In Canzhong Wu, Christian Matthiessen, and Maria Herke (eds.)
Proceedings of ISFC 2008: Voices Around the World. ISFC Sydney, July 2008: 7-12.
ISBN 978-0-9805447-1-8.

A Systemic Functional Micro-Grammar of Spanish Clitics

Víctor M. Castel

Consejo Nacional de Investigaciones Científicas y Técnicas
Universidad Nacional de Cuyo
Mendoza, Argentina
vcastel@lab.cricyt.edu.ar

Abstract

The word order patterns and participant role distribution of Spanish clitics are two well-known phenomena which have been thoroughly studied in Hispanic Linguistics from the perspective of both formal and functional grammars. Arús (2006) is a Sydney Grammar approach to the semantics of *se* but with no references to explicit realization rules or integration with the morpho-syntactic and semantic distributional properties of other clitics. Thus, there are currently no descriptions integrating these phenomena from the viewpoint of Systemic Functional Linguistics, let alone the Cardiff Grammar framework and the specifics of River Plate Spanish (RPS). A set of data illustrating the two phenomena is here accounted for adequately and elegantly in terms of the Cardiff Grammar framework (Fawcett 2000, 2008), more specifically within the new computational version of GeneSys, the Cardiff Grammar Generator of Fawcett & Castel (2006). The paper presents a micro-grammar capable of generating representations that capture both the patterns constraining word order and the expression of participant roles of RPS clitics.

1 Introduction

The purpose of this paper is to specify a micro-grammar of River Plate Spanish clitics (GSC) within GeneSys, the Cardiff Grammar (CG) environment for the development of generation oriented grammars (Fawcett *et al* 1993, Fawcett & Castel 2006). CG is a theoretical framework based on Halliday's leading ideas on the systemic-functional approach to natural languages (Fawcett 2000, 2008).

Since the CG framework has not yet been applied to Spanish (Fawcett 2008: 3), let alone River Plate Spanish (RPS), this paper addresses the problem posed by the word order patterns and participant role distribution of RPS clitics from the CG perspective.[1]

2 Word Order Patterns and Expression of Participant Roles

The word order constraints underlying the distribution of RPS clitics can be abbreviated by the schemas (1)-(3):

(1) III II I A/D, (Modified version of Perlmutter (1972)'s surface structure filter.)
(2) III II I A/D + V [non-imperative finite form],
(3) V [imperative or infinitive or gerund] + III II I A/D,

where 'III II I A/D' is a growing monotonic sequence in which roman numbers stand for clitic person values, 'A/D' is either an accusative (A) or dative (D) clitic, 'V' stands for a verb, the expressions within square brackets refer to forms of such a verb, and the sign '+' marks the relative order between the clitic sequence and 'V'. The schemas capture in an abbreviated way two aspects of word order: (i) patterns governing the occurrence of

[1] Arús (2006) is a Sydney Grammar approach to the semantics of *se* but with no references to explicit realization rules or integration with the semantic and morpho-syntactic properties of other clitics.

clitics with other clitics (cf. (1)), and (ii) patterns governing the occurrence of clitics with their governing verbs (cf. (2)-(3)).

According to (1), the sequences in Table 1 are all the possible well-formed clitic sequences that can occur in RPS clauses, i.e. any other clitic sequence is ill-formed:

One clitic	**Two clitics**		**Three clitics**		**Four clitics**
se	se te	te le(s)	se te me	se me lo/a(s)	se te me le(s)
te	se me	te lo/a(s)	se te nos	se nos lo/a(s)	se te nos le(s)
me	se nos	me le(s)	se te le(s)	te me le(s)	se te me lo/a(s)
nos	se le(s)	nos le(s)	se te lo/a(s)	te nos le(s)	se te nos lo/a(s)
lo/a(s)	se lo/a(s)	me lo/a(s)	se me le(s)	te nos lo/a(s)	
le(s)	te me	nos lo/a(s)	se nos le(s)		
	te nos		te me lo/a(s)		

Table 1: Word Order Patterns among Clitics

Table 2 presents a few examples illustrating the well-formedness conditions captured by the schemas (2) and (3) on the occurrence of clitics in relation to their governing verbs:

Clauses satisfying (2)	**Clauses violating (2)**	**Clauses satisfying (3)**	**Clauses violating (3)**
Me saludás.	*Saludásme.	Saludame.	*Me saludá.
Lo saludás.	*Saludáslo.	Saludalo.	*Lo saludá.
Me lo regalás.	*Regalásmelo.	Regalámelo.	*Me lo regalá.
Te las quieren regalar.	*Quieren te las regalar.	Quieren regalártelas.	*Quieren te las regalar.
Se estaban saludando.	*Estaban se saludando.	Estaban saludándose.	*Estaban se saludando.

Table 2: Word Order Patterns between Clitics and their Governing Verbs

The following table abbreviates a few core correlations between participant roles (**PR**), and the clitics and verb endings that express them:

Me lo regalo/ás/a. $me_{i,\ Af\text{-}Ca}\ lo_{n,\ Af\text{-}Po}\ regalo_i/ás_k/a_{k,\ Ag}$	Me le regalo/ás/a. $me_{i,\ Af\text{-}Po}\ le_{j,\ Af\text{-}Ca}\ regalo_i/ás_k/a_{k,\ Ag}$	Me regalo/ás/a. $me_{i,\ Af\text{-}Ca}\ regalo_i/ás_k/a_{k,\ Ag}$ $me_{i,\ Af\text{-}Po}\ regalo_i/ás_k/a_{k,\ Ag}$
Te lo regalo/ás/a. $te_{i,\ Af\text{-}Ca}\ lo_{n,\ Af\text{-}Po}\ regalo_k/ás_i/a_{k,\ Ag}$	Te le regalo/ás/a. $te_{i,\ Af\text{-}Po}\ le_{n,\ Af\text{-}Ca}\ regalo_k/ás_i/a_{k,\ Ag}$	Te regalo/ás/a. $te_{i,\ Af\text{-}Ca}\ regalo_k/ás_i/a_{k,\ Ag}$ $te_{i,\ Af\text{-}Po}\ regalo_k/ás_i/a_{k,\ Ag}$
Se lo regalo/ás/a. Spurious *se*: $se_{i,\ Af\text{-}Ca}\ lo_{n,\ Af\text{-}Po}\ regalo/ás/a_{k,\ Ag}$ Reflexive *se*: $se_{i,\ Af\text{-}Ca}\ lo_{n,\ Af\text{-}Po}\ regala_{i,\ Ag}$ Impersonal *se*: $se_{i,\ Ag}\ lo_{n,\ Af\text{-}Po}\ regala_{i,\ Ag}$	*Se le regalo/ás. // Se le regala. Reflexive *se*: $se_{i,\ Af\text{-}Po}\ le_{n,\ Af\text{-}Ca}\ regala_{i,\ Ag}$ Impersonal *se*: $se_{i,\ Ag}\ le_{n,\ Af\text{-}Ca}\ regala_{i,\ Ag}$	*Se regalo/ás. // Se regala. Reflexive *se*: $se_{i,\ Af\text{-}Po}\ regala_{i,\ Ag}$ $se_{i,\ Af\text{-}Ca}\ regala_{i,\ Ag}$ Impersonal *se*: $se_{i,\ Ag}\ regala_{i,\ Ag}$

Table 3: Correlations between PRs, and Clitics and Verb Endings

Note that no examples are given of clauses containing three or four clitic sequences, for no provision is made here for *ethical* participants.

3 System Network and Realization Rules

The system network and the realization rules (RRs) needed to account for the data in §2 are given in Figures 1-7, and Figure 8, respectively. The RRs assume that the syntactic unit clause (**Cl**) filling the text-sentence element '**Σ'** is made up of a certain number of *places*. When a 'Cl' is introduced into the linguistic representation being constructed, a structure like (4) is defined:

(4) $_{\Sigma}$[[selection_expression] $_{Cl}$[1, …, 120]$_{Cl}$]$_{\Sigma}$

where 'selection_expression' is a variable ranging over semantic features selected from the system network (cf. Figures 1-7), and '…' ranges over 2 through 119.[2] RRs are defined for the elements Operator (**O**), Verb Root (**M**), and Verb Ending (**Vnd**) to occupy places 99, 100 and 101, respectively:

(5)

$_{Cl}$[…	95	96	97	98	99	100	101	102	103	104	105	…]$_{Cl}$
	Clt3	Clt2	Clt1	CltA CltD	O	M	Vnd	Clt3	Clt2	Clt1	CltA CltD	

The places 95-98 are reserved for the occurrence of clitics with non-imperative finite verb forms (cf. schema (2)), and 102-105 for the occurrence of clitics with imperatives (= *proposal_for_action*), infinitives and gerunds (cf. schema (3)). See lines L2-L3 of RRs 20.11-12, 20 and 20.2-3 of Figure 8, responsible for the exponence of 'CltA', 'CltD', 'Clt3', 'Clt1', and 'Clt2' as *lo/a(s)*, *le(s)*, *se*, *me*/*nos,* and *te*, respectively. Note that only the clitic sequences in Table 1 are allowed.

GSC generates text-sentences like (6), (8) and (10), after deletion of unused places and stripping of the corresponding linguistic representations (7), (9), and (11), respectively:

(6) Lo saludás. (Cf. *Saludáslo. See second row in Table 2.)

(7) $_{\Sigma}$[[…, giver, …] $_{Cl}$[1 … 97 $_{CltA/Af}$[lo] 99 $_{M}$[salud] $_{Vnd/Ag}$[ás] … 119 $_{E}$[.]]$_{Cl}$]$_{\Sigma}$

(8) Saludálo. (Cf. *Lo saludá. See second row in Table 2.)

(9) $_{\Sigma}$[[…, proposal_for_action, …] $_{Cl}$[1 … 99 $_{M}$[salud] $_{Vnd/Ag}$[á] 102 103 104 $_{CltA/Af}$[lo] … 119 $_{E}$[.]]$_{Cl}$]$_{\Sigma}$

(10) Se lo saluda. Impersonal *se*. (Cf. *Lo se saluda.)

(11) $_{\Sigma}$[[…, agent_and_affected, non_coref_rel_ag_af, agent_fcs, outsider_fcs, low_deixis_fcs, outsider_low_deixis, outsider_fcs_recoverable, …, singular_fcs, …, agent_stated, affected_stated, …, outsider_af_recoverable, …, singular_af, high_deixis_af, least_active, outsider_high_deixis_least, non_feminine_af] $_{Cl}$[1 … 94 $_{Clt3/Ag}$[se] 96 97 $_{CltA/Af}$[lo] 99 $_{M}$[salud] $_{Vnd/Ag}$[a] … 119 $_{E}$[.]]$_{Cl}$]$_{\Sigma}$

RR 20.11 of Figure 8 is triggered by the feature *outsider_high_deixis_least*. This feature corresponds to the meanings 'other', 'high deixis', and 'least' that García (1975) assigns to the clitics *lo/a(s)*. The system of Figure 5 introduces it as a gate on the basis of the disjunctive features *high_deixis_af* and *high_deixis_af_po* subcategorizing outsiders and a subset of addressees in the semantic specification of the PRs Affected (Af) and Affected-Possessed (Af-Po), respectively. These PRs 'Af' and 'Af-Po' conflate with the element 'CltA' depending on whether the clause selection expression contains the feature *high_deixis_af* or *high_deixis_af_po*, respectively. Cf. L4-L5 of RR 20.11. The element 'CltA' is expounded by the lexical item *lo/a(s)*, independently of the

[2] It is assumed here that GSC requires 120 places for the clause.

position it occupies with respect to the governing verb, and also independently of whether it has been conflated with 'Af' or 'Af-Po'. Cf. L6-L17 of RR 20.11. Whether the exponence of 'CltA' is *lo*, *la*, *los*, o *las* depends on the selection of specific features from the systems PERSON, NUMBER, GENDER, and TENOR. 'CltA' is expounded by *lo* (cf. L6-L8) if the clause selection expression contains either (i) a singular outsider (*singular_af* or *singular_af_po*), non-feminine (*non_feminine_af* or *non_feminine_af_po*), or (ii) a singular addressee (*singular_addressee_af* or *singular_addressee_af_po*), non-feminine (*non_feminine_af* or *non_feminine_af_po*), and the tenor is formal. *Mutatis mutandis*, 'CltA' is expounded by *la* (Cf. L9-L11), *los* (Cf. L12-L14), and *las* (Cf. L15-L17).

The occurrence of *se* in (10) results from the application of RR 20, which is called by the feature *outsider_low_deixis* introduced by the rule of Figure 7. The features *low_deixis_fcs, low_deixis_af*, *low_deixis_af_po*, *low_deixis_af_ca* subcategorize outsiders and a subset of addressees in the semantic specification of the PRs 'Ag', 'Af', 'Af-Po', and 'Af-Ca', respectively. The rule in Figure 7 reduces the disjunction of features to the gate *outsider_low_deixis* and thus it captures García (1975)'s proposal that the meaning of *se* is made up of 'other' and 'low deixis', fulfilling in (10) a defocussing function.

4 Conclusions

The paper has presented the system network and realization rules of a generation oriented micro-grammar of RPS clitics. The specification is based on Castel (2007)'s rule typology of the Cardiff Grammar framework (Fawcett 2008, 2000; Fawcett *et al* 1993). GSC is written within GeneSys, the development environment of the Cardiff Grammar Generator (Fawcett & Castel 2006). GSC accounts for core phenomena of RPS clitics in simple clauses: the word order patterns governing the occurrence of clitics with other clitics and with their governing verbs, and the distribution of participant roles.

References

Arús, Jorge (2006) "Perspectiva sistémico-funcional de los usos de *se* en español", *Signos* 39 (61): 131-159.

Castel, Víctor M. (2007) "Rule Types in a Systemic Functional Grammar: An XML Definition of the Cardiff Lexicogrammar Generator". In Barbara, L. y T. Berber Sardinha (ed.) *Proceedings of the 33rd International Systemic Functional Congress*. San Pablo: PUCSP.

Fawcett, Robin P. (2000) *A Theory of Syntax for Systemic Functional Linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Fawcett, Robin P. (2008) *Invitation to Systemic Functional Linguistics. The Cardiff Grammar as an extension and simplification of Halliday's Systemic Functional Grammar*. London: Equinox.

Fawcett, Robin P., G.H. Tucker y Y.Q. Lin (1993) "How a systemic functional grammar works: the role of realization in realization." En Horacek, H. y M. Zock (eds.) *New Concepts in Natural Language Generation*. London: Pinter.

Fawcett, Robin P. and V. M. Castel (2006) *Software for GeneSys: Prototype Generator 3*. Cardiff UK: Computational Linguistics Unit, Cardiff University.

García, Érica C. (1975) *The Role of Theory in Linguistic Analysis. The Spanish Pronoun System*. Amsterdam/Oxford: North-Holland Publishing Company.

Perlmutter, David M. (1972) *Deep and Surface Structure Constraints in Syntax*. New York: Holt and Co.

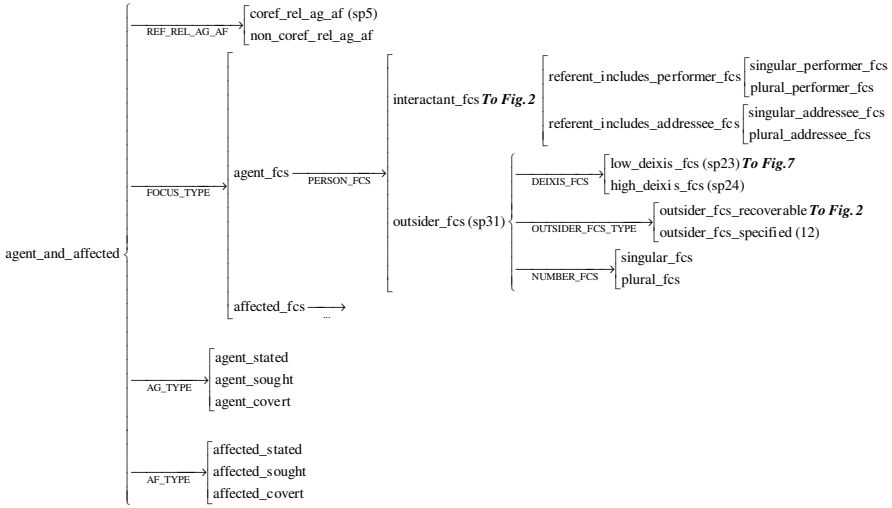

Figure 1

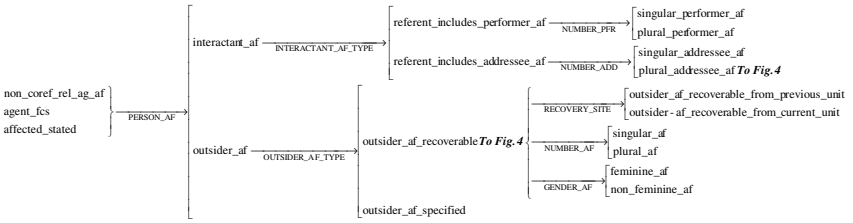

Figure 3

Figure 2

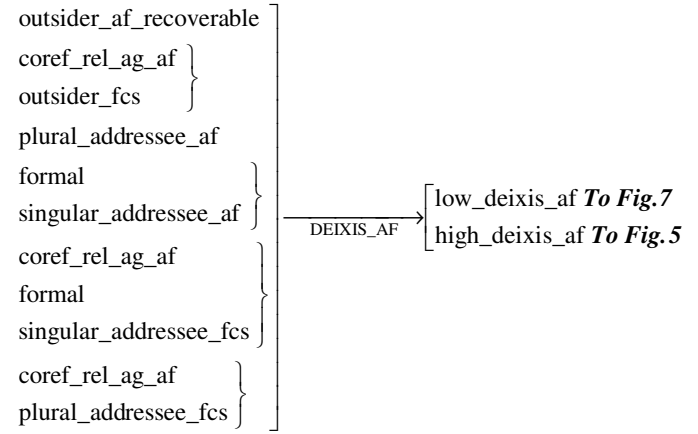

Figure 4

Figure 5

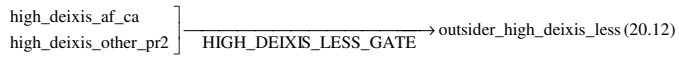

Figure 6

Figure 7

Figures 1-7: Systems Relevant for Clitic Realization Rules

```
L1   20.11 : outsider_high_deixis_least :
L2      [if not (proposal_for_action or infinitive or gerund) then CltA @ 98,
L3         else CltA @ 105],
L4      [if high_deixis_af then Af by CltA],
L5      [if high_deixis_af_po then Af-Po by CltA],
L6      [if (singular_af or singular_af_po or (formal and (singular_addressee_af
L7         or singular_addressee_af_po))) and (non_feminine_af or
L8         non_feminine_af_po) then CltA < "lo"],
L9      [if (singular_af or singular_af_po or (formal and (singular_addressee_af
L10        or singular_addressee_af_po))) and (feminine_af or
L11        feminine_af_po) then CltA < "la"],
L12     [if (plural_af or plural_af_po or plural_addressee_af or
L13        plural_addressee_af_po) and (non_feminine_af or
L14        non_feminine_af_po) then CltA < "los"],
L15     [if (plural_af or plural_af_po or plural_addressee_af or
L16        plural_addressee_af_po) and (feminine_af or feminine_af_po)
L17        then CltA < "las"].

L1   20.12 : outsider_high_deixis_less :
L2      [if not (proposal_for_action or infinitive or gerund) then CltD @ 98,
L3         else CltD @ 105],
L4      [if high_deixis_af_ca then Af-Ca by CltD],
L5      [if (singular_af_ca or (formal and singular_addressee_af_ca))
L6         and not outsider_af_po_recoverable then CltD < "le"],
L7      [if (plural_af_ca or plural_addressee_af_ca) and not
L8         outsider_af_po_recoverable  then CltD < "les"].

L1   20 : outsider_low_deixis :
L2      [if not (proposal_for_action or infinitive or gerund) then Clt3 @ 95,
L3         else Clt3 @ 102],
L4      [if low_deixis_fcs then Ag by Clt3],
L5      [if low_deixis_af then Af by Clt3],
L6      [if low_deixis_af_po then Af-Po by Clt3],
L7      [if low_deixis_af_ca then Af-Ca by Clt3],
L8      Clt3 < "se".

L1   20.2 : referent_includes_performer :
L2      [if not (proposal_for_action or infinitive or gerund) then Clt1 @ 97,
L3         else Clt1 @ 104],
L4      [if referent_includes_performer_af or (coref_rel_ag_af and
L5         referent_includes_performer_fcs) then Af by Clt1],
L6      [if referent_includes_performer_af_po or
L7      (agent_and_affected_possessed_coref and
L8         referent_includes_performer_fcs) then Af-Po by Clt1],
L9      [if referent_includes_performer_af_ca or
L10     (agent_and_affected_carrier_coref and referent_includes_performer_fcs)
L11        then Af-Ca by Clt1],
L12     [if singular_performer_af or singular_performer_af_po or
L13        singular_performer_af_ca or (coref_rel_ag_af and
L14        singular_performer_fcs) or (agent_and_affected_possessed_coref
L15        and singular_performer_fcs) or (agent_and_affected_carrier_coref
L16        and singular_performer_fcs) then Clt1 < "me"],
L17     [if plural_performer_af or plural_performer_af_po or
L18        plural_performer_af_ca or (coref_rel_ag_af and
L19        plural_performer_fcs) or (agent_and_affected_possessed_coref
L20        and plural_performer_fcs) or (agent_and_affected_carrier_coref
L21        and plural_performer_fcs) then Clt1 < "nos"].

L1   20.3 : singular_addressee :
L2      [if casual then [if not (proposal_for_action or infinitive or gerund)
L3         then Clt2 @ 96, else Clt2 @ 103],
L4      [if singular_addressee_af or (coref_rel_ag_af and
L5         singular_addressee_fcs) then Af by Clt2],
L6      [if singular_addressee_af_po or
L7         (agent_and_affected_possessed_coref and
L8         singular_addressee_fcs) then Af-Po by Clt2],
L9      [if singular_addressee_af_ca or
L10        (agent_and_affected_carrier_coref and
L11        singular_addressee_fcs) then Af-Ca by Clt2],
L12     Clt2 < "te"].
```

Figure 8: Clitic Realization Rules