Investigations of the Syntax–Semantics–Pragmatics Interface

Edited by

Robert D. Van Valin, Jr.
Heinrich Heine University, Düsseldorf & University at Buffalo,
The State University of New York

John Benjamins Publishing Company
Amsterdam/Philadelphia
“Floating plurals”, prodrop and agreement –
an optimality-based RRG approach

Rolf Kailuweit
Albert Ludwigs University Freiburg

In this paper, I shall try to provide a RRG description of agreement marking and "morphological" argument realization in Spanish. In Spanish, subject agreement can be expressed by a suffix, but the same suffix can realize the PSA-argument. Object clitics, too, can realize arguments or serve as agreement markers. Valeria Belloro, following the research put forth by Company, pointed out an interesting problem concerning argument realization and object clitics in standard Latin American Spanish. Plurality of the Non-Macrorole-Argument can be realized at the clitic denoting the Undergoer-Macrorole. My contribution will thus be to sketch a general RRG-framework for clitics, argument realization and agreement. As far as Belloro's problem is concerned, I would like to show that the data can be described without presuming a kind of functional head at the level of constituent structure. In fact, speakers use the optimal clitic in each speech variety. What differs is the lexical entry for the clitics, and Grimshaw's approach to Spanish clitics and optimality must therefore be reformulated. Choosing the optimal clitic is part of applying a constructional schema that comes into play as part of the linking algorithm within the RRG framework.

1. Introduction

Romance languages, with the well-known exception of French, are instances of the so-called pro-drop-parameter. In Spanish, for example, not only subject agreement can be expressed by a suffix, but the same suffix can also realize the Privileged Syntactic Argument (PSA) on its own.

(1)  
\begin{align*}
\text{a. } & \text{Maria} \_ \text{cant-ó [agreement]} \\
& \text{Mary} \_ \text{sing.3sg.PAST} \\
& \text{"Mary sang"} \\
\text{b. } & \text{Can-ó [suffix as PSA]} \\
& \text{sing.3sg.PAST} \\
& \text{"(he/she) sang"}
\end{align*}

Object clitics, too, can realize arguments or serve as agreement markers. In Spanish and some other Romance varieties, clitics and full NPs co-occur, realizing the same
argument. In Spanish, clitic redundancy is obligatory with the personal indirect object and even with the direct object when it is realized as a stressed personal pronoun:

\[(2) \quad a. \quad \text{Le, gustó el partido [a Juan]}_{1} \]
\[3\text{DAT.sg. like-3sg.past det-masc match to John} \]
'John liked the match'

\[b. \quad \text{Lo vimos [a él]}_{1} \]
\[3\text{MASC.acc.sg see-1pl.past to 3MASC.sg} \]
'We saw him'

Following Company (1998), Belloro (2004) pointed out an interesting problem concerning argument realization and object clitics in standard Latin American Spanish. Plurality of the Non-MR-Argument can be realized in the clitic denoting the Undergoer-MR:

\[(3) \quad a. \quad \text{Juan compró una casa para sus hijos} \]
\[3\text{DAT.pl. buy.3sg.past indef.fem house for 3pl.poss.pl son.pl} \]
'John bought a house for his children'

\[b. \quad *\text{Juan les la compró} \]
\[3\text{DAT.pl. 3fem.acc.sg buy.3sg.past} \]

\[c. \quad \text{Juan se la compró} \]
\[3\text{DAT.pl. 3fem.acc.sg buy.3sg.past} \]

\[d. \quad \text{Juan se las compró} \]
\[3\text{fem.acc.pl buy.3sg.past} \]
(c.f. Belloro 2004: 22, examples 40d-d(i))

My paper will sketch a general RRG framework for clitics, argument realization and agreement. As far as Belloro's problem is concerned, I will show that the data can be described without presuming a kind of functional head at the level of constituent structure. In fact, the speakers use the optimal clitic in each variety. The problem has been already accounted for by Grimshaw's (1997) approach to Spanish clitics and optimality, which I will discuss in the second part of this paper. But, as we will see, Grimshaw's proposal of a unique lexical entry for the clitic se, which is valid for all constructions and all varieties of Spanish, does not match the data. I will prove that a different lexical entry will be needed. Hence, Grimshaw's approach will be reformulated. In the last part, I will try to integrate optimality into the RRG framework. Choosing the optimal clitic can be considered part of applying a constructional schema that comes into play as part of the linking algorithm.

2. Morphology and RRG

In a paper dating from 2000, Joan Bresnan considers RRG, in addition to LFG and some other theories of grammar, an instance of a "Parallel Correspondence Theory." Following Dan Everett:

As a Parallel Correspondence Theory, RRG does not build grammatical representations up serially via a recipe of derivational steps (as in, say, Chomsky's (1995) Minimalist Program), but it instead simultaneously generates separate structures, viz. the components of the Layered Structure of the Clause, the lexical-logical representation (lexical semantics), the operator projection (tense, aspect, mood, definiteness, etc.), focus structure (discourse and interactional salience and scope relations) [...] These parallel structures are connected via a set of Linking Rules - an algorithm connecting the structures, as the name implies....

(Everett 2005: 12)

Unlike LFG (cf. Bresnan ibd.), which only combines two layers of representation, namely constituent structure (c-structure) and a structure of functional features (f-structure), RRG allots four layers: Constituent Projection, Operator Projection, a conceptual-semantic structure (Logical Structure) and Information Structure (cf. Fig. 1).

How can we justify this proliferation of levels in comparison with LFG? From my point of view, the advantage lies in the consequent separation of the different functional levels of sentence structure, whereas the feature-value-pairs of LFG's f-structure comprise morphosyntactic, semantic and pragmatic information, distinguishing just one other projection, c-structure, as an arrangement of all free morphemes.

In contrast, the Constituent Projection of RRG restricts the concept, representing only constituents that are motivated by referential semantics. Hence, auxiliary verbs such as a ("has") and pu (past participle of pouvoir "can") in Fig. 1 are not linked to any node of the Constituent Structure. Bearing grammatical and not referential meaning, they are only related to the Operator Projection.

There is another point that distinguishes the representation of the Constituent and Operator projections in RRG from the well-known X-Bar-trees of Generative Grammar.

RRG trees only fix the semantically motivated layers on a vertical axis: (NUCLEUS > CORE > CLAUSE > SENTENCE). Free as well as bound morphemes of an input sequence can be linked to different nodes of a certain layer. There is no one-to-one mapping between a morpheme and a lexical or functional head. Thus, RRG trees enable the representation of amalgams and discontinuous morphemes:

For the Spanish input sequence estuve durmiendo ("I was sleeping") in Fig. 2, the verbal suffix -e is linked to the argument position of the Constituent Structure.

However, the verbal stem of the auxiliary does not form part of the Constituent but of the Operator Projection. Together, the auxiliary stem and the gerund termination -iendo of the full verb dormir ("to sleep") mark Tense and Aspect of the predicate.

According to what we have seen so far, the Operator Projection seems to be the one and only structural level for representing grammatical morphemes that do not bear referential meaning. However, this is only one side of the coin. The Operator Projection deals with only a part of the morphemes bearing grammatical meaning, i.e., morphemes that are not "relational". Non-relational morphemes modify the Nucleus, Core or Clause, but they do not indicate a relation between two constituents.

The verbal categories of Person and Number are not classes of operators. Besides Case as part of noun inflection, they constitute mainly relational categories. As far as relational categories are concerned, the semantic representation comes into play. Relational categories reflect the relation of the predicate and its arguments. Therefore, the
realization (from semantics to syntax) or interpretation (from syntax to semantics) of Person and Number (Agreement) as well as Case cannot be represented in a "static" way, but must be dealt with as part of the Linking-Algorithm (cf. Van Valin 2005b: 25).

3. "Head-marking" vs. "Dependent-marking"

In Fig. 2, we linked the verbal suffix -e (1PSg) of the auxiliary *estuve* ("was") to an argument position of the Constituent Structure. This was structurally required due to the fact that the argument slot of the full verb *dormir* ("to sleep") could not be filled by any other morpheme. RRG does not assume empty categories such as (little) "pro" in Generative Grammar. Hence, the example in Fig. 2 illustrates that a morpheme expressing Person and Number needs not to be a morpheme of Agreement, but could...
denote an argument of the predicate in the center of the semantic representation. On the other hand, (4) as well as (1a) is an example for Agreement:

(4) María estuv-o durmiendo [Agreement]
    Mary be-3sg.Past sleep-Gerund
    ‘Mary was sleeping’

Concerning the morphologic marking of arguments at the verbal stem, RRG distinguishes between two types of languages: one type that marks the relationship between the predicate and its arguments exclusively at the verb as head of the construction (head-marking languages), and languages that mark this relation at the verb-dependent constituents, e.g., via Case morphemes (dependent-marking languages).

Being a dependent-marking language, German (cf. example 5) always flags syntactic functions by means of Case morphemes or adpositions. On the other hand, head-marking Tzotzil, a Mayan language of Central America, displays argument-indexing affixes bound to the verbal stem (cf. example 6).

(5) German (dependent marking):
   a. Er brachte sie ihr
      3Nom:sg bring 3sg.past 3acc.pl 3dat.fem.sg
      ‘He brought them to her’
   b. Der Mann brachte
      det.nom.masc.sg man bring-3sg.past
      der frau blumen
      det.dat.fem.sg woman flower-acc.pl
      ‘The man brought the woman flowers’

   a. ?-∅-s-pet
      asp-3abs-3erg-carry
      ‘He/she carried him/her/it away’
   b. ?-∅-s-pet lokel yantzi ju ti ful-e
      asp-3abs-3erg-carry away woman def rabbit def
      ‘The rabbit carried the women away’

In German (5a), pronouns are free morphemes receiving intonational stress, in contrast to Tzotzil’s argument-indexing affixes (6a). In the b-examples of both languages (5b and 6b), full NPs do appear, but in Tzotzil (6b) the NPs are not marked by Case affixes or adpositions. In actuality, they appear outside the CORE, but inside the CLAUSE. Therefore, they differ from topomialized NPs appearing in the Right- or Left-Detached-Position. Nonetheless, in Tzotzil it is always the verbal affix that is mapped to the argument position, independent of the presence or absence of a corresponding full NP (cf. Van Valin 2005a: 16–18).
"Floating plurals", pro-drop and agreement – an optimality-based RRG approach

Figure 3. Head-marking in Tzotzil (cf. Van Valin 2005a: 17).

Note that Romance languages do not fit easily into one of the two types described:

(7) Italian:
   a. Gliela porta
      3.DAT.SG+3.ACC.FEM.SG brought-3.SG.PRES
      'He/she brings it to him/her'
   b. Gliela porta, la marmellata,
      3.DAT.SG+3.ACC.FEM.SG brought-3.SG.PRES, DET.FEM.SG jam
      a sua sorella
to 3SG.POSS.FEM.SG sister
      'He/she brings it to her, the jam, to his/her sister'
   c. Porta la marmellata a sua sorella
      bring-3.SG.PRES DET.FEM.SG jam to 3SG.POSS.FEM.SG sister
      'He/she brings the marmalade to his/her sister'

On one hand, the verbal suffix -a of portare ("to bring") is linked to the PSA (cf. example 1b and Spanish estuve in Fig. 2). Van Valin (2005a: 19) considers this phenomenon, called “pro-drop” in Generative Grammar, to be a “head-marking feature” that appears as an exception in a certain construction of a generally dependent-marking language. However, the fact that Romance languages allow for object clitics shows that things are even more complex.

In (7a) all syntactic functions are expressed by morphemes that do not receive stress. There is an evident parallelism between Italian (7a) and Tzotzil (5a), although clitics hold an intermediate position between free morphemes and affixes. At the first glance, (7b) too seems to be analogous to the corresponding Tzotzil example in (6b). Italian, however, unlike Tzotzil, marks the Dative twice, namely by the clitic 
   gli- on one hand and on the other hand by the preposition a, head of the PP a sua sorella that appears in the Right-Detached-Position. Hence, the Italian construction (7b) seems to bear resemblance to what is the case in so-called double-marking languages such as the North America Choctaw (cf. Van Valin ibid.). Double-marking
languages flag syntactic function by means of verbal affixes, but the full NPs, which are realized inside the CLAUSE as in Tzotzil, are additionally marked by Case.

Finally, example (7c) gives evidence for the fact that in Italian, objects can be solely marked by word order and prepositions in clear contrast to what happens in head-marking languages or double-marking languages.

In summary, Italian could be classified in line with Van Valin (2005a: 19) as basically dependent-marking, resorting to head-marking constructions whenever the subject is not realized as a full NP or an object is expressed by a clitic.

### 4. “Pronominal Redundancy” and “Cannibalistic Datives”

The Spanish examples in (2) that we considered at the beginning differ from the Italian example in (7) in the fact that the full NPs, being coreferential with the clitics, are not outside, but rather inside the (CORE). This can be proved by comparing the intonation patterns. Whereas the full NPs in (7b) are separated by a pause, there is no such a pause in (2).

Further examples are given by Belloro:

(8) a. Le duele la cabeza a Juan
   3DAT.SG ache-3SG.PRES DET.FEM.SG head to John
   'John has a headache'

b. Le preparé una tarta a mi amigo
   3DAT.SG prepare-1SG.PAST INDEF.FEM.SG cake to 1SG.POSS friend
   'I prepared my friend a cake'

c. Le gusta el cine a Juan
   3DAT.SG appeal-3SG.PRES DET.MASC.SG cinema to John
   'John likes the movies'
d. Lo vi a él
3ACC.SG see-1SG.PAST to 3SG
'I saw him' (cf. Belloro 2004: 8, examples 2–5)3

Nonetheless, the two corresponding NPs are not topicalized. They do not appear in the Right-Detached-Position (cf. Belloro ibid.: 19s). Thus, unlike in example (7b), the two NPs, not the clitics, should be considered arguments of the main predicates. The description of the examples in (8) should not differ from that of an Italian or Spanish construction containing a full NP as the PSA. All syntactic functions, subject, direct object, indirect object, etc. would be assigned to full NPs, which appear in the CORE according to the intonation pattern. Only if there is no corresponding full NP can the functions be fulfilled by affixes denoting Person and Number or clitics attached to the verb. Otherwise, clitics would have to be considered agreement markers in the same manner as Person and Number affixes.

Having established a basis for our description, let us now account for the specific data from American Spanish. The examples under (3), repeated for convenience in (9), show that American Spanish not only uses dative and accusative clitics as agreement markers for three-place constructions, it is also characterized by the striking tendency to flag a plural RECIPENT-argument and a singular THEME-Argument by means the clitic combination se las or se los:

(9) = (3)

a. Juan compró una casa para sus hijos
   John buy.3SG.PAST INDEF.FEM house for 3PL.POSS.FEM SCN.PL
   'John bought a house for his children'

b. Juan les la compró
   John 3DAT.PL 3FEM.ACC.SG buy.3SG.PAST

European Spanish:

c. Juan se la compró
   John se 3FEM.ACC.SG buy.3SG.PAST

American Spanish:

d. Juan se la/ se las compró
   John se 3SG.FEM.ACC/SE 3FEM.ACC.PL buy.3SG.PAST
   (cf. Belloro 2004: 22, examples 40d–d(i))

e. [do’(x, buy’(x, y))] PURP [BECOME have’(z, y)]

3. It is not my objective to establish a clear boundary between the cases of obligatory and facultative pronoun doubling in Spanish. For this puzzling issue, variation among different sub-norms of Spanish should be taken into account (cf. Belloro 2004). For the prescriptive norm of European Spanish, see the considerations of the Real Academia (RAE 1973: 422–424).
According to the general linking rules, in unmarked active constructions the most patient-like argument of comprar ("to buy"), the y-argument appearing at the right edge of the LS in (9e), receives the accusative case, whereas the z-argument, featuring a middle level of activity, receives the dative case.

In combination with accusative los/la(s), the dative clitic le(s) that allows Number variation is substituted by invariable se (cf. 9b). Hence, the Number of the z-argument cannot be expressed. It seems that for speakers of European Spanish this problem does not matter. However, American Spanish allows for expression of the Number of the z-Argument at the accusative clitic, although this renders this construction highly ambiguous:

\[(10)\]

\[a. \text{Juan se las compró}
\]
\[\rightarrow \text{John bought a house for his children}\]
\[b. \text{John bought a house for his children}\]
\[d. \text{John bought houses for his child}\]

If we assume, as is the case in Generative Grammar (cf. Manzini & Savoia 2004), that every clitic is the head of its own functional projection, it is difficult to explain why plurality of the dative argument could be expressed at the accusative clitic.

In a RRG framework, Valeria Belloro has proposed the following quite elegant solution:

I will [...] 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. (Belloro 2004: 43).

4. In this paper, I will make use of a simplified LS for verbs of buying (cf. Van Valin 2005a: 157). The PURP phrase codes the intention of the buyer to get himself or another person in possession of the bought object (cf. Van Valin & Lapolla 1997: 382–386). Hence it introduces a third argument, traditionally labeled BENEFACTIVE or BENEFICIARY, that in Spanish is realized as a dative CORE argument or as an Argument-Adjunct marked by the preposition para. Belloro (2004: 47) uses a slightly modified LS. Because the semantics of verbs of buying is not a topic of this paper, I will not discuss the nuances of the different LS.

5. The RRG linking algorithm processes the LS, assigning the Actor-Macrorole to the most active argument of a transitive i.e., in an active construction, the Actor of a transitive verb receives nominative and the Undergoer accusative. A third argument, that by definition does not receive a Macrorole, is realized as a dative or a PP. In RRG, traditional semantic roles such as BENEFACTIVE or THEME are just labels that correspond to different argument positions in a given LS (cf. Van Valin 2005a: 53–67; Kailuweit 2004).

6. Hence, it is not correct when Company (1998: 544) declares in a lump-sum way that we are dealing with a case of "reanalysis of the morpheme -s which in this area of grammar adds the value of animacy-humanness typical of datives to its plurality."
Belloro's approach accounts for the fact that Spanish clitics can be realized as argument functions in the absence of corresponding full NPs. Therefore, they have to be linked to the respective nodes of the Constituent Projection (cf. Fig. 5).

![Figure 5. AGX-node (cf. Belloro 2004: 47).](image)

Whether the AGX-node represents arguments or not must be verified during the linking process. For Semantics-to-Syntax-Linking, the algorithm will use discourse-pragmatic information to be able to tell whether a certain full NP should be realized inside or outside the CORE, and therefore stands for an argument or not. As a result, the algorithm will retrieve a corresponding syntactic template, a Constituent Structure having one, two or even three argument slots. In Fig 6. el regalo ("the present") is topicalized and stays outside the CORE, whereas María is focal and realized inside the CORE. Hence, the BENEFACTIVE-argument of the LS is represented by the full NP María, but the THEME-argument is not.

The AGX-node stands for all arguments that, according to Information Structure, are not represented by full NPs. It contains the Person and Number features of all arguments of the LS, whereas Case features are ascribed during the linking process as a result of Macrorole assignment according to the choice of a marked or unmarked diathesis (active or passive construction) (cf. Fig. 7).

In the end, a morphophonological rule determines the realization of the AGX-features as verbal clitics (cf. Belloro 2004: 48).

Belloro's considerations about Syntax-to-Semantics-Linking are less detailed. Informally, we could assume the following: At the beginning of the Syntax-to-Semantics-Linking Algorithm, the parser marks the CORE as a unit of information structure. If the CORE contains full NPs, these NPs will be linked to argument positions of the LS. Argument positions that are still empty after this step will be filled by clitics or Person and Number affixes of the verb.
The morphology-component that Belloro puts forward is based on Stump's (2001) Inferential-Realizational Approach. According to Stump (2001), inflectional features are not lexical units (morphemes) that, having a determined form and a corresponding grammatical content, could be attached to a lexical root. It is rather the case that a lexical entry receives inflection features by morphophonological rules that combine inflected forms with the root. Hence, the approach is inferential and not lexical (Stump 2001: 1). Furthermore, Stump (ibid.: 2) assumes that it is the association of a root with certain morphosyntactic properties at the level of content that licenses a certain morphophonological realization at the level of expression. Thus, the morphosyntactic properties are neither the result of attaching morphemes nor of applying morphophonological rules. In fact, they precede the morphophonological realization that makes it possible in the first place. Therefore, Stump's approach is realizational and not incremental.
In Fig. 7 we have already shed light on how Belloro makes use of Stump’s approach. The argument slots of the predicate are associated with the agreement features. Now the anti-lexical component of Stump’s approach comes into play, helping to describe the phenomenon of “cannibalistic datives” in the clitic chain *se los/se las*.

The different morphosyntactic properties that stem from the arguments and are associated with the predicate can be bundled and condensed in the form of realization rules that operate over the entirety of morphological material and are not bound to individual morphemes. Thus, the plurality of the Non-Macrorole-Argument can be realized as one feature of the clitic cluster *se los/se las*. Los is not a clitic having the features \{-1, -2, +ACC, +PL, +M\} as a traditional morphologic analysis would assume, but *se los* must be considered as a unit, realizing the whole of the morphological features associated with the arguments: \{-1, -2, +DAT, +ACC, +PL, +M\}.

Belloro’s account leaves the question open as to why the plurality of the Non-Macrorole-Argument is realized and not the singularity of the Undergoer-Argument. Without being more explicit as far as the realization rules are concerned, Belloro (2004: 56) refers to Company (1998), who claims that the Agreement features of the Non-Macrorole-Argument are always stronger than those of the Undergoer-Argument. According to Company (1998), this would also be proved by fact that clitic redundancy is always obligatory with the Non-Macrorole-Argument, whereas the Undergoer-Argument requires a redundant clitic only in a certain number of cases that are well determined by Information Structure rules.

Nonetheless, neither Company (1998) nor Belloro (2004) account for the fact that the Undergoer-Argument imposes its plurality when the Non-Macrorole-Argument is singular (cf. example (11)):

(11) a. Juan compró los juguetes para su hijo
    John buy-3SG.PAST DET.MASC.PL toy-PL for 3POS,3SG son
    ‘John bought the toys for his son’

7. A similarly interesting phenomenon that could be compared to the “floating” datives in Latin American Spanish is that of certain genitive constructions in modern English. For example, in some cases involving group genitives, such as “the mother of John’s friend”, the possessive form (3) could be considered as a type of “floating genitive”. In this example, the possessor and possessem implied by the possessive (3) is ambiguous (it could be understood as either Case 1: “[the mother of John]’s friend”, or Case 2: “the mother of [John]’s friend”). Case 1 is a group genitive, i.e., the possessive (3) does not directly follow its possessor noun but rather comes at the end of the entire NP, and therefore could be seen as “floating”. For a more detailed explanation of the possessive (3) in genitive constructions of English, please refer to Allen (2003) and Rosenbach (2002).

8. Belloro (2004: 53) argues that Object Agreement of the Undergoer-Argument only occurs when it is “accessible” or “inactive” in the sense of Chafe (1987), having a middle degree of activity.

By virtue of these data, the concept of a cannibalistic dative in Spanish has to be more than relativized. It is not the dative that imposes its number on the accusative. It is rather the fact that the feature plurality must always be realized, regardless of which of the two object arguments from which it stems. Hence, it seems to me more appropriate to call the phenomenon under examination a “floating plural”. I will come back to this problem at the end of section 5 of this paper.

5. Linking, agreement and “Floating plurals”

In this section, I will claim that in RRG the phenomena under examination can be described in a somewhat more “conservative” way. In addition, my approach will be more precise than Belloro’s. I will account for some data for which Belloro does not give an adequate description.

I will not deny that Belloro’s approach is very interesting from a theoretical point of view and almost convincing as far as the descriptive results are concerned. Nonetheless, in my opinion, it unnecessarily clashes with some axioms of RRG in the field of morphology. The main problem is the postulated AGX-node at the level of Constituent Structure. This node, bearing some resemblance to a functional head in Generative Grammar, only represents Constituents when the argument positions of the predicate are not filled by full NPs. However, this case is well known in RRG and perfectly described for head-marking languages without any need for an AGX-node. In head-marking languages, argument-indexing verbal affixes still represent the arguments of the predicate (cf. Fig. 3). Languages showing the “pro-drop-parameter” in generative terms can be considered as basically dependent-marking languages with subject-verb-agreement. If and only if there is no full NP or stressed pronoun realizing the subject function, the verbal affix denoting Person and Number fulfills the subject function (cf. Van Valin 2005a: 19). Thus, for all sentences that represent the arguments of the predicate as free morphemes, an AGX-node is dispensable. In (12), the relation between the subject Juan (“John”) and the verbal affix is a relation of

10. Grimshaw’s (1997: 188) concept of “floating number” is too fuzzy, although her analysis that we will discuss in section 5 deals with the fact that it is only Plural that “floats”.
agreement. As an instance of relational morphology, it should be dealt with by applying the linking rules.

(12) Juan via a Maria
    John see-3sg.past a Mary
    'John saw Mary'

An argument put forward by Belloro (2004) to justify her assumption of the AGX-node is precisely the problem of "floating plurals" in American Spanish. As we have already seen, for Belloro (2004), in the clitic cluster se los/le las, the dative and the accusative clitics are not independent morphemes, but form a unit realizing the morphosyntactic properties of the Non-Macrorole-Argument and the Undergoer-Argument simultaneously.

Nonetheless, the Actor-Argument is obviously not affected by this process of amalgamation. Neither of the object arguments has any impact on the verbal suffix denoting Person and Number, nor does the latter influence the preverbal clitic cluster. Thus, it is not comprehensible why in Belloro’s representation of example (13) (cf. Fig. 8) the morphosyntactic information given by the Actor-Argument is related to the clitic cluster se lo and stored to the AGX-node.

(13) Juan se lo compró a María, el regalo
    John se 3acc:sg buy:3sg.past for Mary det:3sg present
    'John bought it for Mary, the present'

Figure 8. Association of morphosyntactic information with AGX through se lo (cf. Belloro 2004: 49).
In summary, in the Romance varieties under examination, subject agreement must be sharply separated from object agreement. Subject agreement is obligatorily realized via bound suffixes, whereas object agreement requires preverbal clitics. There is no evidence for a combination or interference of the two phenomena. Hence, a unified description of agreement by means of an AGX-node is problematic in my view.

But how can we otherwise account for the fact that the plurality of the Non-Macrorole-Argument is realized by the -s of the clitic chain se los/se las? My proposal consists of assuming a discontinuous constituent. In section 2, we dealt with discontinuous morphemes at the level of Operator Projection. Figure 2 illustrated the case that in Spanish, the category “Progressive Aspect” is realized by the auxiliary estar and gerundial suffix -iendo. At the level of Constituent Projection, Van Valin (2005a: 178) proposes a discontinuous analysis for the partitive construction in Italian:

![Partitive in Italian](image)

Figure 9. Partitive in Italian (cf. Van Valin 2005a: 178).

In *Maria ne ha comprato due* ("Mary has bought two of them"), one part of the Undergoer-Argument is focal (*due"two"), but the other part is a clitic and therefore necessarily topical (*ne"partitive*). Thus, the argument has to be split up into two segments. It becomes a discontinuous constituent. Note that only *ne* is connected with Constituent Structure, whereas *due* (two) is classified as a nominal operator attached to the Operator Projection of the constituent *ne* (cf. Van Valin 2005a: 178).
However, classifying a clitic as a head of a NP poses a problem. As far as Operator Projection is concerned, unlike full NPs, \textit{ne} only allows the category of Quantification. Other nominal operators, such as Nominal Aspect, Number, Negation, Definiteness or Deixis, are excluded. To resolve this problem, one could assume a category \textit{pro} at the level Constituent Projection that displays a deficient Operator Projection, permitting only a part of the operators of a standard full NP.

If we transfer these considerations to the problem of “floating” plurals, the clitic chain \textit{se los} could be considered a discontinuous materialization of two arguments:

One major advantage of this view is that it still works when the clitics function as agreement markers. When an argument is represented by a full NP, the corresponding clitic is not connected to the Constituent Projection being part of the relational verbal morphology generated in the linking process. However, it can still have an Operator Projection. Hence, the lower part of Fig. 10 would be the same, independent of an argument reading or agreement marker reading of one or both clitics.

As mentioned previously, the number of the Non-Macrorole-Argument will only be materialized if it is plural. If the Non-Macrorole-Argument is singular, but the Undergoer-Argument plural, the clitic cluster has to be \textit{se los/se las} and not \textit{se lo/la}.

\[ (14) = (11) \]

\begin{enumerate}
  \item a. Juan compró \textit{los} juguetes para su hijo
      John buy-3SG.PAST DET,MASC.PL toy-PL for 3POSS.SC son
      \textit{John bought the toys for his son}
\end{enumerate}

\[ ^{11} \text{A list of NP Operators can be found in Van Valin 2005a: 24.} \]
Morphophonological rules must take these data into account and correctly predict the differences between European and American Spanish. These rules will be based on Grimshaw's proposal discussed in the following lines.

In an Optimality framework, Grimshaw (1997) proposes a formalization that, at first glance, deals with the distribution of se los/la: se los/la in European and American Spanish. Nonetheless, in the following I will show that her approach is both deficient at the level of description and problematic at the level of explanation.

Grimshaw (1997) assumes an identical inventory of morphemes for both varieties. In addition, she posits a unique feature value structure for the clitic se in all contexts: se is supposed to be unmarked for all features, i.e., for Reflexive (R), Person (P), Number (N), Gender (G) and Case (c). In contrast, lo is marked as non-reflexive, masculine and accusative, los as non-reflexive, plural, masculine and accusative.2

\[
\begin{align*}
\text{se} & = (R), (P), (N), (G), (C) \\
\text{lo} & = -R, (P), (N), \text{MASC, ACC} \\
\text{los} & = -R, (P), \text{PL, MASC, ACC}
\end{align*}
\]

We can see immediately that this description is inadequate for various reasons. First, it does not do justice to the data. As Belloro (2004:32) also points out, “floating” plurals only appear with se in a non-reflexive construction. If se has to be interpreted as a reflexive, then plurality of the Non-Maro role-Argument cannot be expressed:

\[
\begin{align*}
\text{a. Ellos se comprarון un libro} \\
\text{3|nom.masc.pl refl buy-3pl.past indef.masc.sg book} \\
\text{′They bought themselves a book′}
\end{align*}
\]

\[
\begin{align*}
\text{b. Ellos se lo comprarón} \\
\text{3|nom.masc.pl refl 3acc.sg buy-3pl.past}
\end{align*}
\]

\[
\begin{align*}
\text{c. *Ellos se los comprarón} \\
\text{3|nom.masc.pl refl 3acc.pl buy-3pl.past}
\end{align*}
\]

(cf. Belloro ibid., example (48))

Following Grimshaw (1997), (16c) should be correct. Assuming a unique meaning for se in all contexts, there is no way of describing the differences in syntactic behaviour resulting from its reflexive or non-reflexive use. The second point concerns the feature

\[12.\text{For convenience, I will only consider se los.} \]
structure of the accusative clitics. In my opinion, it is inadequate to ascribe plurality to los in the varieties of American Spanish. Using los in the context of “floating” plural, speakers of American Spanish do not choose a “wrong” clitic that is plural and accusative because of the impact of the plurality of the Non-Macrorole-Argument. Rather, they choose the “right” clitic se los, expressing simultaneously the features dative, accusative and plurality of one of the two arguments. In other words, se los is not a combination of two clitics, but one complex clitic expressing properties of two arguments. Hence, the feature-value-structure of se los is not identical in American and European Spanish, as Grimshaw (1997) assumes.

In line with the mainstream Optimality Theory, Grimshaw (1997) proposes different rankings of constraints that are responsible for the variation. In contrast, I will account for the facts, assuming identical constraints but different feature-value-structures of the morphological material. The relevant features are illustrated in (17).

In European Spanish, se los displays the feature-value-structure [+DAT, +ACC +PL]. However, the identical sequence in American Spanish shows the feature-value-structure [+DAT, +ACC, +PL].

(17) European Spanish: se los = [+DAT, +ACC, +PL]
    American Spanish: se los = [+DAT, +ACC, +PL]

Acting on this assumption, an identical chain of constraints can explain the different use of se los/se los in the two varieties. Following Grimshaw (1997), I will make use of the distinction of Fill-Rules and Parse-Rules.

(18) Fill: Only features in the input can appear in the output Parse:
    All features in the input must appear in the output (cf. Grimshaw 1997: 170)

From the semantic point of view, there are four different cases of combining a singular or plural Non-Macrorole- and Undergoer-Argument:

(19) a. NMR-ARG = [+DAT +PL]; U-ARG = [+ACC +SG]
    b. NMR-ARG = [+DAT +SG]; U-ARG = [+ACC +PL]
    c. NMR-ARG = [+DAT +SG]; U-ARG = [+ACC +SG]
    d. NMR-ARG = [+DAT +PL]; U-ARG = [+ACC +PL]

Ranking three constraints in the following way: Fill Plural > Fill Accusative Plural > Parse Plural, the use that each variety makes of se los can be correctly predicted.

The Fill-Plural-Rule prohibits the realization of a plural that is not given in the input. In (19c) both arguments are singular. Hence, for an input (19c) the output se los infringes upon the rule in both varieties.

---
13. The assumption of Company (1998: 547): “se los seems to be lexicalized, a single pronoun, solos, totally unanalyzable for most speakers” seems to be too radical. In fact, se los can be considered one complex clitic. But it refers to two entities and can be analyzed as having the feature-value-structure [+DAT, +ACC, +PL].
<table>
<thead>
<tr>
<th>American Spanish</th>
<th>Output</th>
<th>Fill PL.</th>
<th>Fill ACC. PL.</th>
<th>Parse PL.</th>
</tr>
</thead>
<tbody>
<tr>
<td>se los = +DAT, +ACC, +PL</td>
<td>*c!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>se lo = +DAT, +ACC</td>
<td></td>
<td></td>
<td>*a!, *b!, *c!</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>European Spanish</th>
<th>Output</th>
<th>Fill PL.</th>
<th>Fill ACC. PL.</th>
<th>Parse PL.</th>
</tr>
</thead>
<tbody>
<tr>
<td>se los = +DAT, +ACC +PL</td>
<td>*c!</td>
<td></td>
<td>*a!, *c!</td>
<td></td>
</tr>
<tr>
<td>se lo = +DAT, +ACC</td>
<td></td>
<td></td>
<td>*a!, *b!, *c!</td>
<td></td>
</tr>
</tbody>
</table>

The **Fill-Accusative-Plural Rule** prohibits the realization of an accusative plural that is not given in the input. In American Spanish, neither se los nor se lo realize an accusative plural. Se los only realizes the plural of either of the two arguments. In contrast, in European Spanish se los realizes an accusative plural. For the input of (19a) and (19c), the Undergoer-Argument is singular. Hence, the rule is infringed upon if se los appears in the output. Finally, the **Parse-Plural-Rule** requires the realization of a plural if it is given in the input. In (19a), (19b) and (19c) one or both of the arguments are plural. Therefore, the constraint is violated in both varieties if the output sequence is se lo.

In American Spanish, it is the **Fill-Plural-Rule** that accounts for the choice of se lo in (19c). The **Parse-Plural-Rule** is decisive for the choice of se los in (19a), (19b) and (19d). In European Spanish, the same Fill-Plural-Rule causes se lo to be chosen in (19c), whereas the choice of se lo in (19a) results from the Fill-Accusative-Plural-Rule. Finally, the **Parse-Plural-Rule** guarantees that the optimal clitic is se los in (19b) and (19d).

The last question to be raised is how these considerations could be integrated into a RRG framework. Belloro (2004) does not explicitly say at which point morphophonological rules come into play. RRG's Linking Algorithms are only dealt with in a rudimentary way, as far as the filling of the AGX-node is concerned. In my approach, finding the optimal clitic as a morphophonological process is part of a so-called **Constructional Schema**. Constructional Schemas spell out the language-specific particularities of the general Linking-Rules (cf. Van Valin 2005a: 131–135):

However, the interaction of general Linking-Rules and Constructional Schemas is only sketched in Van Valin (2005a). As a handbook, Van Valin (2005a) lacks explicit algorithms that could shed light on the integration of Constructional Schemas in specific languages. In following, I will propose that Constructional Schemas concerning the coding of the PSA by a verbal affix as well as the choice of optimal preverbal object clitics come into play at point 3 of the general Linking-Algorithm.

---

14. See the discussion of the role of constructional schemas in linking, using Sama examples, in Van Valin 2005a: Chapter V.
"Floating plurals", prodrop and agreement – an optimality-based RRG approach

Figure 12. Architecture of Role and Reference Grammar (cf. Van Valin 2005a: 134).


1. Construct the semantic representation of the sentence, based on the LS of the predicator.
2. Determine the actor and undergoer assignments, following the Actor-Undergoer Hierarchy.
3. Determine the morphosyntactic coding of the arguments
   a. Select the PSA [...]
   b. Assign the XPs the appropriate case markers and/or adpositions.
   c. Assign the agreement marking to the main or auxiliary verb, as appropriate.

In order to determine the morphosyntactic coding of the argument, the LS of the predicator must be checked. If the LS does not contain a full NP that fulfills the PSA function, a “pro-drop”-Constructional Schema will be retrieved. In (21) such a schema will be sketched:

(21) CONSTRUCTION: “pro-drop”

SEMANTICS:
The PSA-Argument is not represented by a full NP or this full NP has to be realized in a peripheral position because discourse pragmatics interfere and a topicalization constructional schema is retrieved.

MORPHOLOGY:
The PSA-Argument is realized as a verbal suffix of Person and Number.

SYNTAX:
The verbal suffix of Person and Number fills an argument position of the CORE.
PRAGMATICS:
The PSA-Argument is topical.

If the Undergoer-Argument is not selected as PSA and if there is a Non-Macrorole-Argument in the LS of the predicate, a Constructional Schema for object clitics has to be retrieved. (22) contains an outline of a Constructional Schema for Spanish object clitics.

(22) CONSTRUCTION: Spanish object clitics

SEMAN TICS:
Clitics are agreement markers if they are co-referent with a full NP that, due to discourse pragmatics, has to be realized inside the CORE. If there is no such full NP, or if this NP has to be realized in a peripheral position because of discourse pragmatics considerations (retrieving a topicalization constructional schema), the clitics are arguments. They stand for the Undergoer and/or the Non-Macrorole-Argument.

SYNTAX:
If clitics stand for arguments, they fill one or two argument slots of the CORE.

MORPHOLOGY:
If a dative Non-Macrorole-Argument is realized as a full NP, a dative clitic is attached to the verb. For an Undergoer-Argument realized as a full NP, this is only the case under special conditions depending on discourse pragmatics. The optimal clitic is chosen from a particular inventory according to the variety of Spanish (European or American). If two arguments require a clitic, a complex clitic is chosen.

PRAGMATICS:
The Undergoer-Argument only requires a clitic as agreement marker if it is "accessible" or "inactive" (cf. Belloro 2004: 51–54).

If the Constructional Schemas are processed, the Linking-Algorithm can be fulfilled:

1. Select the syntactic template(s) for the sentence [...].
2. Assign XPs to positions in the syntactic representation of the sentence [...].

It would go beyond the scope of this paper to deal with Syntax-to-Semantics-Linking in almost the same manner. As far as the problem of “floating” plural is concerned, from a Syntax-to-Semantics point of view se los is always ambiguous, if the clitics stand for arguments. The Parser (cf. Fig. 12) would have to generate three different Operator Projections, connecting the plural marker -s either with se (cf. Fig. 10) or with lo or with both. For disambiguation we would need a Linking Theory at the text level. At the moment, RRG is still lacking a detailed text-linguistic component (cf. Butler 2003: 42).15

15. Some considerations concerning the relation of text linguistics and information structure can be found in Van Valin (2005a: 170–174).
6. Conclusion

In this paper I have sketched a RRG approach for the description of some of the most-researched characteristics of Romance languages in the field of morphology: pro-drop and object clitics. I wanted to show that RRG provides two major descriptive advantages. On one hand, there is a sharp distinction between a level of referential meaning (Constituent Projection) and a level of grammatical meaning (Operator Projection). On the other hand, as far as morphology is concerned, non-relational morphology (operators) is separated from relational morphology. The latter is accounted for during the Linking-Process by means of Constructional Schemas. I will not deny that a more detailed description of morphology in the RRG framework has yet to be worked out.\(^{16}\)

With regard to clitics, the realization of the plurality of the Non-Macrorole-Argument by means of the clitic cluster se los/se las in American Spanish is a puzzling problem for all theories of morphology. Together with Company (1998) and Belloro (2004), I have claimed that the phenomenon can only be described adequately if we assume a complex clitic se los/se las as realizing two arguments or being the agreement marker for both of them. Hence, a modification of the Optimality approach put forward by Grimshaw (1997) was required. The modified Optimality approach presented in this paper is able to account for a different distribution for se lo and se los in European and American Spanish and also takes into consideration that “floating” does not occur in reflexive constructions and that it affects only the plural and not the singular of the Non-Macrorole-Argument.

In contrast to Belloro (2004), I argued for a more “conservative” RRG account, avoiding the postulation of some kind of a functional head at the level of Constituent Projection. It has been shown that instead of using an AGX-node, the problem can be dealt with by assuming a discontinuous morpheme at the level of Operator Projection and a Constructional Schema that is retrieved during the Linking-Process. The double function of clitics as arguments and agreement markers can be described if we consider Romance languages as basically dependent-marking languages, allowing head-marking constructions under certain conditions that the present paper has sufficiently specified.

References


\(^{16}\) Everett (in progress) will probably fill this gap.


