

The Syntax of Italian Clitics

Louis-Harry Desouvrey

First draft¹

January 2018

In this paper, a syntactic account of Italian clitics is presented. It is suggested that clitics in Italian are different from other major Romance languages because of their specification for two pervasive syntactic features, namely $[\pi]$ and $[\omega]$. All clitics but *ne* carry the feature $[\pi]$, whereas the feature $[\omega]$ is present in third person accusative-specified clitics as well as past participles. Due to the feature $[\pi]$, simple clitics must undergo an unusual long outbound movement under OCP. However, the most remarkable combined effect of these feature is to be found in clusters of object clitics, which must be processed by a compounding rule (CCR), and then licensed by a phonological sandhi rule (ICIS), in order to avoid an overload derivation ($DEN > 3$). The feature $[\omega]$ has a further effect in the grammar: it forces the agreement of past participles and direct object clitics. Various cooccurrence restrictions, traditionally considered to be arbitrary or idiosyncratic, are accounted for with simplicity and elegance.

Keywords: syntactic features, clitics, clitic clusters, derivation equivalence number, vector, scalar, thematic role, compounding, Italian clitic internal sandhi, OCP, pronoun, anaphor, lexical phonology, syntax.

1. Introduction

The paradigm of Italian personal pronouns is perhaps the most complex of all Romance languages. It displays a rare clutter of morphemes and features. As a result, clitic clusters show a variety of combinations which include numerous opaque forms, those cases where a non-expected element shows up in the cluster and does not even have an independent existence elsewhere. Traditional theories, based on parameters and stipulations on abstract phrase structures, or morphological template (e.g., Monachesi 1998, Pescarini 2010, Seuren 2009, among many others), are absolutely unarmed to tackle such a conundrum. In this paper, I will show that, although the paradigm is messy, it shows a regular

¹ Comments are welcome and will be acknowledged in the next version of this paper.

pattern with respect to the tenets of the strictly feature-based syntax advocated in Desouvrey (2000) and subsequent works. After a presentation of the paradigm and some theoretical tools in the next section, I shall identify the source of all of the complications in Italian clitic clusters, namely the impossible licensing of the direct object clitic (§3), which triggers an overload problem (§4). In section 5, I shall take a fresh look on the so-called Person-Case Constraint, showing that it is a ban on two thematic features in the syntactic structure. Section 6 addresses various vector effects including the past participle agreement. Along the way, various co-occurrence restrictions as well as the ordering of the clitics within clusters will be shown to follow from the proposed analysis. Finally section 7 concludes the paper.

2. Feature inventory and theoretical tools

In order to process the feature contents of Italian personal pronouns, it is necessary to take a descriptive overview of the paradigm. Classifying Italian personal pronouns according to their grammatical functions, one obtains the following Table, which can be found in normative grammars, particularly those designed for foreign learners.

Table 1

Person / gender	Subject	Direct object	Indirect object	Reflexive	Tonic
1	io	mi	mi	mi	me
2	tu	ti	ti	ti	te
3 fem.	lei	la	le	si	lei
3 masc.	lui	lo	gli	si	lui
1	noi	ci	ci	ci	noi
2	voi	vi	vi	vi	voi
3 fem.	loro	le	loro / gli	si	loro
3 masc.	loro	li	loro / gli	si	loro
locative			ci / vi		
partitive		ne			

This Table contains information on person, gender, grammatical functions, as well as reflexive and tonic forms. A striking aspect of the paradigm is the fact that it includes many cases of homophony, that

is, different morphemes have the same phonetic shape. Thus, the pronouns *ci* et *vi* are not only used as first and second plural, but also as locative. Notice also that the pronoun *si* may appear in various constructions referred to as impersonal, reflexive and *passivante*, which points to two different morphemes (see below).

The elements in the subject and the object columns, except the first and second singular, are similar. I take them to be unmarked for case, which allows them to be used as either subject, object of a verb or a preposition. Being deprived of case features, they are not clitics and I ignore them throughout. Similarly, I set aside the elements *io* and *tu*, first and second singular respectively, which presumably play no significant role in the system, since Italian is a null subject language.²

First and second person elements have a single form which is used as direct object, indirect object and reflexive. Unlike the tonic forms, they behave like clitics in that they cannot be the object of a preposition and nor can they be used in isolation. I take them to be specified for both Accusative (direct object) and Oblique (indirect object). The third person singular *la* (feminine) and *lo* (masculine) are the only clitics that do not have cognates elsewhere in the paradigm. As a matter of fact, they can only be a direct object bearing a single case, namely Accusative. However, although the clitic *gli* is found only in the indirect object column, it behaves just like first and second person clitics; it must be therefore specified in the same way (see footnote 9). As for the clitic *le*, it shows up in both direct and indirect object columns, which means it is doubly specified for case, Accusative and Oblique, just like first and second person clitics. The reflexive clitic *si* may be a direct object or an indirect object, which leads us to the assumption that it is specified for both Oblique and Accusative. I take it to be intrinsically distinct from the impersonal *si*, which can only be used as third person subject of a singular verb. Finally, the paradigm contains a partitive clitic *ne* and two locative clitics *ci* and *vi*. Based on its use with direct object, one can assume that *ne* is an accusative clitic. As for *ci* and *vi*, they are both locative, but only the latter has a thematic feature, [λ]. In fact, *vi* has a doublet, as I will show, a second locative *ci*, which is specified alike.

Setting aside tonic and subject pronouns, the clitics, i.e., the case-bearing elements, can be handled according to their number and person features, as shown in the following Table. The letters A, N and O stand for Accusative, Nominative and Oblique respectively. The letters δ and λ stand for the thematic features Dative and Locative respectively; they play a crucial role in the referential system.³

2 What features differentiate *io/tu* from *me/te* is not clear at this point due to lack of data.

3 In this paper, I do not intend to present a full feature structure tree of the paradigm, like the one discussed in Desouvrey

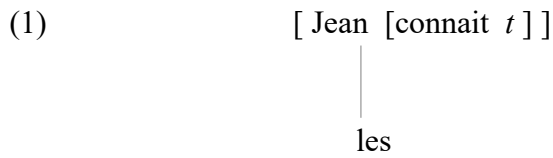
Table 2

	Singular	plural
1	mi [A O δ π]	ci [A δ O π]
2	ti [A O δ π]	vi [A δ O π]
3 fem.	la [A π ω], le [A O δ π ω]	gli [δ O π]
3 masc.	lo [A π ω] gli [A O δ π ω]	li [A π ω] gli [A O δ π ω]
3 reflexive	si [A O δ π ω]	
3 impersonal	si [N φ]	
Others	ne [A] ci [O π] / ci [A O λ π] vi [A λ O π]	

Italian clitics are also specified for two further features, namely $[\pi]$ and $[\omega]$. As can be seen in Table 2, every clitics except *ne* are π -specified, while third person direct object clitics are in addition ω -specified. These features are not specific to Italian neither to Romance languages (cf. Desouvrey 2000, 2005, etc.). Given that in the input to a derivation the morphemes are linearly ordered so as to make a tier, the adjunction of a morpheme A to another morpheme B automatically put the former to a distinct timing tier, which is linked to the original one by an association line. This is illustrated in (1) with the derivation of a simple French sentence, which is cyclically built out of real morphemes, without functional heads and the like. The object clitic moves from its original position *t* to the right edge of the host, always the first element outside the complement domain of the verb.⁴ As can be seen, the last segment of *Jean* links up with the first segment of the clitic. In order to be spelled out, as in (3), the two-tiered structure is linearized by the following convention, (2). (The symbol '=' in (3) indicates a temporal adjunction, an adjunction to a head in a familiar parlance.)

(2000) for French, which anyway has to be reviewed in the light of more recent work on Korean and ergative languages (cf. Desouvrey 2010, 2012). However, the readers should keep in mind that the representation of syntactic features is similar to nonlinear phonological representations.

4 The letter *t* is sometimes shown in the examples for ease of exposition; there are no traces or abstract functional heads in the representation.

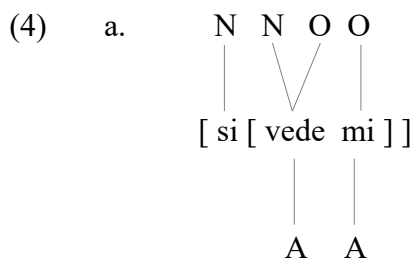


(2) **Linearization (or leveling) convention**

Embed the adjunct tier into the main tier, to the right of the host element.

(3) Jean = les connaît

The landing site of the clitic *les* is within the sentence; such a movement will be referred to as inbound, as opposed to outbound. However, π -specified morphemes cannot undergo this process; instead, they can only make an outbound movement, that is to say, they must move to the left edge of the structure, as illustrated in derivation (4), which gives us a first glimpse into the system. As can be seen, the verb is specified for three cases, each of which being paired up with the relevant case of the arguments so as to yield a plane. Under OCP-1 (cf. (5)), the object clitic must exit the minimal domain it makes up with the verb. Unlike the example above, the object must move outbound, since it is π -specified, (3b). It should be noted that impersonal *si* can stay in situ between the outbound clitic and the verb because it is specified for case, otherwise it would be forced to move leftward or rightward under OCP-2 (also see below). (If we are correct so far, every clitic but *ne* is expected to behave the same way as *mi* in (4).)



one sees me

b. *mi* [*si* [*vede* *t*]] (output)

'One sees me.'

(5) **Obligatory Contour Principle (OCP)**

- (1) Two identical case features are forbidden in the complement domain of the verb.
- (2) Case-specified arguments must be adjacent to the verb.

As for the feature $[\omega]$, it confers scope properties to the elements bearing it, which are therefore similar to typical ω -bearers, namely *wh*-elements and negation, referred to as vectors, as opposed to scalar elements (0 - or φ -specified). The most obvious effect of this feature, which can be seen when there are two or more such elements in a structure, is the strict ordering that it imposes on the elements. Scopal elements cannot commute, that is, their position within the initial domain is preserved throughout the derivation, as stated in (6). This restriction holds even when it comes to conflict with the OCP, which is then barred. A lower vector can only move outbound or to another clearly delimited domain, as explicated in (7). That is, in a configuration $[A [B [C D]]]$, where C and D are vectors, D can move past B , but not to B by head adjunction, a vacuous movement. It turns out that this feature has considerable effects in Italian grammar.

(6) **Vector (or Superiority) Effect**

Ω -specified elements cannot commute in their minimal domain.

(7) **Well-formedness condition on vector movement**

A lower vector cannot move vacuously across a boundary.

Assuming that infinitive verbs and past participles (see below) are ω -specified, (6) and (7) account for both enclitic to infinitive verbs and clitic climbing to a higher verb. This is illustrated in (8). In (8a), the clitic is in situ, but it may move past the higher verb, as in (8b). (Conveniently, features are sometimes appended to the morpheme with an underscore.)

- (8) a. [posso [comprarlo]] (input: [posso [comprare_ ω lo_ ω]]
 b. lo [posso [comprare t]]

In French, clitics do not bear this feature, and therefore neither encliticization to infinitive verbs nor clitic climbing is possible:

- (9) a. Je peux l'acheter
 I can it to buy
 'I can buy it.'
 b. *Je peux acheter le
 c. *Je le peux acheter

As mentioned above, an element bearing the omega feature is likely to interact with negation,

which is a typical vector. Thus in Portuguese, where verbs and clitics are ω -specified (Desouvrey 2008 and references therein), movement under OCP is barred by virtue of (6), as shown in (10a). However, the clitic must move to negation, (10b), without inducing a superiority effect, for the adjunction leads to the mutual disablement of their ω -feature (see below a similar effect in Italian).

- (10) a. [O Pedro [magoou_ ω o_ ω]] / *O Pedro o magoou
 Pedro hurt him.
- b. O Pedro não = o matou
 'Pedro didn't hurt him.'

To summarize, the features of the paradigm are detailed while introducing some basic tenets of the theory I build on. Along the way, a few behaviors of Italian clitics are accounted for. In the next section, I will focus on clitic clusters.

3. A licensing problem

In the present perspective, a clitic is in general a light element, monosyllabic, which carries a set of features among which case is the most prominent. It is important to keep in mind that a clitic is light because of its wealth of features not because of its number of syllables.⁵ As an argument a clitic can only be selected by a head which shows compatible features. If a verb does not have the relevant case features, it may not merge with a clitic. In other terms, a clitic must be crucially licensed by a relevant case of a verb. This is different from lexical NPs which may be licensed by the sole lexical structure of the verb. This is emphasized in (11).

- (11) **Argument Licensing**
- a) A clitic must be licensed by a verb with at least a relevant and valid case.
- b) A lexical NP is licensed by the lexical structure of the verb.

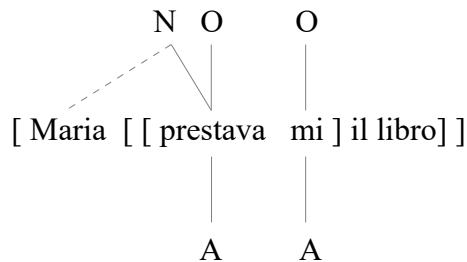
Consider the sentences in (12), which is derived from input (13), where all features except cases are set aside for convenience. As can be seen, the lexical subject *Maria* is assigned nominative case by the verb, while the first person clitic is paired up with both oblique and accusative case of the verb. Therefore, the caseless lexical direct object is not assigned a case, but it is still licensed by the lexical

⁵ Most likely an element with many abstract features need not be expressed with a long sequence of segments, vowels and consonants.

structure of the verb, which anyway needs two complements. This ruins the famous Case Filter, which requires every NP to be assigned a case.

- (12) . Maria mi prestava il libro.
 Maria me lent the book
 'Maria lent me the book.'

- (13) a.



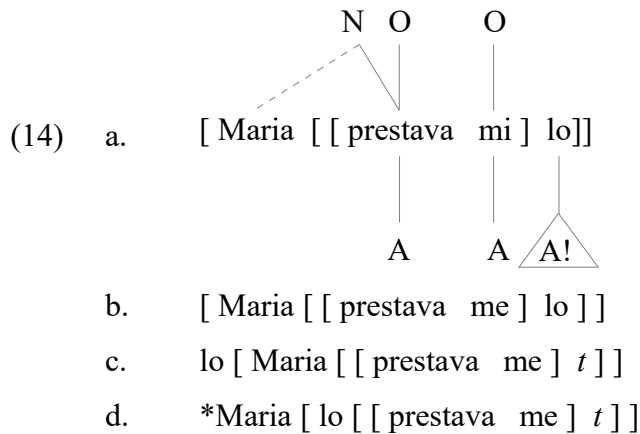
- b. Mi [Maria [[prestava *t*] il libro]
 c. Maria [mi [[prestava *t*] il libro]] (output)

The derivation continues, moving the outbound clitic (π -specified) out of the complement domain of the verb, consistent with OCP, as shown in (13b). Now since the subject comes to be sandwiched between two case-specified elements, it must move to an edge of the structure, thereby allowing the clitic to stay adjacent to the verb, consistent with OCP-2, (13c).

In any derivation, the number of derivational steps, including the input, may not be greater than 3, consistent with the Derivation Equivalence Number (DEN).⁶ If a sentence contains two object clitics and a lexical subject, it is impossible for all the constraints to be satisfied on time. Consider, indeed, the case where the lexical direct object is replaced by a corresponding clitic, as seen in input (14). The indirect object clitic takes both case of the verb and therefore the accusative case of clitic *lo* cannot pair up with a relevant and valid case of the verb (which is indicated by enclosing the feature into a triangle with an exclamation mark). In this theory, it is possible to manipulate the features of a morpheme to rescue an ill-structure. For instance in French, the clitic *me* becomes *moi* (tonic form) by feature manipulation in positive imperative verbs, as discussed in Desouvrey (2000) (see also Bonet 1991). Feature manipulation of the first clitic may take place; eliminating both cases of *mi* returns the tonic

6 The DEN (also Derivation Extension Number) is at this point an empirical number (Cf. Desouvrey, 2007, 2008). I do not know at this point if it is derivable from other notions. Notice that more than one operation may take place in each step if there is no feeding relationship between them, that is, the movement of one element is not the cause of the movement of another (see a handbook of phonology for more on the notion of *feeding* vs. *bleeding*, for instance Kenstowicz 1994).

form, (14b). Then the newly licensed direct object clitic moves to the left edge, (14c), which triggers movement of the lexical subject, yielding an unwanted result, (14d). The last step is shown for the purpose of illustration, since the derivation is cut off at the third step, consistent with the DEN. We shall refer to this as an overload derivation.⁷ This strategy could have been used if it were not overload.



It appears that Italian cannot easily derive two object clitics in a sentence, unlike other Romance languages. This is due to the feature $[\pi]$, which forces a long outbound movement of the clitic and the feature $[\omega]$ which leads to the computation of the indirect object before the direct object (see below). Let us compare the derivation of a similar sentence in French. The clitics in the following derivation are similar to their Italian counterparts, except for the features $[\pi]$ and $[\omega]$ which are absent in the French pronominal system. Notice also the subject is assigned Nominative by spreading and that the extra Accusative of the second clitic cannot (and need not) pair up with the Accusative of the verb (hence the absence of the “!”), since it is Oblique-licensed. Although the direct object is more embedded than the indirect object clitic, the latter moves first and adjoins to the subject (inbound movement). The reason is that the heaviest clitic (the most specified) must become the head for the second inbound clitic, as seen in the final step, which is within the limit of the DEN.

7 It could be possible to cut a derivational step by dropping the subject, which is possible in pro-drop languages, like Italian. The grammar would have to rule in a derivation if the subject is dropped and rule it out otherwise. However, there is no reasons to assume this sort of parochial rule, which implies the use of conditional constraints. Quite contrary, a principle applies or fails to a complete paradigm, not to specific structures.

- (15) a.
$$\begin{array}{c} \text{N} \quad \text{O} \quad \text{O} \\ \diagdown \quad | \quad | \\ [\text{Maria} [[\text{prêtait} \text{le}] \text{me}]] \\ | \quad | \quad | \\ \text{A} \quad \text{A} \quad \triangle \end{array}$$
- b. Maria = me prêtait le
- c. Maria = me = le prêtait (output)

To conclude, the features $[\pi]$ and $[\omega]$ (see below) cause the derivation of a cluster of object clitics to overload. In the next section, I turn to show how this problem is avoided in Italian.

4. Circumventing the overload problem

I argue that certain sequences of two clitics are coalesced into a single element in order to be operational in the syntax so as to avoid the violation of several constraints. To put it in another way, object clitic clusters in Italian are reanalyzed as compound. I claim that a morphological rule takes two adjacent Accusative-specified clitics and returns a two-morpheme word with a single accusative feature, as shown in (16). ('#' stands for a word boundary.)

- (16) **Clitic Compound Rule (CCR)**
- $$\begin{array}{c} \text{C1} \# \text{C2} \rightarrow \text{C1} \sim \text{C2} \\ | \quad | \quad \diagdown \quad / \\ \text{A} \quad \text{A} \quad \text{A} \end{array}$$

To differentiate the compound thus formed to the original sequence, that is, to avoid a string-vacuous operation, a sandhi rule ties up both morphemes together; see (17). This rule transforms the vowel of the first morpheme, [i], into [e]. The symbol '+' stands for a morpheme boundary, which replaces either a tilde ('~') for compounds created with CCR or an equal sign ('=') for those created by adjunction. It is interesting to notice that this rule does not refer to specific segments beyond the boundary (always *l* and *n*), since its rationale is known.⁸

⁸ Clitics seem to make a compound with infinitive verbs as well, as suggested by the rule that deletes the final vowel of the infinitive before a clitic: cf. *comprarelo → comprarlo (to buy it).

(17) Italian Clitic Internal Sandhi (ICIS)

$$i \rightarrow e / \text{ ____ } +$$

In the light of what precedes, let us return to derivation (14). From the input (18a), automatically modified by (16) and (17), the compound moves outbound, i.e., to the left edge of the structure, forcing the subject to clear up, as seen in (18d), which is the desired result.

- (18) a. [Maria [[prestava mi] lo] → [Maria [prestava me~lo]]
 b. me~lo [Maria prestava]
 d. Maria [me~lo prestava] (output)

Italian distinguishes at the third person a feminine and a masculine indirect object clitic, namely *le* and *gli* respectively. However the feminine form never appears in a cluster (cf. **le la/lo/li*), as well-known in the literature; a compound with the masculine is used instead (cf. *glielo, gliela*, etc.). This gap straightforwardly follows from the present analysis. If a derivation proceeds with indirect object *le* in a cluster, CCR should apply to create a compound. However, the obligatory sandhi rule may not process it since its structural description is not met. The grammar must use instead an alternative input with the masculine form of the clitic, as shown in (19b-c). In the same vein, since the sequence **le lo* is ruled out, the grammar may not allow it to be derived either by an application of ICIS. Thus the sequence *li la/lo*, where *li* is the dative plural of *le*, may not be seen to yield **le~la/lo*. This is consistent with the assumption that constraints are context free, that is, the grammar for the ease of its acquisition, does not use conditional constraints.⁹

- (19) a. *[Maria [[presto le] lo]] (failed natural input)
 b. [Maria [[presto gli] lo]] → [Maria [[presto glie ~ lo]]]

9 The ill-formedness of any compound involving *le* may explain the double case specification of the clitic *gli*. Just like its French counterpart *lui*, it can never be used as a direct object. Thus, one may wonder why it has to be specified for Accusative, unlike French *lui*. It is possible that the failed compound *le+DO-cl* is repaired, instead of being dismissed altogether, by replacing the string *le* by *gli* without altering the original features. Once it is created, it can be broken up into its parts, which results in a new *gli* with two case features. The original *gli* is then supplanted, given the harmonization process. Notice that in French clusters, the order is IO>DO, except in clusters involving *lui* and a third person clitic; cf. *la/le/les lui* (see Desouvrey 2000, 2005). If *gli* were an Oblique only clitic, it would move outbound before the accusative clitic, resulting in the order **lo/la gli*. The failure of the *le* cluster appears to have a benefit elsewhere for the grammar: the ordering is the same within all clitic clusters (IO>DO). French does not have bad clusters, but it still has to harmonize the paradigm. This process seems to be in progress, according to data reported in Seuren (2009); cf. the *lui-le-lui* construction. Seuren also points out that Italian has an analogous construction, which is less frequent than the French one. Since in Italian, the order is always IO>DO, this construction may not have the same purpose as in French; rather, it seems to be a strategy to defuse the gender ambiguity of the *gli* clusters.

- c. Maria glie~lo presto (output)

ICIS is a surface lexical phonological rule that mandatorily applies across a morpheme boundary. The compound to which it applies need not arise from the CCR. A compound created in the course of the derivation must be processed as well. To see this, consider the derivation (20), which shows impersonal *si* with the partitive clitic *ne*. Recall that *ne* is an inbound clitic in that it is not specified for $[\pi]$. Thus, unlike the other clitics, it simply adjoins to the impersonal subject, a process that creates a syntactic two-tier compound. Then ICIS automatically transforms $*si = ne$ into $se = ne$.¹⁰

- (20) [si [prende ne]]
 [si = ne prende] → se = ne prende (output)
 'One takes some.'

Moreover, the clitic *ne* offers a further piece of evidence for the compound analysis. Used alone, it adjoins to a head, but in a cluster with a π -specified clitic it cannot be licensed, as seen in (21). Therefore CCR applies, and the clitic becomes a part of a compound that must move outbound. Clearly if the clitics were not in a compound, they would move separately, and the output would be $*mi se = ne$.

- (21) a. $\begin{array}{ccccc} & N & N/O & O & \\ & | & | & | & \\ [& si & [[& d\grave{a} & mi &] & ne &] &] \\ & & & | & | & & / & \\ & & & A & A & & \triangle & \\ & & & & & & A & \end{array}$

one gives me some.

- b. $mi \sim ne [si [d\grave{a} t]] \rightarrow me \sim ne si d\grave{a}$ (output)
 'One gives me some.'

Now consider (22a), which exhibits a reflexive cluster and impersonal *si*. ICIS applies to the input, (22b), then the compound moves outbound, yielding the desired result, (22c). Notice that both object clitics are vector, which means there are ω -specified. Such scope-bearing elements are rigidly

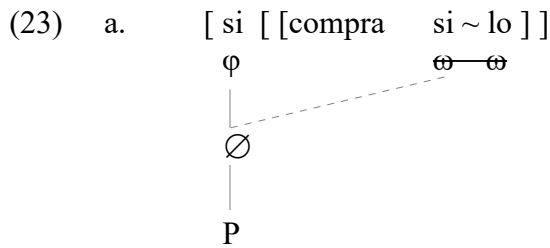
10 Why is it that the clitic *ne* is alone in the paradigm without the feature $[\pi]$? It seems that the grammar uses it as a hint at learners, for it makes it possible to reckon the feature $[\pi]$ elsewhere. Without the *ne* exception, it would be indeed very difficult to sustain the long outbound movement of the other clitics. The outbound movement is not easy with impersonal *si*, which prevents the adjacency of π -specified clitics to the verb, in violation of OCP-2.

ordered and therefore may not commute during the derivation. Thus, if the clitics were not in a compound, the reflexive would have to move first and then the other clitic, which would yield the unwanted order **lo si*, in violation of the Superiority Effect.¹¹

- (22) a. Se lo si compra
 refl. it one buys
 'One buys it for oneself.'
- b. [si [[compra si] lo]] → [si [compra se ~ lo]]
- c. se ~ lo [si compra] (output)

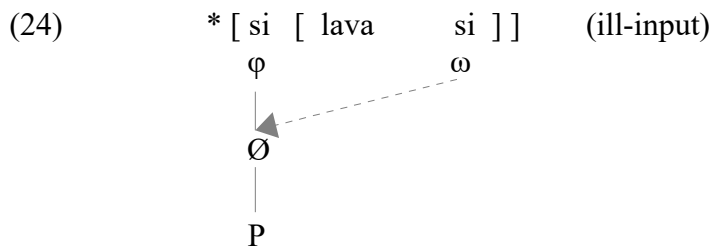
In (22) a further event takes place: CCR causes the feature [ω] of both clitics to be disabled, a process that normally takes place when two vectors are morphologically bound together (see Desouvrey 2010 and also 2008, where it is overlooked). This is a welcome result, but to appreciate it, let us take a look on coreference. In Desouvrey (2003, 2006), referring elements, i.e., NPs, pronouns and anaphors, are differentiated from each other by their referential feature structure. NPs and certain pronouns, always referentially autonomous, bear a terminal referential feature (noted with capitals like P, Q, etc.), which is the virtualization of any real world (or some specific universe) element. As can be seen in (23), this feature, P, expands from the null thematic feature of the subject (on thematic feature, see below), which originates from a φ -node, since impersonal *si* is (by assumption) a scalar. Unlike NPs and pronouns, anaphors have only a referential root node, either [φ] (scalar), [ω] (vector) or [\emptyset] (neither). The coreference relation is simply realized by spreading a thematic node or else a terminal feature: a thematic node from an NP or a pronoun spreads onto the root node of an anaphor; a terminal feature of an NP spreads onto the thematic node of a pronoun. Regarding anaphors, there are a few restrictions on this process, one of which is that the donor and the receiver must be compatible. It appears that no coreference relation can exist between a scalar pronoun and a vector anaphor, and vice versa. However, in (23) the vectors are disabled by compounding, and their feature amounts to \emptyset , which makes the coreference possible.

11 The Italian orthography fuses every clusters involving *gli*. In the light of the present analysis, every clusters whose the second clitic begins with an *l* or an *n* should be fused as well.



b. se ~ lo [si compra] (output)

This analysis straightforwardly predicts that if reflexive *si* is not in a compound, it may not acquire its scalar antecedent, impersonal *si*. This is borne out, as shown in (24).



It is not possible to derive a well-formed sentence from input (24), cf. **si si lava*. The grammar uses rather an alternative input where reflexive *si* is replaced by first person plural *ci*, which is a scalar (either 0- or φ-specified), (25). Why *ci*, but not another clitic? It appears that the clitic *ci* has almost the same semantic as the impersonal *si*. So in the impossibility to use the reflexive *si*, it is the only element that can fulfill this role, just like *mi* and *ti* are used as reflexive for first and second person subject for lack of a dedicated morpheme. Moreover, in French, which has a clitic similar to the impersonal *si*, namely *on*, *nous* (first plural) is used as its corresponding dislocated tonic form; compare *Moi, je suis heureux* 'As for me, I am happy' and *Nous, on est heureux* 'As for us, we are (lit. one is) happy'. In the literature, this is accounted for with an OCP-triggered rule: *si* → *ci* / ___ *si*, which is unlikely since the domain is the sentence, not the VP. In addition, this rule appears to be purely descriptive and very parochial, for the cooccurrence of identical clitics is not excluded in Romance (cf. French *nous nous lavons* 'we wash ourselves.')

- (25) a. [si [lava ci]]
 ϕ ϕ
 | |
 \emptyset
 |
 P
- b. ci [si lava] (output)

An independent piece of evidence for the coreference analysis of **si si* is provided by certain well-known phenomena in the literature, namely the so-called *that-t* effect in English and the *que-qui* alternation in French, and more generally embedded questions. Consider the structures shown in (26). Under the analysis that *que* is a relative anaphor referring to *Marie*, (26a) is perfect since the relative and its antecedent are both scalar elements. On the other hand, (26b), where the antecedent of anaphor *que* is a wh-vector, becomes acceptable by deletion of the anaphor (see Desouvrey 2007, 2008a, 2008b).

- (26) a. [Je sais que] [Marie est venue]
 ϕ ϕ
 | |
 \emptyset
 |
 P
- b. [Je sais ~~que~~] [qui est venue]
 ϕ ω
 | |
 \emptyset
 |
 P

To summarize, object clitic clusters are syntactic compound in that they behave as a single element with respect to movement. A remarkable cooccurrence restriction, namely **si si* is due to a feature clash between a scalar antecedent and a vector anaphor. The next section is devoted to further restrictions.

5. On the Person Case Constraint

The PCC is standardly invoked to rule out certain clitic clusters since Bonet (1991). According to this

constraint, the indirect object cannot be third person when the direct object is second or first person. In fact, it appears that in a cluster the direct object can only be a third person clitic. In Desouvrey (2000), based mainly on French data, this constraint is overlooked, since its effects follow from the proposed analysis and the distribution of features in the paradigm of French personal pronouns. It appears that first and second person clitics in French are specified for Accusative and Oblique, whereas third person clitics use a distinct morpheme for each case. Since in French the order of the clitics in the input structure is DO>IO (which is another reason why French does not need something like CCR), a first or second person direct object will saturate the verb, pairing up with both Accusative and Oblique, hence leaving no room for a further clitic to be licensed. In other terms, PCC is then taken to be a licensing problem.

With respect to case features, Italian opposes French in two ways: (i) every argumental indirect object is specified for both Accusative and Oblique and (ii) the ordering in the input is IO>DO (see below). As a result, no second clitic can ever be licensed, hence the cluster must be processed by CCR and ICIS in the course of the derivation, as discussed above.

Now if PCC were simply due to a case-licensing effect, as assumed in Desouvrey (2000), one would expect a cluster such as **ti gli/gli mi* to be treated by CCR, just like **mi la (me~la)* or **si lo (se~lo)*. Since this is not the case, there must be a further reason that precludes such clusters.

The formulation of the PCC is just a description of the problem, it is by no means an explanation. An account of this phenomenon appeals to a fresh look of the notion of case, grammatical relation and thematic features in the grammar. The subject is universally marked with Nominative, the direct object with Accusative, while an indirect object can take various cases, the most common being Dative in languages with a limited use of case morphology, like English and Romance. As for the thematic features, the subject is traditionally the agent while the direct object is the theme. However, it turns out that, the subject and the direct object are unspecified for thematic features, unlike the indirect object which includes thematic features like Ablative, Locative, Benefactive, etc (cf. Desouvrey 2010, 2012/2013).¹² It is reasonable to assume that in languages where cases are confined in a closed paradigm, like clitics, the pool may contain at most one thematic feature.¹³ From this perspective, and

12 The assumption that Nominative and Accusative are underspecified for thematic features is supported by the fact that languages with enough morphological power do not use them as the subject or the object or certain verbs, for instance psych-verbs.

13 The pool is the mechanism (represented as a sine wave) that allows grammatical relations to align with cases, as discussed in Desouvrey (2013). I assume here, pending further research, that the pool in such languages like Korean is expandable to accommodate more than one thematic, while in languages like Romance it is rigidly fixed to one and a

consistent with the Pool Theory, the sequences (27b-c) may not be possible in Romance-like languages. It appears that PCC is a constraint on thematic features, which can be formulated as in (28).

- (27) a. Nom-Acc-Dat
 b. *Nom-Dat-Dat
 c. *Nom-Acc-Acc
 d. *Nom-Dat-Locative/Ablative

(28) **The Person Case Constraint (PCC)**

A cluster of object clitics must contain one and only one thematic feature.

It is important to note that this constraint will apply to real features of a morpheme and not on the interpretation we may give to a morpheme. For instance, although the Nominative is not thematic in the sense just defined one can feel whether it is an agent or an experiencer, according to the verb. Similarly, a locative morpheme may or may not bear a thematic role; it can be left unspecified if another Oblique-specified element in the paradigm has a thematic feature.

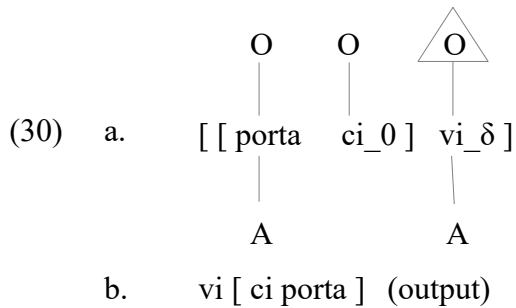
With this in mind, let us turn now to clusters involving a direct object and a locative indirect object. Even though there are two locative clitics, *ci* and *vi*, only the former can be used in a cluster, as shown in (29) (cf. Monachesi 1998). Based on what we already know about this clitic paradigm, suppose that locative *vi* is specified for both Accusative and Oblique with a locative thematic feature, as indicated in Table 2. As for *ci*, I suggest that it represents two different morphemes: one instance is specified for oblique, but not for a thematic feature; the other is specified for both Accusative and Oblique, as well as the locative thematic, just as *vi*.¹⁴ From this perspective, the clusters in (29b) includes two thematic features, namely a Dative and a Locative, and therefore are ruled out by TRC

- (29) a. mi ci porta / ti ci porta / vi ci porta
 me there porta / you.acc there porta / you.acc there porta
 b. *mi vi porta ; *ti vi porta; *vi vi porta
 'It takes you/me there.'

half cycle. Also, notice that the case-morphemes in languages like Korean, Japanese, etc., are in fact thematic pronouns (see Desouvrey 2010).

14 Monachesi points out that locative *vi* is mainly used in the literary language. This is expected under the hypothesis that it bears the same features as one instance of *ci*. In addition the string /ci/ is already involved in various alternations; so given the harmonization process, which is intended to ease the acquisition of the grammar, it is generalized as the only operational clitic (for substitution operation) in the system.

The derivation of *vi ci porta*, for instance, proceeds from the natural input (30a). The unspecified locative indirect object (with a 0 thematic) moves first and then the dative-specified direct object, yielding the desired result, (30b). Notice that each case of the arguments correctly pairs up with a case of the verb, except the extra Oblique of *vi*, which is ignored. I assume also that the verb is endowed with the relevant thematic features in order to accommodate the clitics.



However, the following cluster with a thematic *ci* (direct object) and a non-thematic *ci* (indirect object) is not possible. A well-formed sentence can obtain by dropping the indirect object. In this context, the locative indirect object is not crucial for the interpretation of the sentence. Plausibly the dismissal of the indirect object makes it possible to avoid the repetition of the morpheme.¹⁵

- (31) (**Ci*) *ci porta* Micol.
 there.CL us.CL bring.3SG Micol
 ‘Micol brings us there.’

The unspecified locative morpheme *ci* can also be used as an indirect object to non-locative verbs in order to avoid a violation of PCC, as illustrated in (32a) (cf. Pescarini's (7)). The morphemes *ti* and *gli* both bear the dative thematic role and therefore are barred to cooccur in the input (32b). The derivation must rather proceed from the input (33a), which yields the correct result after the rightward movement of *io* (OCP-2). Notice that, as Pescarini points out, *ci* cannot cooccur with first person plural *ci*, nor can it be dropped to rescue the structure, which is presumably due to an OCP effect (see

15 This may be due to an OCP effect, as Pescarini points out. Indeed, both morphemes are adjacent under the VP domain, unlike the case of **si si lava* discussed above. Be that as it may, there seems to be a general tendency to avoid a useless repetition of the same morpheme in languages. That is to say, a morpheme is repeated only if it conveys a significant piece of information, as are for instance the sequence *vous vous* and *nous nous* in French (cf. *nous nous lavons* 'we wash ourselves' *vous vous lavez* 'you wash yourselves'). This tendency, which might well be due to hypercorection by literary people, accounts for the fact that clitic *y* cannot be used with the verb *aller* 'to go' in the future tense: **y'irai* 'I will go there'. In this case, clitic *y* is phonetically identical to the first syllable of the verb, *i-rai*. It is mandatory to drop the clitic in this context. Generally students learning French are often puzzled when they first meet with a sentence like *Vous vous appelez comment?* ('What is your name?')

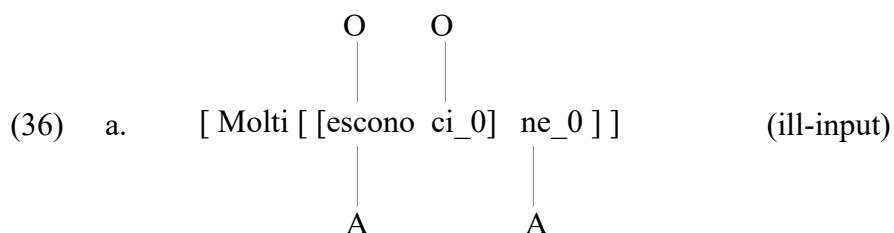
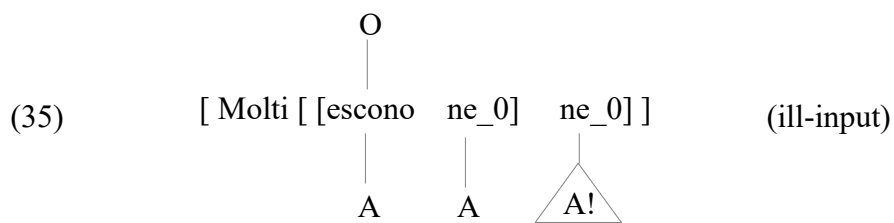
footnote 15).

- (32) a. Ti *ci*/**gli* presento io, al direttore.
 you.CL to.him.CL introduce.1SG I to.the director
 ‘I will introduce you to him (the director).’
 b. *[io [[presento gli_δ] ti_δ]]
- (33) a. [io [[presento ci_0] ti_δ]]
 b. ti [ci [io [[presento t] t]]
 c. ti [ci [t [[presento t] t]] io (output)

Furthermore, locative clitic *ci* is used instead of ablative *ne* when the direct object is partitive *ne*, as shown in (34). Pescarini, who takes it to be an OCP effect, assumes that ablative *ne* becomes *ci*. However, given our assumptions, this sentence needs a further look. In effect, OCP alone cannot rule out **ne ne* in (34) any more than the clusters **la la* and **lo lo* (since **la lo* is not good either). The problem is that, unlike the cluster **ti gli*, the clusters just mentioned do not contain any thematic feature: *la*, *lo*, and *ne* are all 0-specified. Thus both inputs shown in (35) and (36a) can be dismissed. However, as can be seen, a derivation with (36a) yields a well-formed output, (36c). This result is purely accidental, as I argue.

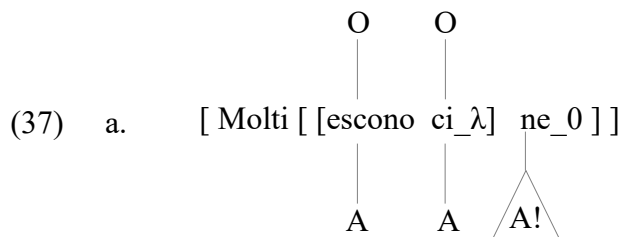
- (34) Ce/**ne* *ne* escono molti.
 from.there.CL of.them.CL come.out.3PL many
 ‘Many of them come out from there.’

In input (35), both clitics bear Accusative, but CCR fails (**ne~ne*) for the same reason as the cluster **le~lo*: ICIS cannot seal them since its structural description is not met. In the input (36a), which violates (28), *ci* must be a sister to the verb, as is normally the case for indirect object clitics in Italian. Since it is π -specified, it moves outbound, forcing *molti* to clear the intermediate position between the verb and the clitic, (36c). Finally, the inbound clitic *ne* adjoins to *ci*, making up a compound which is automatically processed by ICIS, (36c). Actually, under the assumption that a derivation is cut off at the third step, consistent with DEN, this correct output may not be used.



- b. ci [molti [escono ne]]
 c. ci [escono ne] molti
 d. ci = ne escono molti → ce = ne escono molti (correct output)

The correct input for (34) is given in (37a), with the locative-specified *ci* instead. CCR and ICIS apply, then the compound moves outbound, triggering the movement of *molti*, as seen in (37d), which is the desired result.



- b. Molti escono ce~ne
 c. ce ~ ne molti escono
 d. ce ~ ne escono molti

In order to show that the well-formedness of (36d) is accidental, consider now the sentence shown in (38), where the locative-specified *ci* must be crucially used. If (38) proceeds from an input with the non-thematic *ci*, as in (39a), both clitics pair up with the verb, and therefore CCR may not apply. Thus the first clitic moves outbound, followed by the accusative clitic, which results in an ill-output.

- (38) Ce lo vedo
 there it.acc I.see.
 'I see it there.'

- (39) a. $\begin{array}{c} \text{O} \quad \text{O} \\ | \quad | \\ [[\text{vedo} \quad \text{ci}_0] \text{lo}_0] \\ | \quad | \\ \text{A} \quad \text{A} \end{array}$ (ill-input)
- b. *lo [ci vedo]

The derivation of (38) must rather proceed from input (40a), which contains the locative-specified *ci*, as required under (28). Since the second clitic is not licensed, CCR creates a compound which is sealed by ICIS. Then, the compound moves outbound, yielding the desired result.

- (40) a. $\begin{array}{c} \text{O} \quad \text{O} \\ | \quad | \\ [[\text{vedo} \quad \text{ci}_\lambda] \text{lo}_0] \\ | \quad | \quad \triangle \\ \text{A} \quad \text{A} \quad \text{A!} \end{array}$
- b. ce ~ lo vedo

To conclude, it appears that the PCC, reformulated in (28), is a constraint on thematic features in the sense discussed above. A double object verb must have a thematic argument and a non-thematic one. Two specified thematic features are thus banned. In many instances, this constraint is avoided by using one or the other instance of the clitic *ci*. In the next section, I turn to another remarkable property of Italian, the agreement of direct object clitics with past participles.

6. Vector Effects

We have seen that third person accusative-specified clitics are vectors, as opposed to scalar elements, in that they are also specified for the feature $[\omega]$. It appears that this feature, unknown in other linguistic theories, is involved in all the complexities of Italian grammar, including CCR, and the various restriction on cooccurrence of clitics.¹⁶

Recall that in simple tense verbs, the direct object is computed after the indirect object, and as a result the former is not case-licensed. If the verb were merged first with the DO and then with the IO,

¹⁶ This feature $[\omega]$ seems to be inherent to negation and perhaps *wh-elements*. For more ω -effects, see Desouvrey (2002, 2009, 2008a, 2008b, 2010 and 2013).

just like in French, both arguments would be licensed, and there would be no need to use the compound strategy and the sandhi rule. The obvious question is why Italian verbs, with respect to clitics, cannot compute the direct object before the indirect object. This is tackled in the next subsection.

6.1 Past Participle agreement

Italian has a further complexity that makes it unique among well-studied Romance languages: the past participle of compound tense verbs mandatorily agrees with third person direct objects and optionally with other persons. Past participle agreement exists as well in standard French, but it appears to be a calque from Italian. (Based on the analysis of this phenomenon in Italian, it will become clear that such a rule is not genuine to French grammar.)

The building of the syntactic structure is regulated by strict rules on merger. What is relevant here is that the verb will merge first with the best match, that is, the element that does not oppose (or is maximally compatible with) its own features. Now, simple tense verbs are not ω -specified, but they might be either φ -specified or not specified at all. If these verbs were φ -specified, they would never merge with third person clitics. Since this is not the case, one can assume that they bear a null (neutral) specification, which allow them to emulate the feature of the complement. On the assumption that non-third person clitics are neutral just like simple tense verbs, ditransitive verbs must postpone in second place the computation of third persons clitics, yielding the order IO>DO in the input, as seen above and below. (The barred zero indicates a neutral element, neither φ - nor ω -specified.)

- (41) a. [[presto_θ mi_θ] lo_ω]
s/he.lent me it
- b. *[[presto_θ lo_ω] mi_θ]
s/he.lent it mi

Suppose that past participles are specified for $[\omega]$, just like third person direct object clitics. Therefore the past participle is more compatible with the clitic than with the auxiliary. As a result, a transitive verbs in a compound tense has to use two different computational patterns, according to whether or not its argument is a vector. If the complement is a vector, the past participle and the clitic must make a constituent which then merges with the neutral auxiliary, as in (42a) instead of (42b). Since the past participle and the clitic are equipped with the same morphology, namely gender and number, they normally agree under this configuration. Clearly, it happens that past participle agreement is a simple reflex of the agreement between a head and its complement, a normal phenomenon across

Romance.

- (42) a. [ha_θ [vista_ω la_ω]]
 (s/he) has seen.fem. her
 b. *[[ha_θ vista_ω] la_ω]

On the other hand, if the argument is not a vector, the auxiliary may make up a constituent with the participle as its sister complement and then that constituent merges with the clitic, as shown in (43a). Such a pattern is the normal case for all languages that use compound tenses. However, it is not quite optimal in Italian, due to the presence of the feature [ω], which loosens it. Indeed, there is no principled reasons that prevent the computation from proceeding the other way around, as in the case of third person clitics, (43b). Under (43b), the non-third person clitic and the past participle are expected to agree normally. This is correct, since agreement of past participles with non-third person clitics are apparently frequent in colloquial speeches.

- (43) a. [[ha_θ visto_ω] mi_θ]
 b. [ha_θ [visto_ω mi_θ]]

Let us illustrate this with a cluster complement, for instance the sentence in (44). From the above discussion, the input structure must be as shown in (45a). Being π-specified, the first clitic moves outbound, (45b), without violating the superiority effect, which is relevant for movement within the initial domain (inbound). Then the second clitic follows, yielding the sequence *mi la*, which meets the structural description of the sandhi rule, (45c).

- (44) Me la ha presta.

- (45) a.
$$\begin{array}{c} \text{O} \qquad \qquad \text{O} \\ | \qquad \qquad | \\ [\text{ha} [[\text{presta}_\omega \text{ la}_\omega] \text{mi}]] \\ | \qquad | \qquad \triangle \\ \text{A} \qquad \text{A} \qquad \text{A} \end{array}$$

 b. la_ω [ha [presta_ω mi]]
 c. mi_ω [la_ω [ha presta_ω]] → me~la ha presta

As can be seen in the derivation (45), ICIS applies after the movement of the clitics to their final position. In the input, the cluster *la mi*, although consistent with CCR, cannot be licensed by the sandhi

rule, which allows the clitics to move separately. Incidentally, this shows that ICIS and CCR are two sides of the same coin. CCR applies to two different clitics if they have each an accusative case. ICIS validates the cluster only if it is consistent with its own structural description.

A further interesting case is the fact that past participles cannot agree with the relative *che*, as discussed in Belletti (2006), see (46) below (her (17)). Based on passed studies of relative clauses, it is likely that relative pronouns are universally φ -specified, which contrasts them with *wh*-operators (see the discussion of (25)-(26)). On this view, the relative and the past participle may never agree, since they cannot merge, (47a). To avoid a feature clash, the relative must merge with the complex auxiliary-past participle, as seen in (47b), the correct input from which (46b) is derived.

- (46) a. *I libri che ho letti.
 the books (Masc, Pl) that (I) have read (Masc, Pl)
- b. I libri che ho letto.
 the books that (I) have read
- (47) a. *[I libri ...] [ho [letti_ ω che_ φ]]
- b. [I libri ...] [[ho_ θ letto_ ω] che_ φ]

6.2 Further complexities

The ω -effect makes it possible to account for certain facts that are considered as arbitrary in Monachesi (1998). Consider first the fact that third person accusative clitics cannot be the complement of a present participle; compare (48) and (49).

- (48) *Gli argomenti riguardanti-lo /-la /-le /?-li
 The topics concerning cl. (acc)
 'The topics concerning him / her / them'
- (49) Gli argomenti riguardanti-ci /-mi /-vi /-ti
 The topics concerning cl. (acc)
 'The topics concerning us / me / you'

This restriction is presumably due to the feature specification of the present participle. If infinitives and past participles are ω -specified, as seen above, it is very likely that present participles bear this feature as well. If so, it is a matter to know why the interaction of two vectors fails. Recall past participles

agree with third person accusative clitics. This agreement is purely circumstantial, that is to say, the agreement is not the rationale of the construction. Nevertheless, its role changes from an epiphenomenon to a crucial licenser of a head complement relation between two vectors, as I argue. Being a vector, the present participle perfectly merges with a third person accusative clitic. However, unlike past participles, present participles agree with their subject not with their complement, which means that these agreement features (gender and number) are invisible to the object clitic. By a pervert effect of the harmonization process, reminiscent of ICIS vs. **le lo*, the structure is rejected, since the clitic cannot agree with the head. Monachesi points out that the third person plural accusative clitic *li* is marginally acceptable, as shown in (48). This is precisely what can be expected under this analysis. The plural clitic, which is also doubly specified for case, like the IO clitics, coincidentally agrees with the participle, hence the illusion of acceptability. Predictably this illusion disappears with a singular present participle.

A related question is why every clitic complement of the present participle must stay in situ. Unfortunately, I do not have enough data to clearly answer this question. But tentatively, it seems that the outbound movement of the clitic cannot be redeemed, presumably because the lexical subject cannot move out of its intermediate position between the clitic and the verb, unlike the case discussed above. In other use of the present participle without a subject, which is likely possible, the movement of the clitic would create a fatal vector effect. Recall that a lower vector cannot vacuously cross a boundary, since this movement amounts to a commutation of two vectors:

- (50) a. [vector_1 vector_2]
 b. *vector_2 [vector_1 t]

Monachesi mentions a further type of construction referred to as verb left-detachment. The detached verb can be either a bare or a cliticized infinitive, as illustrated in (51). (52) shows that a lexical NP, *Martina*, cannot be the complement of the infinitive verb (Monachesi p. 6). Tentatively, I suggest that an effect involving vectors takes place. In (52), the lexical NP stands between two vectors, the infinitive verb and the third person clitic, a configuration which is barred under OCP-2. Dropping the lexical direct object is the only way to redeem this structure. This problem does not arise in (51), since the clitic and the infinitive make up a compound (recall footnote 9).

- (51) Vedere/vederla, la vedo ogni giorno.
 to see/to see.cl, cl see every day

'As for seeing her, I see her every day.'

(52) *Vedere Martina, la vedo ogni giorno. (Underlining added)

7. Conclusion

I have presented a syntactic analysis of clitic clusters based on a few features and their interaction according to general universal principles. The specificity of the Italian clitic paradigm is due to the prominence of the features $[\pi]$ and $[\omega]$. The latter is the main source of complications in the grammar. It forces the verb to merge first with the indirect object, which leaves the direct object clitic unlicensed. In addition, in compound tenses, it triggers the agreement of the past participle with the accusative clitic. The grammar avoids the licensing problem with a compound rule (CCR), which must be processed by a sandhi rule (ICIS). As for the feature $[\pi]$, we have seen that it forces all clitics but *ne* to move outbound, hence augmenting the overload of the derivation. Finally, the Person-Case Constraint, which accounts for certain cooccurrence restrictions, turns out to be a constraint that bans the occurrence of two thematic features.

8. References

- Belletti, Adriana (2006). (Past) Participle Agreement. <https://unige.ch/lettres/linguistique/files/4514/7436/7268/Belletti/forthcoming>.
- Bonet, Eulalia (1991). Morphology After Syntax: Pronominal Clitics in Romance. Doctoral dissertation, MIT.
- Desouvrey, Louis-Harry (2000). Romance Clitics and Feature Asymmetry: An Autosegmental Based-Approach. Doctoral dissertation. <http://ling.auf.net/lingBuzz>.
- Desouvrey, Louis-Harry (2002). Adverbs, Negation and OCP Effects. www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2003). The Proper Treatment of Coreference Relations. www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2005). Romance Clitic Clusters: The Case Connection. In Heggie, L. and F. Ordóñez (eds), *Clitic and Affix Combinations*. Amsterdam: John Benjamins.
- Desouvrey, Louis-Harry (2006). Underspecification and Long-Distance Antecedent: The Case of Chinese Ziji. www.semanticsarchive.net
- Desouvrey, Louis-Harry (2007). Wh-Interrogatives: The OCP Cycle. www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2009). Nominal Predication, Relative Clauses and Wh-Extraction: The Amazing Syntax of Haitian Creole Se. <http://ling.auf.net/lingBuzz>.
- Desouvrey, Louis-Harry (2008a). Vector Effects on Wh-Interrogatives. <http://ling.auf.net/lingBuzz/000755/> and www.semanticsarchive.net.

- Desouvrey, Louis-Harry (2008b). Superiority Effect and Clitic Placement in European Portuguese. <http://ling.auf.net/lingBuzz/000764>.
- Desouvrey, Louis-Harry (2010). The Syntax of the So-Called Internally-Headed Relative Clauses: A Study of Korean Clause Structures. <http://ling.auf.net/lingbuzz/001170>.
- Desouvrey, Louis-Harry (2013). Periodic Linking: The How, Where and Why of Ergativity and Analogous Phenomena. <Http://ling.auf.net/lingbuzz/001655>
- Kenstowicz, Michael (1994). Phonology in Generative Grammar. Cambridge, MA: Blackwell.
- Monachesi, Paola (1998). Decomposing Italian Clitics. www.let.uu.nl/~Paola.Monachesi/personal/papers/csli.pdf
- Pescarini, Diego (2010). Elsewhere in Romance: Evidence from Clitic Clusters. *Linguistic Inquiry*, 41, 3, 427–444.
- Seuren, Pieter A.M. (2009). The clitics mechanism in French and Italian. *Probus* 21, 83-142.