with them under each condition. Highlighted, the possible instances of "direct object doubling", predicted in this model by the activation levels of the referents:

	ACTOR		UNDERGOER		NMR	
	AGX	NP	AGX	NP	AGX	NP
Active	Yes	No	Yes	No	Yes	No
Accessible	Yes	Pronom. Phrase	Yes	Pronom. Phrase	Yes	Pronom. Phrase
Inactive	Yes	Lexical Phrase	Yes	Lexical Phrase	Yes	Lexical Phrase
Brand-new anchored	Yes	Lexical Phrase	No	Lexical Phrase	Yes	Lexical Phrase
Brand-new unanchored	Yes	Lexical Phrase	No	Lexical Phrase	Yes	Lexical Phrase

Figure 11: "Doubling" Patterns and Activation Levels.

RRG incorporates the information regarding the activation level the referents in the LS. An illustration of the result of the semantics to syntax linking including this information (and therefore motivating how the arguments will be realized) is illustrated in (13). The actor argument is *active*, and therefore realized by the bundle features interpreted by the AGX. The undergoer argument is *inactive*, and accordingly realized both by features bundle occupying the corresponding argument position and by a lexical phrase that occurs in the syntax in the Left-Detached Position (the corresponding position for "left-dislocated" topicalized elements reviewed in section 3.1). The non-macrorole argument is *accessible*, and therefore the bundle features linked to the AGX co-occur with the pronominal information realized in the syntactic structure by the strong pronoun.
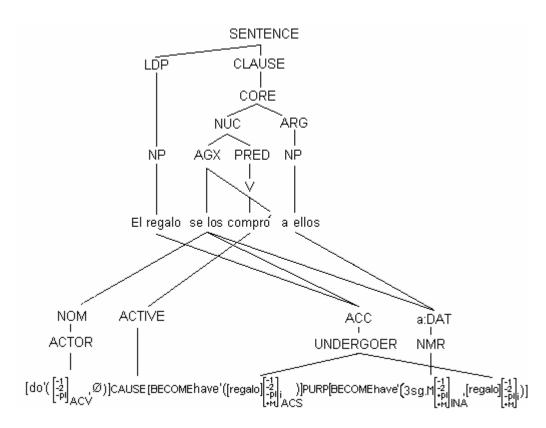
Figure 12: Coding of Activation Levels in LS
('The present (he) bought it for them')

7. CONCLUSION

In this paper I addressed several problems raised by clitic constructions in Spanish, paying particular attention to the ones that involved third-person clitic clusters. In particular, I focused in the Spanish spoken in Argentina, which presents two features that distinguish it from other dialectal varieties. One of them, studied quite extensively, is the co-occurrence of the accusative clitic and the independent accusative NP in monotransitive clauses (i.e. cases of "direct object doubling"). We reviewed the contributions to the explanation of this phenomenon advanced within the generativist framework, and in doing so we summarized different proposals varying in terms of whether the clitics are treated as pronominal heads or as agreement indexes, and

reviewed the advantages and limitations of the different "rules" or abstract representations needed to account for the constructions in which they enter.

I also mentioned attempts to approach cliticized constructions from a discourse perspective, and to explain the co-occurrence of clitics and NPs in terms of the distribution of information in the sentence or in discourse representation. We saw that some of the results presented by these studies seemed contradictory, in that "doubling" constructions were treated as basically marking either focus or topic, depending on the researcher. We also suggested that this inconsistency may be due to a lack of explicitness with regards to whether the "doubling" element is considered to be the clitic or the independent NP, and also to the tendency to lump together in the results of the corpora analysis the "doubling" of accusatives and datives which, as we suggested in section 3.1., at least in the cases in which both clitics occur, seem to respond to particular constraints in focus types.

The second "idiosyncrasy" of Argentinean Spanish (shared also with the dialects spoken in Mexico and Colombia) has received much less theoretical attention. It consists on the marking of plurality into the accusative clitic, regardless of whether it is the accusative or the dative referent in the clause or discourse representation which serves as the plural antecedent. We saw that this tendency is not an "error" ascribable to "uncultured speakers", as some traditional Spanish Grammars suggest, but on the contrary the preferred strategy in about half of the potential contexts. Further, we summarized the case for "cannibalistic" datives advocated by Company (1998) and confronted its predictions with a corpus of oral interviews edited as "El Habla Culta de la Ciudad de Buenos Aires".

Finally, I suggested that the model of Role and Reference Grammar presents itself as an ideal candidate for explaining both the formal and functional aspects of cliticized constructions, and advanced the idea of treating Spanish clitics as a result of a

realizational rule that spells out the morphosyntactic properties associated with the verb, depending on the activation level of the PSA, undergoer and NMR arguments, and the corresponding sentence's focus structure. I proposed the incorporation of an Agreement Index node dependent of the Nucleus, and a modification of the Completeness Constrain so that the features coded in AGX are able to satisfy it.

It may be also noteworthy that the cross-dialectal differences in terms of the extension of "subject", "indirect object" and "direct object" agreement coding on the verb with co-occurring independent NPs that grammatical studies refer to seems to follow the same ranking principles than RRG's PSA hierarchy. As mentioned before, Spanish finite verbs always trigger verb agreement with the PSA ("*Juan* compró el regalo"; 'Juan bought-3sg. the present'). In turn, verb agreement with the second-highest argument (a co-occurring non-macrorole direct core argument) is widespread across many Spanish dialectal varieties, although not obligatory ("Juan *le* compró un regalo *a María*"; 'Juan bought-3sg.-3sg. the present for María'). Finally, verb agreement with the lowest-ranking argument (a co-occurring (lexical) undergoer in two-place verb constructions) is limited to the Rio de la Plata dialect ("Juan *la* vio-3sg. *a María*"; 'Juan saw-3sg-3sg. María') and, further, to inactive referents. Thus, this cross-dialectal pattern seems to follow the same distribution that ranks which arguments will be privileged in particular constructions, with verb agreement being obligatory with the PSA, widespread with the second highest ranking argument ("indirect object"), and most restricted with the lowest ranking argument ("direct object"). Likewise, the same ordering seems to be responsible for the preference for coding the plurality of the second highest ranking argument (the dative), overriding the singular features of the lower-ranking one (the accusative), thus yielding the occurrence of the "cannibalistic" cliticization reviewed in section 4.

Again, a detailed study of spontaneous texts needs to be carried out to further support these ideas, and many of their details need to be refined. However, I hope that

the present proposal may open a possible way for treating these "loose threads" that

clitics seem to represent in Spanish grammar within a more homogeneous fabric.

REFERENCES

Alonso, A. and P. Henríquez Ureña.1951. Gramática Castellana. Segundo Curso. 10th Edition. Buenos Aires: Editorial Losada.

Barrenechea, A. (ed.) 1987. El habla culta de la ciudad de Buenos Aires. Buenos Aires: Universidad Nacional de Buenos Aires.

Beninca, P. 1988. L'ordine degli elementi nella frase e le costruzione marcate. In Renzi, L.(ed.) Grande Gramatica Italiana di Consultazione. Vol. 1:153-194. Bologna:Il Mulino.

Bleam, T. (2001). Properties of the Double Object Construction in Spanish. Proceedings of LSRL (31)

Bresnan J. and S. Mchombo. 1987. Topic, Pronoun, and Agreement in Chicheŵa. Language, 63(4):741-782.

Chomsy, N. 1970. Remarks on nominalization. In Studies on Smenatics in Generative Grammar, The Hague: Mouton, 11-61.

Chafe, W. 1987. Cognitive Constraints on Information Flow. In Tomlin, R. (ed.) Coherence and Grounding in Discourse, 21-51, Amsterdam: John Benjamins.

Colantoni, L. 2002. Clitic Doubling, Null Objects and Clitic Climbing in the Spanish of Corrientes. In Gutierrez-Rexach, J. (ed) From words to discourse: Trends in Spanish Semantics and Pragmatics. Amsterdam: Elsevier.

Company, C. 1998. The interplay between form and meaning in language change. Grammaticalization of cannibalistic datives in Spanish. Studies in Language, 22 (3):529-565.

Company, C. 2001. Multiple dative-marking grammaticalization: Spanish as a special kind of primary object language. Studies in Language, 25 (1):1-47.

Corbett, G. 2000. Number. Cambridge: CUP.

De Mello, G. 1992. Se los for se lo in the spoken cultured Spanish of eleven cities. Hispanic Journal, 13 (1):165-179.

Fernandez Soriano, O. 1993. Los pronombres átonos en la teoría gramatical. Repaso y balance. In Ferdandez Soriano (ed) Los pronombres átonos, Madrid: Taurus.

Flórez, L. 1977. Apuntes de español. Bogotá: Instituto Caro y Cuervo.

Franco, J. (2000). Agreement as a Continuum. In Beukema, F. and M. den Dikken (eds) Clitic Phenomena in European Languages. Amsterdam: John Benjamins.

García, E. 1975. The role of theory in linguistic analysis: The Spanish pronoun system. Amsterdam: North Holland Publishing Company.

García-Miguel, J.M. 1991. La duplicación de complemento directo e indirecto como concordancia. Verba 18:375-410.

Givón, T. 1983. Topic continuity in discourse: An introduccion. In T. Givón (ed.) Topic continuity in discourse: A quantitative cross-language study, 1-42. Amsterdam & Philadelphia: John Benjamins.

Jaeggli, O. 1981. Topics in Romance Syntax Dordrecht: Foris.

Lambrecht, K. 1994. Information structure and sentence form. Cambridge: CUP.

Lenz, R. 1920. La oración y sus partes. Madrid:Anejo de R.F.E.

Levinson, S. 1987. Pragmatics and the grammar of anaphora: A partial pragmatic reduction of binding and control phenomena. Journal of Linguistics 23: 379-434.

Kany, Ch. 1951. American Spanish Syntax. Chicago: UChP.

Kayne, R. 1975. French Syntax: The Transformational Cycle. Cambridge: MIT Press.

Monachesi, P. (ms.) The verbal complex in Romance: a case study in gramatical interfases.

Perlmutter, D. 1971. Deep and surface structure constrains in syntax. New Cork: Holt, Rinehart and Winston Inc.

Prince, E. 1981b. Toward a Taxonomy of Given-New Information. In Cole, P. (ed.). Radical Pragmatics. New York: Academia Press, 223-256.

Real Academia Española.1999. Gramática descriptiva de la lengua Española. Madrid: Espasa.

Rivarola, J. 1985. *Se los* por *se lo*. Leix, IX (2):239-244.

Rivas, A. 1977. A Theory of Clitics. Doctoral Dissertation, MIT.

Silva-Corvalán, C. 1981. The Diffusion of Object-Verb Agreement in Spanish. Papers in Romance, 3:163-76.

Silva-Corvalán, C. 1984. Semantic and pragmatic factors in syntactic change. In Fisiak, J. (ed.) Historical Syntax. Berlin: Mouton, 555-597.

Stump, G. 2001. Inflectional Morphology. A Theory of Paradigm Structure. Cambridge: CUP.

Suñer, M. 1988. The role of agreement in clitic-doubled constructions. Natural Language and Linguistic Theory 6:391-434.

Vallduví, E. 2001. La dislocació del verb en catala. in Bover, M. and M. Vidal-Tibbits (eds.) Actes del Nove Colloqui d'Estudis Catalans a Nord-America, 467-494. Barcelona: Abadia de Montserrat.

Van Valin, R. 1999. A Typology of Focus Structure and Syntax. In Raxilina, E. and J. Testelec (eds.) Typology and the Theory of Language: From Description to Explanation. Moscow.

Van Valin, R. 2001. The Role and Reference Grammar Analysis of Three-Place Predicates. Unpublished ms. [available on RRG web site].

Van Valin, R. (in press) Introduction to RRG, Cambridge: CUP.

Van Valin, R. and R. LaPolla 1997. Syntax. Structure, meaning and function. Cambridge:CUP.

Weissenrieder, M. 1995. Indirect Object Doubling: Saying Things Twice in Spanish. Hispania 78:169-177.

Zubizarreta, M. 1999. Las funciones informativas: Tema y Foco. In Bosque, I. and Demonte, V. (dir.) Gramática Descriptiva de la Lengua Española, v. 3. Madrid: Espasa.

Zwicky and Pullum. 1983. Cliticization vs. Inflection: English n't. Language 59(3):502-513.