Optimal Clitic Positions and the Lexicon in Romance Clitic Systems[1]

Jane Grimshaw, Rutgers University

October 1999

[Note: this paper is to appear as a chapter of G. Legendre, J.Grimshaw and S. Vikner eds. *OT Syntax*, in press, MIT Press.]

When pronominal clitics combine in the Romance languages, they occur in fixed orders, and in different fixed orders in different languages (and dialects), even when attention is limited, as here, to clitic placement in finite clauses.[2] The research program of explicating the systematicity within these systems, begun within generative grammar in Perlmutter (1971), has been challenged by the complexity of the facts, with many researchers simply concluding that the possible orders must be stipulated in a "template", or surface structure constraint. Monachesi 1995, Bonet 1995, Miller and Sag 1997 are recent examples.

The goal of the present paper is to show that some fundamental properties of the clitic systems of French, Italian and Spanish can be understood as a function of a set of ranked, violable morpho-syntactic constraints of three core kinds: faithfulness constraints, markedness constraints, and alignment constraints (Prince and Smolensky 1993, McCarthy and Prince 1993). This proposal is developed under the general theoretical assumptions of Optimality Theory (Prince and Smolensky 1993). Interaction among conflicting constraints is at the heart of the solution I propose for these complex and highly variable systems.

The relative order of clitics in Romance finite clauses follows from alignment constraints which relate morpho-syntactic specifications for case and person to the right and left edge of the clause (and hence indirectly the clitic cluster). Alignment interacts with the lexical representations of the clitics (Bonet 1995, Grimshaw 1997b) in ways which predict the positions of the clitics. The alignment constraints interact also with faithfulness constraints explaining the clitic substitution known as "spurious *se*". Markedness constraints also interact with the faithfulness constraints and this allows us to derive the inventory of clitic specifications seen in the lexicons of the languages under study, i.e. to predict which of the universal possibilities for pronominal specifications will actually appear in grammatical outputs. The functional lexicon is in this view the output of a set of violable universal constraints in interaction, i.e. it is a function of the grammar of a language, and not an independent set of stipulations concerning the morphemes which can appear in well-formed sentences.

1. The Clitic Lexicon

The clitic lexicons of Italian, Spanish and French are given in (1)-(3). (1) shows the non-reflexive third person forms, which mark distinctions in number, gender and case. This is illustrated for Italian by table 1 which shows the distribution of third person forms across the paradigm cells for case, person and number.

(1) The Romance lexicon: 3rd person non-reflexive forms

	It.	Fr.	Sp.
him/it	*lo*	*le*	*lo*
her/it	*la*	*la*	*la*
them (masculine)	*li*	*les*	*los*
them (feminine)	*le*	*les*	*las*
to him/it	*gli*	*lui*	*le*
to her/it	*le*	*lui*	*le*
to them	-	*leur*	*les*

Within first and second person forms, however, there are no case or gender distinctions, and all clitics are used both for reflexive and non-reflexive forms.

(2) The Romance lexicon: 1st and 2nd person forms

	It.	Fr.	Sp.
(to) me(self)	*mi*	*me*	*me*
(to) you(self)	*ti*	*te*	*te*
(to) us(self)	*ci*	*nous*	*nos*
(to) you(self)	*vi*	*vous*	*os*[3]

Finally, there is a clitic which occurs only as a third person reflexive:

(3) The Romance lexicon: 3rd person "reflexive" forms

	It.	Fr.	Sp.
self	*si*	*se*	*se*

Table 1 illustrates the resulting distribution of morphemes across the paradigm in Italian:

Table 1: The Italian clitic system[4]

	1sg	2sg	3sg	1pl	2pl	3pl
acc	mi	ti	lo/la	ci	vi	li/le
dat	mi	ti	gli/le	ci	vi	-
acc-ref	mi	ti	si	ci	vi	si
dat-ref	mi	ti	si	ci	vi	si

A single clitic form occupies all four slots for 1st and 2nd persons, while there are multiple forms for the 3rd persons, reflecting distinctions in case and reflexivity.

In addressing the issue of how these lexical items are represented, we have two basic possibilities. One option is to posit multiple occurrences of all the morphemes that occur more than once in the table. Thus we posit one clitic, *mi*, which is 1st person, accusative, non-reflexive. Then we posit a second clitic, also *mi* by accident, which is 1st person, dative, non-reflexive, and so on. The second option is to posit only one lexical entry for each morpheme, and treat the identity of forms across certain cells as non-accidental. In this case, we posit one morpheme *mi*, with no gender, case or reflexivity specification, rather than multiple clitics, each with a different specification. While either of these options, or indeed a combination of the two, is descriptively adequate for selecting the correct clitic form in isolation, more fully predicting the properties of Romance clitics depends crucially on the exact nature of their lexical representation. In the course of this paper, the lexical representations proposed in the present section will have consequences for constraints on clitic ordering, substitution of one clitic for another, and the markedness constraints which derive the clitic inventories themselves by determining which combinations of morpho-syntactic features are possible in the grammar of these languages.

It turns out to be critical, for example, that the clitic *mi* has a person specification and a number specification, but no specification for reflexivity, gender or case. On the other hand, the clitic *lo* will have specifications for all five properties. The outlined typeface, $\mathbb{R}$ for reflexivity, $\mathbb{P}$ for person, $\mathbb{N}$ for number, $\mathbb{G}$ for gender and $\mathbb{C}$ for case, indicates that the clitic concerned has no specification for that property. Thus a clitic represented as $\mathbb{R}$ means that a clitic is neither reflexive nor non-reflexive, but occurs in both uses. The Italian lexicon, analyzed in this way, is shown in (4).

(4) The Italian lexicon

lo	-R 3 sg m acc	him/it
la	-R 3 sg f acc	her/it
li	-R 3 pl m acc	them (masculine)
le	-R 3 pl f acc	them (feminine)
gli	-R 3 sg m dat	to him/it
le	-R 3 sg f dat	to her/it

mi	ℝ 1 sg 𝔾 ℂ	(to) me(self)
ti	ℝ 2 sg 𝔾 ℂ	(to) you(self)
ci	ℝ 1 pl 𝔾 ℂ	(to) us(self)
vi	ℝ 2 pl 𝔾 ℂ	(to) you(self)
si	+R ℙ ℕ 𝔾 ℂ	self

We can interpret the array in (4) as follows: all specifications are preserved in the least marked configuration, namely a third person non-reflexive clitic. Reflexivity, gender and case are all lost in the marked persons: 1^{st} and 2^{nd}, and all distinctions are lost in the (highly marked) reflexive. This is the consequence of markedness constraints interacting with faithfulness, as we will see in Section 5, where it forms the basis of the derivation of the inventory of clitics from constraint interaction..

2. Selecting the best clitic

Given this analysis of the clitic lexicons, what determines which clitic actually occurs? Why, for example, does an Italian speaker select the clitic *mi* for the 1^{st} person accusative feminine form, rather than *la*, which is both accusative and feminine? The answer is given by a process of optimization, of the kind familiar from Optimality Theory, which selects from the lexicon the clitic which is *most faithful* to the input. (Grimshaw 1997b presents a more detailed discussion, albeit with a slightly different lexical basis.)

The key points here are the following. The **input** is a morpho-syntactic specification constructed from the universal feature inventory for pronouns, e.g. [-R 3 pl f acc]. (See Section 5 for further elaboration.) The **candidates** are the set of pronouns in the language, lexically analyzed as discussed above. The **output** is the clitic with the lexical representation that best analyzes the input. The **selection** of the best clitic is conducted by a set of faithfulness constraints which require a match between the input specification and the output specification. For each relevant morpho-syntactic property admitted into Universal Grammar, there is a FAITH constraint which is violated if the specification in the input and the output differ (in any way).

Whenever the input is third person and non-reflexive, there is a candidate which parses the input perfectly, hence this candidate is always the optimum and is always chosen by the grammar. Consider, for example, table 2, which exemplifies the crucial comparisons:

Table 2: 3rd person non-reflexive inputs

input: [-R 3 sg m acc]	FAITH REFL	FAITH PERS	FAITH NUM	FAITH GEN	FAITH CASE
a. ℝ 1 sg 𝔾 ℂ	*!	*!		*!	*!
b.☞ -R 3 sg m acc					
c. -R 3 sg m dat					*!

Since the specifications of *lo* perfectly match those in the input, it satisfies all the faithfulness constraints, *under any ranking*. Every other candidate pronoun in the system violates at least one. For example, the candidate in a. violates FAITHREFL, FAITHPERS FAITHGEN and FAITHCASE, having no specification for reflexivity, gender or case, and the wrong person. The dative counterpart to the optimal candidate, in c. violates FAITHCASE, so is sub-optimal even though it satisfies all the other constraints.

The situation becomes more interesting as soon as we examine non-third person inputs, since for these inputs there is never a candidate which offers a perfect parse. The clitic *si* in candidate a. violates all faithfulness constraints except FAITHREFL. Since the candidate which violates FAITHREFL but satisfies FAITHPERS and FAITHNUM is in fact the winner, FAITHPERS or FAITHNUM must dominate FAITHREFL. (Subsequently we will see that parsing of reflexive is ranked above parsing of number, so the crucial ranking will prove to be FAITHPERS over FAITHREFL with FAITHREFL over FAITHNUM, as shown in the tableaux.) Among the other candidates, the best 3rd person candidate has better gender and case specifications than the winner, but the wrong person, thus FAITHPERS must dominate FAITHGEN and FAITHCASE.

Table 3: 1st and 2nd person inputs

input: [+R 2 pl m acc]	FAITH PERS	FAITH REFL	FAITH NUM	FAITH GEN	FAITH CASE
a. +R ℙ ℕ 𝔾 ℂ	*!		*!	*	*
b.☞ ℝ 2 pl 𝔾 ℂ		*		*	*
c. -R 3 pl m acc	*!	*			

In Table 4 we examine the same three candidates as in Table 3, but this time corresponding to a non-reflexive input. The same candidate is still selected as optimal. (The tableau shows only the candidates relevant for ranking: in the full tableau candidate b. harmonically bounds all other 3rd person outputs. Candidate a. harmonically bounds all other 1st and 2nd person outputs.)

Table 4: 1st and 2nd person non-reflexive inputs

input: [-R 2 pl m acc]	FAITH PERS	FAITH REFL	FAITH NUM	FAITH GEN	FAITH CASE
a. +R ℙ ℕ 𝔾 ℂ	*!	*	*	*	*
b.☞ ℝ 2 pl 𝔾 ℂ		*		*	*
c. -R 3 pl m acc	*!				

Here we have further evidence for the constraint rankings. Candidate b. is the output, despite the fact that it violates FAITHREFL, FAITHCASE and FAITHGEN, all of which are satisfied by candidate c. The violation of FAITHPERS incurred by candidate c. must be overriding these violations, as will be the case if FAITHPERS dominates FAITHREFL FAITHGEN, FAITHCASE,.

Finally we come to the case of 3rd person reflexive inputs. The winning candidate in Table 5 is faithful to the reflexivity of the input and to nothing else. The other candidates are more successful with respect to number (candidate b.) or to person, gender, case and number (candidate c.). Hence the rankings established so far will incorrectly select candidate c., which is faithful to person, over candidate a. which is faithful to reflexivity. This problem is solved if we conclude that FAITH REFL is really two constraints, one which penalizes absence of a reflexivity specification (PARSE) and one which penalizes presence of the wrong reflexivity specification (FILL). (This is in accordance with Grimshaw 1997b, where it is argued that the constraints at work are really the MAX and DEP families of McCarthy and Prince 1995, 1999.)

FILLREFL will penalize the candidate in c. with a -R specification, and be satisfied by *si* in candidate a. Provided that FILLREFL dominates FAITHPERS, FAITHGEN, FAITHCASE, FAITHNUM, the correct output will be selected. If PARSE REFL dominates FAITH NUM. candidate a. is selected over b. (Note that if FAITHPERS is also broken up into FILLPERS and PARSEPERS, the ranking of

FILLPERS with respect to FAITHNUM could be the critical one.)

Table 5: 3rd person reflexives

input: [+R 3 pl m acc]	FILL REFL	FAITH PERS	PARSE REFL	FAITH NUM	FAITH GEN	FAITH CASE
a.☞ +R ℙ ℕ 𝔾 ℂ		*		*	*	*
b. ℝ 1 pl 𝔾 ℂ		*	* !		*	*
c. -R 3 pl m acc	*!					

The earlier ranking evidence, we can now see, showed that FAITHPERS dominates PARSE REFL (rather than FAITHREFL) ensuring the choice of a 1st or 2nd person output which has no reflexive specification over 3rd person alternatives which may match the input for reflexivity.

Thus, by ranking the faithfulness constraints, and assuming the lexical representations developed in Section 1, we can predict which clitic surfaces as the output for each input. Later (Section 5), we will see that the lexical representations themselves are derived by constraint rankings.

3. Romance Clitic Order: Interaction among Alignment Constraints

When clitics combine, as mentioned above, they do so in fixed orders, different for different linguistic systems. Given the lexicon developed in Sections 1 and 2, the relative order of clitics in combination can be determined by a set of alignment constraints on morpho-syntactic specifications for person and case. For related work proposing that clitics are subject to alignment constraints see Anderson 1996, in press, Gerlach 1998, Legendre (1996, in press, this volume b) Van der Leeuw 1995. It is important to note that these constraints govern *the order* of clitics when they combine with each other, not the well-formedness of the combinations in the first place, which is a matter for markedness constraints (see Section 4).

Since only person and case specifications prove important for alignment, we can simplify the lexicon to that given in (5).

(5) The Italian lexicon showing person and case only

lo	3 acc	him/it
la	3 acc	her/it
li	3 acc	them (masculine)
le	3 acc	them (feminine)
gli	3 dat	to him/it
le	3 dat	to her/it
mi	1 ℂ	(to) me(self)
ti	2 ℂ	(to) you(self)
ci	1 ℂ	(to) us(self)
vi	2 ℂ	(to) you(self)
si	ℙ ℂ	self

The order of clitic combinations has been taken to be essentially inexplicable, with the restrictions normally stated by means of a positive output constraint, or “template”, which stipulates the possible orders and combinations of clitics. Templates proposed in the literature for French, Italian and Spanish are given in (6)- (8):

(6) French (based on Perlmutter 1971:57)

me	3-acc	3-dat	y	en
te				
nous				
vous				
se				

(7) Italian: (Monachesi 1995:122)[5]

mi	ci	si	lo	si	ne
ti		(Ref)	la	(imp)	
gli			li		
le			le (acc)		
ci					
vi					

(8) Spanish: (based on Perlmutter 1971:45)

se	2	1	3

Although the French and Spanish templates as presented by Perlmutter reflect the existence of clear morpho-syntactic patterns in clitic combinations, the template notation clearly renders the positions of the clitics essentially arbitrary, at least in the absence of a theory of what a template can look like. While the Spanish generalizations might lead one to suspect that person specifications are the key, it is clear that person specifications alone are not sufficient to determine the ordering of clitics in French or Italian. 3rd person forms appear in the first and fourth columns in Italian for example, and the first column contains 1st , 2nd and 3rd person clitics. However, even the Italian clitic template is not random: all 3rd person accusative clitics are on the right, for example. It certainly appears that the morpho-syntactic properties of the clitics determine where they appear in the template. But how is this accomplished, why should it be so, and what predicts the exact outcome?

3.1 French

The template in (9) shows the French clitic order temporarily simplified to include only personal clitics.

(9)	A	B	C
	1,2	3-acc	3-dat
	se		

In terms of the lexical analysis given here, the three columns can be summarized as follows: *clitics in column B have an accusative case specification, clitics in column C have a dative case specification, and clitics in column A have neither*. The clitics in column A are thus the complement of those in B and C.

This array can be derived from interacting constraints relating case to position. Specifically, we can posit positional constraints DATRT and ACCRT, which are satisfied when an element with the relevant case specification is on the right hand edge of the clitic cluster, violated otherwise. Such positional constraints are always satisfied by any clitic in isolation; a 3rd person accusative by itself necessarily satisfies ACCRT (and ACCLFT, see below) for example. Hence the constraints have no effect except on clitic combinations.

The alignment constraints DATRT and ACCRT both require that a case-specified clitic occur on the right, hence when only one clitic in a combination has a case specification, that clitic will appear on the right hand edge. However, when there are two, as there can be if one is dative and one is accusative, provided that the constraint targeting datives is the dominant one, the system will settle on the choice of the dative on the edge. Thus, clitics from column A will precede clitics from column B, as illustrated in Table 6, and clitics from B column will precede clitics from C, as illustrated in Table 7.

Table 6: French 1st person + accusative

input: [1dat + 3 acc]	DAT RT	ACC RT
a.☞ 1ℂ 3acc		
b. 3acc 1ℂ		*!

When any 1st or 2nd person clitic combines with an accusative, as in Table 6, the constraint which mentions dative is irrelevant, since neither clitic is dative. Hence ACCRT is the only relevant constraint and it positions the accusative on the right edge. Table 7 illustrates the situation when an accusative and a dative cooccur.

Table 7: French 3rd person -- dative + accusative

input: [3 acc + 3 dat]	DAT RT	ACC RT
a.☞ 3acc 3 dat		*
b. 3 dat 3acc	*!	

The dative clitic must be on the right hand edge to satisfy the dominant DATRT constraint. The position of the 3rd person accusative violates ACCRT but there is no better configuration. Note the argument for the crucial ranking: if ACCRT >> DATRT, then b. is optimal instead of a.

It is apparent that the lexical representation of the clitics is crucial in predicting clitic order; it determines how the clitics interact with the constraints. The clitic corresponding to a [1 dat] input in Table 6, like all 1st and 2nd person clitics and *se,* has no case specification, hence it vacuously satisfies both the constraints: it satisfies them regardless of where it is positioned. When it combines with any clitic with a case specification, that clitic will be positioned by DATRT or ACCRT, hence the clitic with no case will end up on the left. All clitics which have no case specification go in the residual position, on the left, as a result of their satisfying the constraints vacuously. This illustrates the role of constraint interaction in picking out the *complement* of what is targeted by constraints. The clitic corresponding to the [1 dat] input is in the position of an element that is not dative and is not accusative. *Se*, the 3rd person reflexive clitic, occurs in the same slot because it also lacks a case specification. This supports the lexical analysis which posits one clitic with no case for each of the 1st and 2nd person forms, rather than two fully specified clitics. If they were specified for accusative or dative they would behave just like third persons, since they would not satisfy the case constraints vacuously.

I assume, in accordance with the leading ideas of alignment (see in particular Samek-Lodovici 1998), that the two constraints at work in French are members of a universal pool which consists of all possible combinations of case and person specifications with "RT" and "LFT" (we will see direct evidence from Italian that alignment constraints like these can target the left edge of the cluster). It follows that DATLFT and ACCLFT are also constraints, and hence also in the grammar of French: they must be ranked below DATRT and ACCRT respectively. This will guarantee that it is DATRT and ACCRT that make the key decisions in the language. The analysis will thus require the crucial rankings in (10).

(10) French Ranking: DATRT >>DATLFT, ACCRT >>ACCLFT

DATRT >>ACCRT

DATLFT and ACCLFT are omitted from the tableaux above, and the same simplification strategy is followed below. Positioning each constraint anywhere to the right of its opposite will leave the selection of the optimum unchanged. See 4.3 for further analysis of ranking among such matched pairs of alignment constraints.

Finally, clitics *y* (a locative) and *en* (a partitive element) have been omitted from the analysis so far. Since they have no person or case specification, their positioning will not be affected by alignment constraints on positioning these specifications. They can be properly positioned by the constraints PARTITIVE RIGHT and LOCATIVE RIGHT, under a ranking in which PARTRT >> LOCRT, and both constraints dominate all other constraints discussed here.

To sum up, the sequencing information in the clitic "template" for French can be derived by ranking the constraint placing datives on the right over the constraint placing accusatives on the right. In contrast to standard modern French, Old French had systematic accusative-dative order, and some regional French dialects have dative-accusative (Morin 1979). Re-ranking the posited alignment constraints derives these alternative systems. For example, if ACCLFT dominates the constraint which left aligns datives, as well as ACCRT, a sequence will occur in accusative-dative order. If ACCRT dominates ACCLFT and the conflicting dative constraint, a dative accusative order will emerge..

3.2 Italian

The template for Italian personal clitics appears in (11):

(11)	A	B	C
	1,2 dative	si	3-acc

The three columns can be summarized as follows: *clitics in column A have a person specification and are not accusative. Clitics in column C have an accusative case specification, and clitics in column B have neither*. In describing the columns this way we highlight the fact that they do not all contain natural classes of clitics. In particular column A in (11) can only be understood as the complement of column C, which is exactly the effect that constraint interaction will derive. The primary generalization is that all accusatives go on the right. Clitics which have a person specification but are not accusative, i.e. the *remaining* person specified clitics, go on the left, and the still-remaining clitic *si*, which has no properties, goes in the middle. We can derive this pattern if Italian has AccRt as the dominant member of the alignment constraints for accusative, and PersLft as the dominant member of the alignment constraints for person. The accusative dominates the person constraint, and the Italian pattern emerges.

(12) Italian Ranking: AccRt >> AccLft, PersLft >> Per Rt

AccRt >> PersLft

If an accusative clitic combines with any clitic with no accusative specification, the accusative will occur at the right edge, see Tables 8 and 9. The sequencing in Table 9 motivates the ranking: the violation of PersLft in the first candidate is outweighed by the violation of AccRt in the second.

Table 8: Italian – accusative with non-accusative

input: [1dat + 3 acc]	AccRt	PersLft
a.☞ 1ℂ 3acc		*
b. 3 acc 1ℂ	*!	*

Table 9: Italian – accusative with non-accusative

input: [+R 3 + 3 acc]	AccRt	PersLft
a.☞ ℙℂ 3acc		*
b. 3 acc ℙℂ	*!	

If a non-accusative person clitic combines with a clitic with neither case nor person, the latter will follow the former:

Table 10: Italian – non-accusative with clitic with no case, person

input: [+R 3 + 1 acc]	AccRt	PersLft
a. ℙℂ 1ℂ		*!
b..☞ 1ℂ ℙℂ		

In sum, clitics from column A will precede *si* (in column B) and clitics from both A and B will precede those from column C: this system will guarantee that any accusative clitic appears on the right, and that any clitic with a person specification but no accusative specification will appear on the left. The same ordering of clitics will hold regardless of which accusative clitic occurs, because they all have the same specification for case and person. Similarly, the same result will hold regardless of which of the 1st or 2nd person clitics occurs, since nothing relevant distinguishes them. The third person dative clitics *gli* and *le* pattern with the 1st and 2nd person forms because they all satisfy Acc Rt vacuously. Hence Pers Left positions them at the left edge, just like *mi*. Because the alignment constraints relate morpho-syntactic specification to position, the lexical properties of the clitics are crucial in sequencing them.

Positional effects in turn provide evidence about lexical representation. For example, the same morpheme, *le* is used as the feminine dative singular and the feminine accusative plural. It would be incorrect, however, to treat this as a case of under-specification, assigning *le* no case. This would predict that it would always occur on the left edge of the cluster, even when used as a direct object. (The markedness constraint system to be presented in Section 5 correctly analyzes *le* as fully specified.)

The full template in (12) also contains positions for a partitive, an impersonal subject and a locative. Characterizing *ci* as locative, *ne* as partitive, and assuming for simplicity that the impersonal subject *ci* has no case or person specification, the full set of orders is derived from the constraint hierarchy in (13), which includes the constraints on personal forms as a subpart.

(13) PARTRT >>IMPRT>>ACCRT>> PERSLFT >> LOCLFT

3.3. Spanish

The orders found among clitics in Spanish are given in (14), with combinations of 1st and 2nd person forms temporarily omitted from consideration.

(14)

A	B	C
se	1/2	3

The three columns can be summarized as follows: *clitics in column B have person but no case specification, those in column C have both, and those in column A have neither*. Once again, then, we see the effects of positional constraints interacting with the lexical properties of clitics. The Spanish pattern will be the result of right alignment constraints mentioning case and person, regardless of the ranking between them:[6]

(15) CASERT >>CASELFT

PERSRT >>PERSLFT

Table 11 shows the sequencing of a clitic from column A and one from column B: candidate b. wins under either ranking of CASERT and PERSRT.

Table 11: Spanish -- *se* with 1st or 2nd person

input: [3 refl + 1 dat]	CASE RT	PERS RT
a. 1ℂ ℙℂ		*!
b. ☞ ℙℂ 1ℂ		

Table 12 shows the sequencing of a clitic from column B and one from column C: candidate a. wins under either ranking.

Table 12: Spanish – 1st or 2nd person with 3rd person

input: [1 dat + 3 acc]	CASE RT	PERS RT
a. ☞ 1ℂ 3acc		*
b. 3acc 1ℂ	*!	*

Finally, Table 13 establishes the sequencing of a clitic from column A and one from column C:

Table 13: Spanish – *se* with 3rd person

input: [3 refl + 1 dat]	CASE RT	PERS RT
a. ☞ ℙℂ 1ℂ		
b. 1ℂ ℙℂ		*!

The case constraint forces a case marked clitic to the right edge. Of the remaining clitics, one with person goes as far to the right as it can. The then remaining clitic then occupies the remaining position. Thus the grammar does not have to include characterizations such as “has person but not case”: that the facts can be described this way is the result of interacting simple constraints referring only to individual grammatical properties.[7] The complex effects are due to interaction.

There is no possible conflict between CASERT and PERSRT, since the clitics which are affected by CASERT are a subset of those affected by PERSRT, and both constraints position the clitics on the same edge. A clitic which had a case specification but no person specification (non-existent under the present analysis), combined with a clitic with a person specification but no case specification (which does exist under the present analysis), would allow us to detect the ranking: with the person constraint dominant the second clitic would be on the specified edge, with the case constraint dominant the first clitic would be on the specified edge. Thus, this is an instance where there happens to be no direct evidence for the relative ranking of two constraints, and *not* an instance where two constraints are crucially unranked.

To complete the picture, we need to examine the sequencing of 1st and 2nd person clitics, which can cooccur in Spanish, although not in French and Italian (see Section 4.2). 2nd person clitics precede 1st persons in standard Spanish (Perlmutter 1971). This will follow if a constraint aligning 1st person specifications with the right edge (1RT) dominates the constraints 1LFT and 2RT, whose existence follows from the general theory of alignment constraints (cf. Section 4.3).[8] Since the order in which 1st and 2nd clitics occur depends only on the ranking of 1RT and 2RT, we expect dialect variation in this respect.

Table 14 shows that a (perhaps hypothetical) sequence with *se* (the clitic with no person or case) in initial position, followed by a 2nd person, followed by a 1st person will be optimal under this ranking, the revealing comparisons being between the winner and candidate a. in which the 1st person precedes the 2nd and between the winning candidate and any candidate (such as c.) in which a 1st and/or 2nd person clitic is to the left of a clitic with no person marking.

Table 14: Spanish

input: [3 reflexive + 2 acc + 1 dat]	CASE RT	PERS RT	1 RT	2 RT/LFT
a. ℙℂ 1ℂ 2ℂ		*	*!	**
b. ☞ ℙℂ 2ℂ 1ℂ		*		**
c. 1ℂ 2ℂ ℙℂ		**!*	**	**

When 1st and 2nd person clitics combine with a case marked clitic, as in Table 15, (Perlmutter's (96), p 51)), the 2nd person will again precede the 1st, and the case marked clitic will follow, provided that CASERT >> 1RT. Thus we have shown that the 2nd person will always precede 1st in any combination.

Table 15: Spanish – 2nd and 1st persons, gradience of alignment

input: [2 dat + 1 dat + 3 acc]	CASE RT	PERS RT	1 RT	2 RT/LFT
a. 3acc 1ℂ 2ℂ	*!*	***	*	**
b. ☞ 2ℂ 1ℂ 3acc		***	*	**
c. 1ℂ 2ℂ 3acc		***	**!	**

All candidates in which the case marked clitic is not at the right edge are eliminated, since there are competitors (such as b. and c.) which do satisfy CASERT. Of the two candidates which satisfy CASERT, the one which best satisfies 1RT is optimal. The fact that candidate c. is ungrammatical and b. grammatical, shows that 1RT, and the clitic alignment constraints in general are *gradient* constraints, see Section 4.3.

Other variation concerns the relative order of *se*: in Murcian (Heap 1996:231) *se* follows rather than precedes the 1st and 2nd persons, while still preceding 3rd persons. This will follow from an analysis in which CASERT >> CASELFT, as in standard Spanish, but PERSLFT >> PERSRT, with CASERT then dominating PERSLFT.[9] Table 16 illustrates the effect of this in a compact form, for a hypothetical input including a reflexive, a 1st person and an accusative form. It shows only candidates which satisfy CASERT, by having the accusative on the right edge.

Table 16: Spanish – variation in the position of *se*

input: [3 dat + 1 dat + 3 acc]	CASE RT	PERS LFT
a. ☞ 1ℂ ℙℂ 3acc		**
b. ℙℂ 1ℂ 3acc		***!

Candidate a. is optimal under the present ranking, while b. was optimal under the ranking of standard Spanish, where PERSRT is the dominant person alignment constraint.

3. 4 Alignment interacts with faithfulness: A Clitic substitution in Spanish

Other things being equal, the fewer morpho-syntactic properties a clitic has, the fewer alignment constraints it can violate. If S-1 is the set of specifications of clitic 1 and S-2 is the set of specifications of clitic 2, and S-1 is a subset of S-2, then clitic 1 can satisfy positional constraints which are violated by clitic 2, if S-1 does not contain the specification targeted by the positional constraint(s). In this way, positional constraints themselves can force the substitution of one clitic for another, otherwise expected, clitic form.

The combination of two 3rd person forms must violate both the dominant case constraint and the dominant person constraint. Since each clitic has a specification for both case and person, and only one clitic can be in the position demanded by the highest ranking of the case constraint pair and/or the highest ranking of the person constraint pair, such a combination must violate both constraints. By the same token, a combination of one 3rd person form and one 2nd or 1st person form will always violate the person constraint, since again only one of the clitics can be on the edge specified for person. The same is true for the combination of one 2nd or 1st person clitic with another 2nd or 1st person clitic. Spanish offers a well known example of a system which does not tolerate combinations of 3rd person clitics. The data in (16) are taken from Bonet 1995.

(16) a. E1 premio, **lo** dieron a Pedro ayer.

the prize 3rd-acc gave(3rd-pl) to Pedro yesterday

b. A Pedro, **le** dieron el premio ayer.

to Pedro 3rd-dat gave(3rd-pl) the prize yesterday

c. A Pedro, el premio, **se lo** dieron ayer. (***le lo** ***lo le**)

to Pedro the prize se 3rd-acc gave(3rd-pl) yesterday

"they gave the prize to Pedro yesterday"

Where, based on the forms that occur in isolation, we expect to find both a 3rd person dative and a 3rd person accusative, what instead surfaces is a 3rd person accusative accompanied by *se*. Why doesn't *se* occur in place of a 3rd person dative elsewhere? The answer is that the faithfulness constraints will prefer the faithful clitic, which parses person, number, and case. However, as we saw above, *se*, with no case or person specification can improve performance of a combination on the positional constraints, so *se* will "replace" the expected 2rd person clitic if the relevant faithfulness constraints are dominated by the relevant positional constraint.[10] (See Bonet 's work and Grimshaw 1997b for analysis of agreement with spurious *se*.)

Table 17: Spanish – spurious *se*

input: [-R 3 sg dat + -R 3 sg acc]	CASE RT	PERS RT	FAITH REFL	FAITH PERS	FAITH NUM	FAITH CASE
a. -R 3 sg dat + -R 3 sg acc	*!	*!				
b. ☞+R ℙ ℕ ℂ + -R 3 sg acc			*	*	*	*
c. R 𝟙 sg ℂ + -R 3 sg acc		*!	*	**!		*

Table 17 shows that *le + lo* (candidate a.) violates PERSRT and CASERT (one of them fatally) while *se* has neither person nor case so it satisfies both vacuously. Thus the candidate which contains *se* (candidate b.) satisfies both positional constraints. It violates the faithfulness constraints for reflexivity, person number and case, but provided these are dominated by at least one of CASERT and PERSRT, the right candidate will be selected.[11]

(17) Either: CASERT >> FAITHR, FAITHPERS, FAITH NUM, FAITH CASE

Or: PERSRT >> FAITHR, FAITHPERS, FAITH NUM, FAITH CASE

Why replace *le* with *se*, and not, for example, with a 1st or 2nd person clitic (candidate c.), which would be faithful to the number in the input, if not to the other specifications? Like candidate b., this option would satisfy CASERT, since the first clitic in the sequence would have no case. This candidate, however, incurs a violation (one fatal) on both PERSRT and FAITHPERS (since it has two clitics bearing person specifications, and hence violates PERSRT, as well as both the PARSE and the FILL form of faithfulness, indicated by two stars in the FAITHPERS column of Table 17. We conclude that at least one of FAITHPERS and PERSRT must dominate FAITHNUM, so that the violation of FAITHNUM in candidate b. is preferred to the extra violations of PERSRT and FAITHPERS in candidate c.

Table 17 raises an additional question: why is *lo* retained instead of *le*? That is, why is the dative clitic replaced by *se* and not the accusative clitic? This is explained if we posit a universal markedness hierarchy, in which datives are more marked than accusatives: MARKDAT >> MARKACC. (See Sections 4 and 5, Aissen, this volume and Woolford, this volume, for further discussion of markedness in this system. The two candidates (*se le* and *se lo*) are otherwise identical, and *se le* is eliminated by MARKDAT.

In order to complete the picture, one more generalization needs to be addressed. If Spanish prefers *se* to *le* in combination with a 3rd person clitic, why doesn't it also prefer *se* to the 1st and 2nd person datives? That is where we expect *le + lo* we find *se + lo*, but where we expect *me + lo*, that is exactly what we get. Yet *me + lo* contains a PERSRT violation, just as *le + lo* does. Ranking can account for this observation. While *me + lo* does violate PERSRT it does not violate CASERT, since the 1st and 2nd person clitics have no case. So if we adopt a ranking in which CASERT alone dominates FAITHREFL, FAITHPERS, FAITHNUM and FAITHCASE, while PERSRT is dominated by the faithfulness constraints, we will derive the result that faithful parsing will be overridden by the case constraint but not by the person constraint. Thus we have evidence for the first of the two ranking possibilities listed in (17). (A minor variation of this solution posits the ranking: FAITH1,2 >> PERSRT, so that faithful parsing of 1st and 2nd person specifications in the input takes precedence over elimination of the positional violation for person).

Since there is a crucial constraint ranking involved in deriving spurious *se* in Spanish, namely between the faithfulness constraints and the positional constraint(s) it is not a surprise that there are grammars, such as those of French and Italian, which do not have this clitic substitution but simply retain the combination of two 3rd person forms.

Figure 1: Summary of Constraint Rankings

```
French:            DATRT
                    /  \
               ACCRT    DATLFT
                 |
               ACCLFT

Italian:           ACCRT
                    /  \
              ACC LFT   PERS LFT.
                          |
                        PERSRT

Spanish:
           CASERT                       PERSRT
            /  \                          |
     CASELFT   1RT                      PERS LFT
                /  \
            1LFT   2RT
```

Of course, every constraint is present in the grammar of each language: those that play no active role in positioning clitics are ranked below the constraints which are critical: below DATRT and ACCRT for French, below ACCRT and PERSLFT for Italian, and below CASERT PERSRT and 1RT for Spanish.

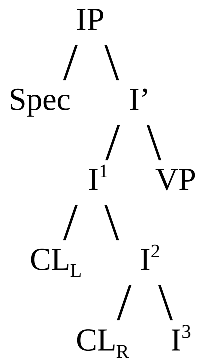
4. Romance Cliticization and Alignment

4.1 The Alignment Constraints

The positional constraints are clearly affiliated with the family of alignment constraints, (McCarthy and Prince 1993), which require that the edge of one constituent occur at the edge of another. See the references cited in Section 3, for related proposals involving the role of alignment constraints in cliticization. In this view, there are no constraints directly imposing order among clitics. Rather the clitics are competing for positions at the edge of the cluster, just as, for instance, heads and Specifiers compete for edges of syntactic projections, X' and XP, in Grimshaw 1997a: 406-409.

Much recent work, e.g. Kayne 1991, Haverkort 1993 takes the position that Romance clitics left-adjoin to a functional head, see Uriagereka 1995 and Terzi 1999 for more detailed analysis. (For concreteness, I will assume the structure given in Figure 2 for all Romance cliticization in finite clauses.[12]) The notation $CL_{L/R}$ designates a clitic with a specification subject to left/right alignment.

Figure 2: Structure of Finite Clauses with Clitics



A left alignment constraint requiring coincidence of the left edge of the clitic with the left edge of I’ would be satisfied by CL_L in Figure 2. Alternatively, and perhaps more interestingly, the alignment may be between the clitic and the left edge of IP, violated in Figure 2 because of the higher ranking SPECLFT constraint (Grimshaw 1997a).[13] In the notation of McCarthy and Prince, the constraints are stated as in (18).

(18) ALIGN (Specifier, Left, XP, Left) >> ALIGN (case/person, Left, IP, Left)

As for rightward alignment, again there is again more than one possibility. The alignment could be to the right edge of I, or I’, or IP. All of these constraints are violated by CL_R in Figure 2. In the interests of maintaining symmetry between the left and right alignment constraints, let us posit the right edge of IP as the target for alignment in this case. Why then is alignment violated (by the intervention of I and VP between the clitic and the right edge of IP)? Switching the order of CL_R and its sister I in Figure 2 would improve alignment for CL_R, so some higher ranked constraint must prevent this. A candidate for this constraint is ADJUNCT-LFT (requiring adjuncts to occur at left edges), which will govern configurations where order is not determined by constraints on heads or specifiers. If adjunct alignment takes priority over the case/person right alignment constraint, as in (19), clitics will precede what they adjoin to.

(19) ALIGN (Adjunct, Left, host, Left) >> ALIGN (case/person, Right, IP, Right)

Reversing the order of I and its complement VP will also improve right alignment of the clitic to IP: this is prevented by the constraint HDLFT (Grimshaw 1997a), which enforces I-VP order rather than VP-I order under I'.

In accordance with McCarthy and Prince's schema, the interpretation of a clitic alignment constraint such as ACCRT is as follows:

(20) $\forall$ accusative-marked X^0 $\exists$ IP such that the right edge of X^0 and the right edge of IP coincide

That these constraints and rankings give the right results is confirmed by Table 18, based on the analysis of Italian and consistent with the rankings just established, where (in the optimal candidate) the first clitic in the sequence is in the position of CL_L in Figure 1, and the second clitic is in the position of CL_R.

Table 18: Alignment of Clitics

input: [1 dat + 3acc]	ADJRT	ACCRT	SPECLFT	PERSLFT
a.☞ Spec 1ℂ 3acc I		**		***
b. Spec 1ℂ I 3acc	*!			****
c. 1C Spec 3acc I		**	*!	**

Comparison between candidates a. and b. shows that reversing the order of the I and the accusative clitic is ruled out because it violates ADJRT, although it satisfies ACCRT. Similarly, reversing the order of the specifier and the person-marked clitic in c. violates SPECLFT, and is thus not possible, even though it improves evaluation with respect to PERSLFT.

One last point of clarification. Both of the clitics in (38) are separated from the right edge of IP by an I (I^2 for CL_L and I^3 for CL_R) and a VP. In order to be sure that CL_R better satisfies right alignment when it is to the right of CL_L than in the opposite order, the calculation of alignment violations must be based on the number of terminal elements which intervene between the aligning element and the specified edge. CL_R is separated from the right edge of IP by I^3 and VP, if it is the rightmost clitic, but by I^3, VP and the other clitic, if another clitic follows it.

In this view, the positioning of Romance clitics is accomplished by alignment constraints relating morpho-syntactic specifications to the right and left edges of syntactic constituents: specifically, by hypotheis, IP. This raises two fundamental typological issues. First, if all of the morpho-syntactic constraints dominate the constraints on specifiers and heads, then the clitics should occur on the edges and the other constituents in the interior of a clause: the sequence Cl_L - Spec - I'- Cl_R satisfies clitic alignment constraints requiring one clitic to be on the left, and one on the right, of IP. Second, if some, but not all, of the morpho-syntactic constraints dominate the constraints on structure, then some clitics could precede, and some follow, either the specifier or the I’, a point made by Nicole Nelson (p.c.). Neither of these two scenarios appears likely: both would be eliminated by a general principle requiring that structural alignment constraints (those which mention only items of syntactic theory, such as head, specifier, X’) must always dominate morpho-syntactic alignment constraints (those which mention specifications such as person or case).

Viewed from a perspective in which clitics are ordered by violable alignment constraints, the sequencing of clitics is not very different from other word order phenomena; the only characteristic which makes clitics different from, say heads or specifiers being that they are positioned by virtue of their morpho-syntactic properties.

4. 2 Alignment, Markedness and Well-formed Clitic Combinations

The analysis developed up to this point concerns the order of clitic combinations: nothing predicts which combinations are possible in the first place. It appears, in fact, that regardless of whether templates or alignment constraints are at work, the well-formedness of a combination on the one hand, and the order in which its elements must occur on the other, are distinct problems. This point is partially disguised in templates, which often seem to contain both kinds of information – what can occur with what and in what order. The illusory nature of this is clear on inspection, however. For example, the template for French, in (6), properly sequences 1st and 2nd persons before, and 3rd person datives after, the 3rd person accusatives. This leads us to expect that 1st and 2nd person forms should precede 3rd person datives. Perlmutter (1971, 63-65) argues that this combination is impossible, and the only way a 3rd person indirect object can be realized in combination with a 1st or 2nd person clitic is by a strong form, which has a non-contrastive interpretation in these circumstances. (Simpson and Withgott 1986: 160-161 cite examples in which 1st or 2nd person ethical datives co-occur with 3rd person datives). Thus, while it may be true that templates can represent the impossibility of some combinations by placing the items which do not co-occur in a single column, they certainly cannot represent *all* restrictions this way. Once an analysis is developed then, to eliminate the impossible combinations from different columns of a template, it is necessary to ask whether it naturally extends to impossible combinations of clitics which can be grouped into a single column. Indeed, this appears to be the case.

Alignment constraints alone will never predict that combinations are impossible: just that the best aligned version of a combination will be grammatical. Alignment constraints in combination with others, however, can predict the impossibility of combinations. When alignment constraints interact with faithfulness, as seen above in Section 3.4, a clitic combination can be ruled out completely, with a less faithful clitic replacing one of the expected elements, in order to improve alignment. Other effects, I argue here, concern markedness, and are closely related to the markedness factors which underlie the lexicon itself (see Section 5). If 1st and 2nd persons, dative, and reflexive are more marked than, respectively, 3rd person, accusative and non-reflexive (see Aissen this volume, Artstein 1998, Woolford this volume, for recent OT work on these issues), it turns out that the impossible combinations of clitics are those which involve the marked values for case, person and/or reflexivity. Space limitations prevent a full analysis here: a sketch of the analysis of one prohibited combination follows.

Italian (Monachesi 1995) and Spanish (Bonet 1994), like French, show prohibitions against combining 1st and 2nd persons with dative clitics. This is illustrated for Spanish in (21), with data from Bonet 1994. (Note that the Spanish template proposed by Perlmutter cannot account for this generalization).

(21) a. Me recomendaron a él porque era el más influyente

b. *Me le recomendaron porque era el más influyente

They recommended me to him/her because s/he was the most influential

It appears that the optimal forms in such circumstances are alternative candidates with full pronouns, such as (21a). When an input is specified as weak, by virtue of its discourse status (cf. Grimshaw and Samek-Lodovici 1998) the best output normally will be a clitic form; I refer to the constraint which mandates this as WEAK (Bonet's 1994 "Avoid Pronoun Strength" embodies the same idea). However, a clitic combination may violate a markedness constraint which conflicts with, and dominates, the constraint mandating a weak form. In this circumstance a strong form (a pronoun) will be optimal instead, even as the realization of a weak element in the input. In support of this view, Bonet 1994:43 observes that in Spanish, a strong pronoun is possible without a contrastive focus interpretation and without a doubled clitic, only when the use of the strong pronoun avoids a violation of what she calls the "Person-Case constraint", which bans datives with 1st and 2nd person clitics.

The core of the analysis of (21b) is that the ill-formed clitic combination violates a markedness constraint against combining the marked features 1st/2nd person and dative (thanks to Ron Artstein p.c. for key steps in this analysis). This markedness constraint dominates WEAK, and forces the non-clitic realization of one of the arguments, as in (21a). In isolation each argument is realized as a clitic, since the combination constraint is not violated, and combinations involving less marked elements are also realized with clitics: 1st or 2nd person with accusative, for example.

In this analysis, then, the combination of a dative and a 1st or 2nd person form is more marked than the combination of a dative and a 3rd person form or a 1st or 2nd person form and an accusative. To predict the impossibility of the relevant combinations, the constraint against combining 1st and 2nd with dative is ranked above WEAK, the relevant constraints violated in the other clusters are ranked below.

Table 19: Realization of input by a strong pronoun

input: [1 + weak dat]	MARK1& DAT	WEAK
a. 1 + dat	*!	
b. 1 + strong		*

4.3 The Gradience of Alignment Constraints

Most of the clitic sequences discussed so far have contained only two clitics. When this is the case, it is impossible to distinguish between two interpretations of the violation of the constraints: that the violations may be absolute – a clitic not at the edge violates the constraint exactly once no matter how many elements separate it from the edge, or they may be gradient - a clitic not at the edge violates the constraint once for each element that separates it from the edge.

In a sequence of two, a mis-aligned element will violate the relevant alignment constraint once under either interpretation, however in larger sequences the distinction is crucial. The decision between gradient and absolute violation requires sequences of at least three clitics, returning us to the question addressed in Section 4.2, of which potential combinations of clitics can in fact occur. Combinations involving highly marked clitics are generally expected to be impossible, and combinations of three clitics are of course more likely to contain marked clitics than combinations of two. It is not clear that the data available from the various languages is strictly comparable: whether the same combinations have been examined, the same standards of well-formedness applied and the same range of interpretations reviewed (see for example, footnote 6 in Wanner 1977).

Perlmutter cites grammatical three-clitic combinations (pp 50-51, although see note 28.), such as the one analyzed in Table 15 above. As noted previously, the ungrammaticality of candidate c. and the grammaticality of candidate b. shows that the constraints are gradiently violable. They must be *gradient* constraints since the violation of 1RT in c. is worse than that in b. only if the separation of the 1st person clitic from the right edge by two elements is worse than the separation of the 1st person clitic and the right edge by only one element. Under an absolute interpretation of the violations, both candidates would violate 1RT exactly once, by virtue of containing one element with a 1st person specification and not on the right edge.

This case illustrates what I take to be the general situation: in so far as combinations of 3 clitics are allowed in the first place (whether personal clitics, partitives or locatives) their positioning shows gradience. The general pattern predicted by gradience is for a clitic to appear as close as possible to the targeted edge, when it cannot be the absolute edgemost item.

The gradience of the alignment constraints establishes the patterns of constraint violation that they are responsible for. As an example, we can look at the constraints PERSLFT, PERSRT. In (22), violations of PERSLFT are shown as subscripts to the left of P, where P is a person marked clitic; violations of PERSRT are shown as subscripts to the right of P. ("X" and "Z" have no person specification.) What (22) shows is that for an input of 4 clitics, one of which has a person specification and is thus targeted by the constraints, the total number of violations of PERSLFT and PERSRT added together is constant: there are always a total of 3 violations.

(22) ${}_0P_3$ X Y Z = 3

X ${}_1P_2$ Y Z = 3

X Y ${}_2P_1$ Z = 3

X Y Z ${}_3P_0$ = 3

The total number of violations is constant because of a particular property of matched alignment constraints with opposite edge specifications: for each clitic with the targeted property elimination of a violation of one always adds a violation of the other. For example, as a clitic with a targeted specification moves rightward (as P does in (22)), the number of XRT violations it incurs decreases and the number of XLFT violations it incurs increases, at exactly the same rate.

The patterns of violation in (23) make the same point for a sequence of four clitics, three of them person marked. Here the violations *for each person-marked clitic* always total 3, and the total number of violations in *the sequence* is always 9.

(23) ${}_0P_3$ ${}_1P_2$ ${}_2P_1$ ${}_0X_0$ = 9

${}_0P_3$ ${}_1P_2$ ${}_0X_0$ ${}_3P_0$ = 9

${}_0P_3$ ${}_0X_0$ ${}_2P_1$ ${}_3P_0$ = 9

$_0X_0$ $_1P_2$ $_2P_1$ $_3P_0$ $= 9$

Where *n* is the total number of clitics in a sequence, there are n-1 violations of a matched opposite alignment pair *for each clitic* with the targeted property. E.g. each clitic specified for person in (23) incurs 3 violations of the matched opposite alignment pair [PERSLFT, PERSRT], because the total number of clitics in this sequence is 4. For *the entire sequence*, the total number of violations of the matched opposite alignment pair is t(n-1), where *t* is the total number of clitics with the targeted property. In (23), then, the total number of violations per constraint pair is 9, i.e. 3 x (4-1).

When the choice is between candidates with the same items, which do not differ in structure in any way, the lower ranked constraint of each pair can never make a decision, i.e. will be invisible in the grammar. Any candidates which are still competing by the time they are assessed by the lower ranked of the pair, must by definition have tied on all higher ranked constraints, including the higher ranked member of the pair. But if they tied on the higher ranked of the pair they will tie on the lower ranked of the pair also. Note that the invisibility of the lower ranked constraint of each pair is not arbitrary: it maintains the absolutely strongest version of the OT universality hypothesis: all constraints are in the grammar of every language. What distinguishes gradient alignment constraints from other types of constraints is that for each pair of constraints, a candidate which incurs n fewer violations on one member of the pair will always incur n more violations on the other member.

The only ranking that can ever be crucial, given this argument, is the ranking of the higher member of a matched opposite alignment pair, since the lower member can never be decisive in determining optimality, and hence cannot be crucially ranked with respect to any other constraint.

5. Markedness and the Derivation of the Clitic Inventory

At several places in the analysis, the nature of the clitic lexicon plays a critical role. The inventory of clitics, i.e. the set of lexical representations which exist in these systems, can itself be derived from ranked markedness and faithfulness constraints; that is, such a ranking predicts the set of morpho-syntactic combinations that are present in, or absent from, these systems. To mention a couple of prominent examples: there are no 1st or 2nd person reflexive forms, there are no 1st or 2nd person case-marked forms, there is no reflexive clitic with gender, case or number specifications.

Of course, the grammar (ie.. ranked constraints) cannot predict the phonological form of the clitics, i.e. the fact that a clitic is *la* rather than *pi*. The grammar can, however, predict that a particular combination of morphological specifications will exist in the lexicon. I assume that in the unmarked case, each distinct cluster of morpho-syntactic specifications is associated with a distinct phonological form.[14]

The derivation of the clitic inventory is very much in the spirit of the derivation of phonemic inventories as proposed in Prince and Smolensky 1993, and pursues the goal suggested in Grimshaw 1997a of deriving the inventory of functional heads in individual grammars from ranked grammatical constraints.

The basic idea is that universal grammar specifies a set of morpho-syntactic features which can be encoded in the pronominal systems of natural languages. Clusters of these features constitute the inputs to the constraints which determine the pronominal system. (The analysis presented here assumes that the input always includes a specification for every feature: modification is required to encompass the more complex situation where inputs can be partially specified.) The ranking of markedness constraints (MARKX), which penalize a pronoun for encoding a specification, and faithfulness constraints (FAITHX), which reward a pronoun for encoding input specifications, determine which combinations of specifications will occur in output forms in a given language. These constraint rankings jointly determine the inventory.

We thus posit FAITH and MARK constraints for all specifications allowed by Universal Grammar. (I assume that there are FAITH and MARK constraints both for the specification and its values. For instance, one constraint in each family targets any number specification, one targets singular and one targets plural. I will discuss the ranking of the general constraint unless it is the more particular constraint that is important.) For those specifications that play no role in distinguishing personal clitics from one another (but which can be active in pronominal systems in general) such as the +/- human distinction, we conclude that the markedness constraint MARKHUMAN dominates the faithfulness constraint FAITHHUMAN, and thus that all inputs with human specifications map to outputs which lack such specifications. (This assumes that no other constraint(s) favor preservation of the +/- human specification.) In cases such as this, where a distinction is irrelevant throughout the clitic system, I omit it from lexical representations, and from further consideration here, altogether.

In contrast, for those specifications which are contrastive in the clitic systems because they are preserved in some output forms, we know that the FAITH constraint must dominate the MARK constraint, The FAITH constraints for reflexivity, person, number, gender and case must all dominate the corresponding MARK constraints.

(24) FAITHREFL >> MARK REFL, FAITHPERS >> MARK PERS,
FAITHNUM >> MARK NUM, FAITHCASE >> MARK CASE,
FAITHGEN >> MARK GEN

The interesting challenge posed by the Romance clitic systems is that they are not uniform: that is one cannot say for each specification that it is always, or never, preserved in the output. This can be seen from re-examination of (4) above. Gender, for example, is preserved in the third persons but not elsewhere, and the same is true for case and reflexivity. This shows that the interaction of simple markedness and faithfulness constraints alone cannot determine the clitic inventory. Such interactions can only generate systems in which a given feature (or value) is always preserved in the output (FAITH >> MARK) or systems in which a given feature (or value) is never preserved in the output (MARK >> FAITH). To characterize a mixed system will require constraints which assess the markedness not just of single features or values, but combinations of features or values (cf. Section 4.2.) .These constraints will be referred to as "combination constraints".[15] Every case where the outline notation has been used in a lexical representation is a case of this kind: the outlined feature, while active in the Romance clitic systems in general, has not been preserved in a particular output. This occurs when a combination markedness constraint conflicts with a faithfulness constraint, and the combination constraint is dominant.

There are three possible types of clitics in the Romance personal clitic lexicon. First, there is a reflexive clitic, in which reflexivity does not co-occur with any other specification. Second, there are 1st and 2nd person clitics, which are specified for number but not for reflexivity, gender or case. Third, there are non-reflexive 3rd persons, in which all features are preserved. I illustrate this case first, in Table 20.

The optimal output here violates the combination markedness constraints for all specifications combined with 3rd persons. The combination constraint can only be satisfied by eliminating the person specification as in c., or eliminating all of the other specifications, as in candidate b. The ranking of the faithfulness constraints rules out either of these options: no matter which specification is omitted (leading to improvement on the combination constraint), a violation is incurred, so that such a candidate cannot be optimal. For 3rd person inputs, then, all specifications are parsed.[16]

Table 20: All specifications parsed in 3rd person non-reflexives

input: [-R 3 sg m acc]	FAITH REFL,PERS,NUM, GEN,CASE	MARK 3&REFL,3&NUM, 3&GEN, 3&CASE
a.☞ -R 3 sg m acc		****
b. ℝ 3 ℕ 𝔾 ℂ	*!***	
c. -R ℙ sg m acc	*!	

We conclude that the combination markedness constraints which target combinations involving 3rd person are all dominated by the conflicting faithfulness constraints.

(25) FAITH REFL/PERS/NUM/GEN/CASE >> MARK 3 & REFL/NUM/GEN/CASE

For 1st and 2nd person inputs, Table 21 shows how person and number are preserved (the tableau shows 1st person but the same basic analysis holds for 2nd persons). The FAITH constraint for 1st and 2nd person, and the combination markedness constraint which bans the combination of 1st and 2nd person with reflexive, gender and case, must both dominate faithfulness for reflexive, gender and case. The conflict will then be resolved by satisfying FAITH1ST/2ND and the combination constraint, eliminating reflexive, gender and case. Unlike the other specifications, number is always preserved in 1st and 2nd persons, so FAITH1ST/2ND and FAITHNUM must both dominate the combination constraint which bans them together.

(26) FAITH1ST/2ND, MARK1ST/2ND & REFL/GEN/CASE >> PARSEREFL/GEN/CASE

FAITH1ST/ 2ND, FAITHNUM >> MARK 1ST/2ND & NUM

Table 21: 1st and 2nd persons preserve person and number

input: [+R 1 sg m acc]	FAITH 1ST/,2ND	FAITH NUM	MARK 1ST/,2ND &NUM	MARK 1ST /2ND &REFL/ GEN/ CASE	FAITH REFL/ GEN/ CASE
a. +R 1 sg m acc			*	*!**	
b.☞ ℝ 1 sg 𝔾 ℂ			*		***
c. ℝ ℙ sg 𝔾 ℂ	*!				***
d. ℝ 1 ℕ 𝔾 ℂ		*!			***

In sum, for any 1st or 2nd person input, the person and number are parsed and the remaining specifications are not. That this is true for non-reflexive inputs as well as reflexives can be seen from re-examining Table 21. The tableau is exactly the same if the input is changed to being -R, as it is if the input is changed to being plural, or dative etc.

The remaining case is that of a 3rd person reflexive input. The lexicon in (4) represents the output as a clitic which is +R but loses all other specifications. In the present terms, the combination markedness constraint which targets +R in combination with others, MARK+REFL&PERS/NUM/GEN/CASE, must dominate the constraints which enforce parsing of person, number, gender and case. These constraints musts also be dominated by FAITHREFL, in order to ensure that the reflexive specification is preserved, and the others lost, due to the conflict with the combination markedness constraint. Table 22 illustrates the analysis.

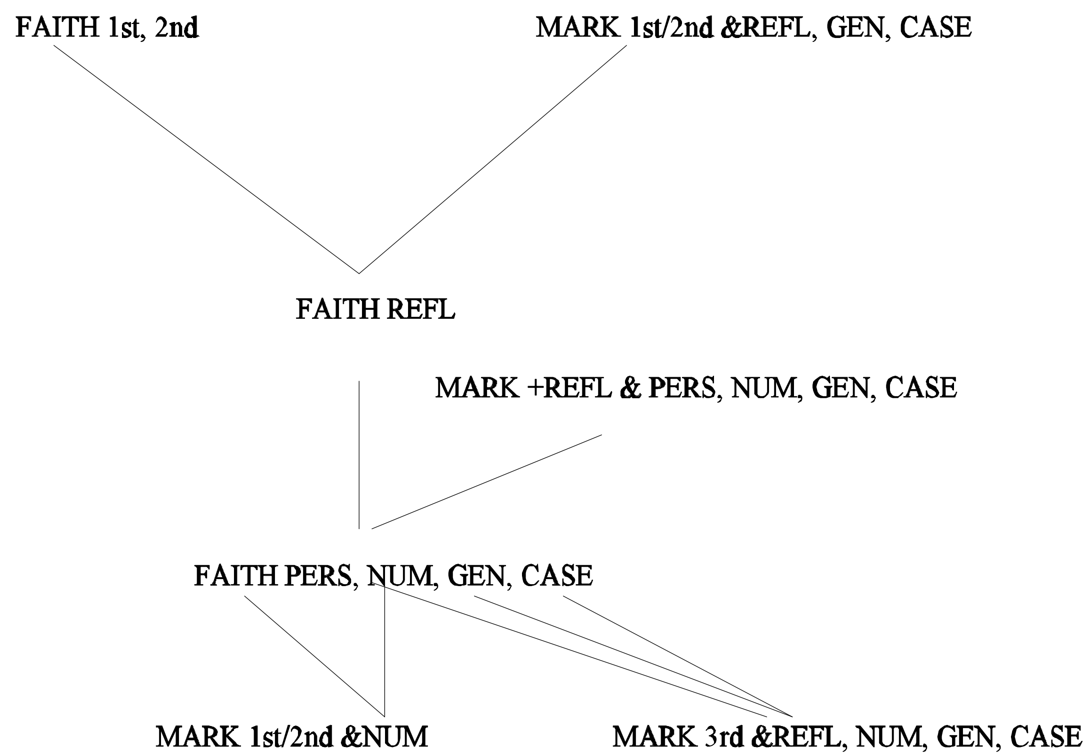
(27) FAITHREFL, MARK +REFL& PERS/NUM/GEN/CASE >> FAITH PERS/NUM/GEN/CASE

Table 22: 3rd person reflexives preserve just reflexivity

input: [+R 3 sg m acc]	FAITH REFL	MARK +REFL& PERS/NUM/ GEN/CASE	FAITH PERS/NUM/ GEN/CASE
a. ☞ +R ℙ ℕ 𝔾 ℂ			****
b. +R 3 sg m acc		*!***	
c. ℝ 3 sg m acc	*!		

With this ranking, the +R specification will be preserved, and the others eliminated, resulting in satisfaction of the combination markedness constraint in the optimal candidate.

Figure 3: Ranking Diagram for Faithfulness and Markedness Constraints



By appeal, then, to markedness constraints against combinations of properties encoded in a single clitic, we can derive the inventory of clitic pronouns used in the languages under study.[17] A clitic specification such as "+R 2 pl 𝔾 ℂ" or "-R 3 sg 𝔾 ℂ" is simply impossible, since the constraints will never select it as optimal for any input.

Note that there is a difference between the outputs in the case of 1^{st} and 2^{nd} person and the case of the reflexive clitic. In the first case both person and number are preserved in the output specification, because the combination constraint involving $1^{st}/2^{nd}$ and number is dominated by FAITH NUM., as well as by FAITH$1^{ST}/2^{ND}$. In the second case only reflexivity is preserved, because only FAITHREFL dominates the combination markedness constraint targeting +Refl: it dominates all of the other faithfulness constraints.

It is worth noting, and suggestive of a course for further research, that the rankings in Figure 3 are broadly consistent with recent work on markedness in morpho-syntactic systems reported in Aissen (this volume) and Artstein (1998). (The same is true of the rankings sketched in Section 4.2.) Note that the rankings in Figure 3 show that the combination markedness constraints of the form: MARK 1ST/2ND&X dominate the combination markedness constraint MARK 3RD &X.

Following a line of research developed in Prince and Smolensky 1993, Smolensky 1995, Aissen this volume, Artstein 1998, the ranking of a markedness constraint against combinations is partly determined by the ranking of the individual markedness constraints. MARK1ST/2ND is univerally ranked higher than markedness constraints for 3rd person: this is what guarantees that 1st and 2nd persons are more marked than 3rd person. I assume also that reflexives are more marked than non-reflexives, and that dative is more marked than accusative (this is important in Section 4.2, see Grimshaw 1997b, and Woolford this volume for further evidence.)

These markedness generalizations follow from the rankings in (28):

(28) MARK 1ST /2ND >> MARK 3RD

MARK DAT >> MARK ACC

MARK +REFL >> MARK -REFL

It is no accident that, for example, 1st and 2nd person clitics, which are marked relative to 3rd, do not support further specifications for reflexivity, gender or case, while the unmarked 3rd person clitics do.

The fundamental prediction of such proposals, transported to the domain of inventory derivation, is that more contrasts will be maintained in less marked forms, because the constraint which bans a marked specification combined with any other will necessarily be ranked above a constraint which bans a less marked specification combined with the same additional property. Thus those properties that can combine in lexical representations of Romance clitics (and those that can combine in clitic combinations as sketched in Section 4.2) are less marked than those that are not allowed to combine.

This general line of explanation is extensible to quite arbitrary-looking problems, suggesting that they are not as arbitrary as they may appear. Note 2 alludes to the gap in Latin American dialects, where there is no 2nd person plural clitic, a 3rd person being used instead. Table 23 shows that this state of affairs is derived if the relevant combination constraint dominates faithfulness to person.. (Note that the FILL version of faithfulness is violated here.)

Table 23: Absence of a 2nd person form

input: [-R 2 pl m acc]	MARK 2ND&PL	FAITH PERS	FAITH REFL/GEN/CASE
a.☞ -R 3 pl m acc		*	
b. ℝ 2 pl 𝔾 ℂ	*!		***

The candidate in a. is the Latin American optimum, the candidate in b. is the winner in *os* dialects. Even the dialect which does not tolerate a 2nd person plural clitic does can preserve a 1st person plural clitic, and a 2nd person singular clitic, because MARK1ST&PL and MARK2ND&SG are both ranked below faithfulness to person. This is again entirely consistent with a general markedness hierarchy in which 2nd person is more marked than 1st (see Grimshaw 1997b for more evidence for this conclusion from Romance clitics), and plural is more marked than singular; hence the combination markedness constraints against 1st person plurals and 2nd person singulars must necessarily be ranked below the combination markedness constraint against 2nd person plurals.

If a marked form, such as 2nd person plural or dative plural, is missing from the lexicon, then there is no prediction concerning the less marked counterparts, 2nd person or dative singular and 1st/3rd person or accusative plural. But if the less marked forms are impossible, then their more marked counterparts must be ruled out, since the constraint ranking which eliminates the less marked combination will eliminate the more marked also.

In this way, the clitic lexicons can be generated by faithfulness constraints, together with two kinds of markedness constraints: simple constraints against feature specifications and markedness constraints against combinations of feature specifications, both ranked in accordance with the markedness hierarchies of Universal Grammar.

The inventory derivation laid out here is, as far as I can see, inconsistent with a theory which either enforces under-specification whenever it is possible or forbids it except when it is absolutely necessary. Instead, whether a property is specified or not is a function of the ranking of the constraints which favor or disfavor it. This means that one cannot simply study the paradigm in Table 1 and determine the lexical representation of the clitics. The data for clitics in isolation is perfectly consistent, for example, with the hypothesis that *se/si* has no specifications at all (indeed that was the analysis given in Grimshaw 1997b.) However, in order to predict the behavior of this clitic in combinations, and to predict the nonexistence of clitics like *se/si* with case specifications, it is necessary to analyze *se/si* as reflexive, which the constraint ranking does.

6 Conclusion

The major points of this paper concern the role of alignment constraints in determining the relative positions of clitics in the three Romance systems studied here, and the role of markedness and faithfulness constraints in determining the lexical representations which form the clitic lexicons.

The order of clitics is governed by alignment constraints, and is thus subject to the same fundamental principles as any other syntactic ordering, including relative order among heads, specifiers and complements (Grimshaw 1997a).

The inventory of clitics in each system is determined by the interaction of markedness constraints, which target combinations of morpho-syntactic specifications, and faithfulness constraints which mandate preservation of input specifications. The derivation of the inventory is a step within the larger program sketched in Grimshaw 1997, in which the functional lexicons of languages are all derived from the interaction of grammatical constraints. In Grimshaw 1997, this logic was applied analyzing the distribution of the complementizer *that*, and the auxiliary verb *do*. In Grimshaw and Samek-Lodovici it was applied to the English expletive *it*.

These two threads in the proposal made here, together support the view that pronominal clitics in Romance are exemplary of the principles governing functional heads: their inventory is derived by ranked violable constraints, and their position is determined by syntactic alignment.

Bibliography.

Aissen, Judith. This volume. Markedness and Subject Choice in Optimality Theory.

Anderson, Stephen. 1992. *A-Morphous Morphology*. Cambridge University Press.

Anderson, Stephen. 1996. Rules and Constraints in Describing the Morphology of Phrases, in Dainora, A., R Hemphill, B. Luka, B. Need and S. Pargman eds. *The Parasession on Clitics*, Papers from the 31st Regional Meeting of the Chicago Linguistic Society, Volume 2, 1995.eds.15 -31.

Anderson, Stephen. In press. Toward an Optimal Account of Second Position Phenomena. In Dekkers et al. eds.

Artstein, Ron. 1998. Hierarchies. Ms., Department of Linguistics, Rutgers University. Available at http://www.eden.rutgers.edu/~artstein/publications.html.

Barbosa, Pilar, Danny Fox, Paul Hagstrom, Martha McGinnis and David Pesetsky eds. 1998. *Is the Best Good Enough? Optimality and Competition in Syntax.* MIT Press.

Bonet, Eulàlia. 1994. “The Person Case Constraint: a Morphological Approach. *MIT Working Papers in Linguistics* 22. 33-52.

Bonet, Eulàlia, 1995. Feature Structure of Romance Clitics. *Natural Language and Linguistic Theory* 13, 607-647.

Cardinaletti, Anna. 1991. On Pronoun Movement: the Italian Dative Loro. *Probus* 3. 127-153.

Dekkers, Joost, Frank van der Leeuw and Jeroen van der Weijer eds. In Press. *Optimality Theory: Syntax, Phonology and Acquisition*. Oxford: Oxford University Press.

Gerlach, Birgit. 1998. Restrictions on Clitic Sequences and Conditions on the Occurrence of Clitics in Romance. Theorie des Lexikons, Arbeiten Des Sonderforschungsbereichs 282, Universität Düsseldorf, Nr. 105. Available as ROA-277-0898, Rutgers University.

Grimshaw, Jane. 1997a. Projection, Heads and Optimality. *Linguistic Inquiry* 28. 373-422

Grimshaw, Jane, 1997b. The Best Clitic: constraint interaction in morphosyntax. In L. Haegeman ed. *Elements of Grammar*: *A Handbook in Contemporary Syntactic Theory*, Kluwer. 169-196

Grimshaw, Jane and Vieri Samek-Lodovici.1998. Optimal Subjects and Subject Universals. In P. Barbosa et al. eds. 193-219.

Haverkort, Marco 1993. *Clitics and Parametrization: Case Studies in the Interaction of Head Movement Phenomena.* Program in Language Typology. European Science Foundation.

Heap, David. 1996. Optimalizing Iberian Clitic Sequences. In *Theoretical Analyses on Romance Languages*, José Lema and Esthela Treviňo eds. Amsterdam/Philadelphia, John Benjamins.

Kayne, Richard. 1991. Romance Clitics, Verb Movement and PRO. *Linguistic Inquiry* 22, 647-686.

Legendre, Géraldine 1996. Clitics, Verb (Non)-Movement, and Optimality in Bulgarian. Johns Hopkins University Technical Report JHU-CogSci-96-5.

Legendre, Géraldine. In press. Morphological and prosodic alignment of Bulgarian clitics. In: Joost Dekkers et al. eds.

Legendre, Géraldine. This volume a. An Introduction to Optimality Theory in Syntax.

Legendre, Géraldine. This volume b. Masked V2 Effects and the Linearization of Functional Features.

Legendre, Géraldine, Paul Smolensky and Colin Wilson.1998. When is Less More? Faithfulness and Minimal Links in *wh*-Chains. In Barbosa et al. eds. 249-289.

McCarthy, John and Alan Prince. 1993. Generalized Alignment. In *Yearbook of Morphology 1993*, ed. Geert Booij and Jaap van Marle, 79-153. Dordrecht: Kluwer.

McCarthy, John and Alan Prince. 1995. Faithfulness and Reduplicative Identity. Ms. University of Massachusetts at Amherst and Rutgers University. ROA # 60.

McCarthy, John and Alan Prince. 1999. Faithfulness and Identity in Prosodic Morphology. In René Kager, H. v. d. Hulst, and W. Zonneveld eds., *The prosody-morphology interface*, 218-309. Cambridge University Press.

Miller, Philip H. and Ivan Sag. 1997. French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15, 573-639.

Monachesi, Paola. 1995 *A grammar of Italian Clitics*, ITK Dissertation Series/TILDIL Dissertation Series.

Morin, Y.-C. 1979 More Remarks on French Clitic Order. *Linguistic Analysis* 5. 293-312.

Perlmutter, David, 1971. *Deep and Surface Structure Constraints in Syntax*, Holt, Rinehart and Winston.

Prince, Alan and Paul Smolensky, 1993. *OptimalityTheory: Constraint interaction in generative grammar*. RuCCS Technical Report #2. Rutgers University Center for Cognitive Science, Piscataway NJ.

Samek-Lodovici, Vieri. 1998. Opposite constraints: Left and right focus-alignment in Kanakuru. *Lingua* 104, 111-130.

Simpson, J. and M. Withgott. 1986. Pronominal Clitic Clusters and Templates. Syntax and Semantics Volume 19: *The Syntax of Pronominal Clitics*. Academic Press.

Smolensky, Paul. 1995. On the Internal Structure of the Constraint Component Con of UG. Handout of talk presented at UCLA. Available as ROA-86-0000, Rutgers Optimality Archive.

Terzi, Arhonto. 1999. Clitic Combinations, their Hosts and their Ordering. *Natural Language and Linguistic Theory* 17. 85-121

Uriagereka, Juan. 1995. Aspects of the Syntax of Clitic Placement in Western Romance.

Linguistic Inquiry 26. 79-123.

Van der Leeuw, Frank. 1995. Alignment and Integrity Constraints in Cliticization. Chicago Linguistic Society 31, Volume 2. *The Parassession on Clitics,*168-197.

Wanner, Dieter, 1977. “On the Order of Clitics in Italian”, *Lingua* 43, 101-128.

Woolford, Ellen. This volume. Case Patterns.

Notes

1. The research reported here was supported by grant SBR-9511891 from the National Science Foundation to Rutgers University. An earlier version of the alignment constraint analysis was presented at the Hopkins Optimality Workshop, the University of Rochester, Stanford University, the University of Texas, Harvard University, the “Lexicon in Focus” workshop at the Universität Düsseldorf, the “Workshop on Clitics” at the Deutsche Gesellschaft fuer Sprachwissenschaft Konstanz., the workshop on clitics at the Linguistic Society of America Linguistics Institute, University of Illinois at Urbana-Champaign. Questions, suggestions and comments from the audiences have sharpened the analysis presented here considerably. Many individuals have contributed to this project: especially Ron Artstein, Eric Bakovic, Edward Keer, Nicole Nelson, Alan Prince, Vieri Samek-Lodovici, Bruce Tesar, and Colin Wilson. I would like to thank Géraldine Legendre and Sten Vikner for comments on earlier versions of this paper.

2. See Uriagareka 1995 and Terzi 1999 for recent discussion of cliticization in non-finite contexts.

3. This clitic does not occur in dialects of Latin America, where a third person replaces it. See Section 5 for an explanation.

4. See Cardinaletti 1991 for an analysis of Italian *loro*.

5. *Ci* in the second column of (7) is a locative clitic. Some Italian clitics have a different phonological shape when they precede a clitic. I will not analyze this effect here.

6. A minor variation of this analysis posits ACCRT rather than CASERT as the dominant constraint. This requires modification of the analysis of spurious *se* (see Section 3.4), with DATRT now the constraint which forces replacement of the dative clitic by *se*.

7. The constraint proposed in Heap 1996 “Least Leafy Last” is, in this view, not a constraint at all, but a description of the state of affairs resulting from the particular constraint ranking of Spanish, which has the effect of positioning clitics with more person and case specifications to the right of clitics with less.

8. Either PERSRT or 2RT must dominate 2LFT, in order to guarantee that the 2nd person clitic follows *se*. This is consistent with, but not included in, the constraint rankings presented here.

9. It also follows if 1LFT >> 1 RT, provided that 1LFT >> PERSLFT also.

10. In an earlier analysis (Grimshaw 1997b) I suggested that the clitic combination violated the Obligatory Contour Principle, despite the fact that the clitics are not completely identical. The present alternative makes such an appeal unnecessary.

11. In discussing spurious *se*, I am assuming that the replacement clitic has to be one drawn from the generally available inventory of Spanish clitics. Thus *le* is replaced by *se*. But why couldn't there be an additional clitic which parses every input specification but case, or at least parses more of the input than *se* does? It would satisfy CASERT vacuously, and be more faithful to the input than *se*. This clitic would occur just in spurious *se* configurations. The constraint system laid out in Section 5 precludes the existence of such a clitic. The reasoning is as follows. The clitic can improve over *se* only on the PARSE constraints which are ranked below PARSEREFL, and the combination markedness constraint (see (27)).
If the hypothetical clitic parses the +R in the input it will fatally violate the combination constraint banning +R together with whatever additional feature in the input hte hypothetical clitic parses. If it does not parse the +R in the input it will fatally violate the PARSEREFL constraint. Hence no such specialized clitic can survive given the rankings that derive the clitic inventory.

12. Alternative structures can be analyzed in a manner which parallels that developed for Figure 2. For example, if the clitcs adjoin to I', then alignment constraints relating them the right and left edges of IP will still give the right orders, with the leftmost clitic separated from the left edge of IP by Spec, and the rightmost clitic separated from the right edge of IP by I and VP.

13. This proposal does raise the possibility that alignment could be (better) satisfied if the subject were dropped, i.e. that the presence of clitics could force omission of the subject. However, if deletion in syntax may occurs only where there is a specific constraint requiring omission, such as DROPTOPIC of Grimshaw and Samek-Lodovici 1998, this problem will not arise.

14. As can be seen in (1) and (2), there are minor differences between the three lexicons, e.g. French has *les* for both genders in the 3rd person plural accusative, while Italian has *li* and *le*. At least for now, I take it that these represent nothing more than arbitrary differences in the mappings between morpho-syntactic representations and phonological forms, so that *les*, for example, accidentally corresponds to two different morpho-syntactic clusters. It is the existence of the clusters themselves that is predictable from the constraint system presented here.

15. This is closely related to Smolensky's 1995 proposal to allow constraint conjunction: see the discussion in Legendre (in press, this volume a.), Legendre et al. 1998, Aissen this volume, Choi this volume. The constraints suggested here ban combinations of properties, and may not involve conjoining constraints.

16. There is some evidence that masculine gender and singular number should not be specified even in these forms, based on agreement phenomena (see Grimshaw 1997b for a constraint-based account.)

17. Other combination constraints such as MARK GENDER & CASE, and the general version of the combination constraints targeting person specifications (MARK PERSON & X), must all be ranked below any faithfulness constraints which they conflict with, and are not shown in Figure 3. Also a crucial part of the system are the constraint rankings established in Section 2, which guarantee the choice of the appropriate morpho-syntactic specification for an input where there is no perfect output to be selected, and hence play a critical role in determining the inventory. In a suitably refined version, e.g. one which distinguishes 1st and 2nd from 3rd person, these rankings, which involve faithfulness constraints, are consistent with Figure 3 but are not shown there.