

ON ITALIAN *SI*

MARIA RITA MANZINI

*Department of Italian
University College London
London WC1, England*

1. INTRODUCTION

The purpose of this essay is to explore the properties of the clitic *si* in Italian, proposing that there is one single lexical item *si* from which four different types of *si* can be obtained: impersonal, reflexive, middle, and middle-reflexive. We will argue that the syntactic and semantic properties of each of these types derive from a possibility of associating *si* optionally with a passivizer property and with either a tree-variable or a dependent-variable property. The interaction of these two options indeed gives rise to four possibilities, all of which are attested. This account differs from standard accounts of Italian *si* in that a unified treatment is proposed for all occurrences of *si*.

The organization of this essay is as follows. In Section 2 we discuss impersonal *si*, showing its distribution and properties and proposing a lexical entry for it. In Section 3 we discuss reflexive *si*; in Section 4 we show that with minimal modification the lexical entry for impersonal *si* and reflexive *si* can be unified. In Section 5 we discuss other kinds of *si* which appear in Italian, showing them to involve the same lexical entry for *si* as well. We conclude by speculating on how this account can be parametrized so as to provide an explanatory insight into similar constructions in other languages, for example, Icelandic and French.

2. IMPERSONAL *SI*

Consider so called impersonal *si*. In its interpretation impersonal *si* seems to be equivalent to PRO_{arb} ; hence, following Chomsky (1981), we can say that the interpretation of impersonal *si*, as of PRO_{arb} , is that of a free variable, as for example in (1):

- (1) *Si lava volentieri i bambini*
One gladly washes the children

Impersonal *si* and PRO_{arb} , furthermore, seem to be identical with respect to their agreement features. In particular, PRO_{arb} and impersonal *si* both agree with an adjective in the plural masculine, as in (2) and (3), though impersonal *si* agrees with a finite verb in the third-person singular, as in (3):

- (2) *E' facile PRO_{arb} essere nervosi*
It is easy to be nervous (pl. masc.)
- (3) *Si e' facilmente nervosi*
One is easily nervous (pl. masc.)

Given that in (3) and similar examples impersonal *si* agrees with a finite verb in the singular and with an adjective in the plural, the grammar seems to be faced with a paradox. The paradox however disappears if one assumes that impersonal *si* is associated not only with a third-person feature, but also with an unspecified number feature, or, taking the number and gender features to belong together naturally, with unspecified number and gender features.

If so indeed, impersonal *si* will agree with a tensed verb in third person, unspecified number, and with an adjective in unspecified number, unspecified gender. But it is only natural to assume that given unspecified features in the grammar, if they have to be phonologically interpreted they will be associated at some point or another of a derivation with conventionally specified values. Suppose then we assume that in Italian unspecified number and gender features on adjectives are associated by convention with the values plural and masculine, and an unspecified number feature on a tensed verb is associated by convention with the value singular. If so, the apparently paradoxical data of the type of (3) are obviously accounted for.

From this point of view, the similarity of the agreement features of PRO_{arb} and impersonal *si* reduces to the fact that they share unspecified features. What is more, the agreement features of PRO_{arb} and the impersonal lexical item seem to be subject to exactly the same crosslinguistic variations. From our point of view, this will mean that PRO_{arb} and the impersonal lexical item share unspecified features universally, while the values conventionally associated with unspecified features change from language to language.

So in Spanish it seems to be the case that both PRO_{arb} and the impersonal element, *se*, agree with an adjective in the singular masculine. From our point of view, this means that in Spanish, as in Italian, PRO_{arb} and the impersonal element are associated with unspecified number and gender features; but in Spanish, contrary to Italian, an unspecified number feature is conventionally associated with the value singular on adjectives as well as on finite verbs.

Consider now the core syntactic properties of impersonal *si*. To begin with, it is possible to argue that impersonal *si* is an argument, hence as an argument is subject to the θ -Criterion.¹ In particular, it is easy to see that in all of the well-formed examples which involve it, if impersonal *si* is an argument the θ -Criterion can be satisfied; for example, both in (1) and in (3) impersonal *si* can be associated, for all we know, with the subject θ -position. In those same examples, on the other hand, if impersonal *si* is not an argument, it is also easy to see that the θ -Criterion cannot be satisfied. So if in (1) and in (3) impersonal *si* is not an argument, hence not associated with the subject θ -position, the subject θ -position cannot be associated with any argument at all.

Similarly, consider (4) and (5), which involve an ergative and passive predicate respectively:

(4) *Si va volentieri*
One gladly goes

(5) *Si e' invitati volentieri*
One is gladly invited (pl. masc.)

In (4) and in (5), if impersonal *si* is an argument, it can be associated with the object θ -position; if it is not an argument, the object θ -position cannot be associated with any argument at all. Hence we conclude that impersonal *si* must, and can, be an argument.

Much in the same way, it is possible to argue that impersonal *si* has categorial features N, hence, as a lexical nominal, it is subject to the Case Filter. Of course this reduces to the fact that impersonal *si* forms chains at S-structure and chains at S-structure are subject to Visibility, if Visibility is substituted for the Case Filter.² Consider again the well-formed examples in (1), (3), (4), and (5). In such examples it is easy to see that impersonal *si* can satisfy the Case Filter; in particular, it can be assigned nominative Case.

¹ We informally take the θ -Criterion to state that every argument is associated with exactly one θ -position and every θ -position is associated with exactly one argument. The θ -Criterion is originally from Chomsky (1981); the notion of θ -position is from the version of the θ -Criterion in Chomsky (1986).

² Informally, we assume that at S-structure the association of θ -positions and arguments is mediated by chains; we take then Visibility to state that the head and only the head of a chain is Case-marked. Visibility is originally from Chomsky (1981); the formulation suggested here is essentially from Chomsky (1986).

If on the other hand impersonal *si* cannot be assigned a Case, it is easy to show that this gives rise to ungrammaticality. Consider for example (6) and (7), essentially from Burzio (1981) and Belletti (1982b):

- (6) **E' bello lavarsi volentieri i bambini*
It is good [one to gladly wash the children]

- (7) **E' bello andarsi volentieri*
It is good [one to gladly go]

In both (6) and (7) there is no Case for impersonal *si* to be associated with. If impersonal *si* is subject to the Case Filter, (6) and (7) are predicted to be ungrammatical. But if impersonal *si* is not subject to the Case Filter, the ungrammaticality of (6) and (7) cannot be predicted. In the same way, nominative Case is available to the head of the chain that impersonal *si* forms in examples like (1) and (3)–(5), but no Case is available to the head of the chain that impersonal *si* forms in (6) and (7). If so, the ungrammaticality of (6) and (7) is explained by Visibility. Hence we conclude that impersonal *si*, or the chains impersonal *si* forms, must be subject to some version of Case Theory, be it the Case Filter or the Visibility Hypothesis.

Finally, notice that in the well-formed examples considered so far, impersonal *si* is associated with different θ -positions, in particular with a subject θ -position in (1) and in (3), and with an object θ -position in (4) and in (5). But in all well-formed examples, impersonal *si* seems to be associated with the same Case, the nominative Case. This is not accidental: indeed, it is possible to show that impersonal *si* can never be associated with any Case but the nominative Case. Thus example (8), where impersonal *si* can only be associated with the object θ -position and the accusative Case, is obviously ill-formed: and examples like (9) or (10), where impersonal *si* can only be associated with a subject θ -position and an accusative Case, are similarly impossible:

- (8) **Si lavo volentieri*
I gladly wash one
- (9) **Si vedo lavare volentieri i bambini*
I see one gladly wash the children
- (10) **Vedo lavarsi volentieri i bambini*
I see one gladly wash the children

Thus, we conclude that it is a property of impersonal *si* that it can only be associated with nominative Case.

One problem is left. Obviously, impersonal *si* is phonologically a clitic on a verb. But the phonological configuration of clitic on a verb can correspond to at least three different syntactic configurations. So impersonal *si* could be generated

as a clitic on a verb already in the syntax, as it is generally assumed to be the case for complement clitics; or alternatively impersonal *si* could be generated in the syntax in subject position, as it has been assumed to be the case in Rizzi (forthcoming) for the subject clitics of French; or alternatively yet, impersonal *si* could be generated in the syntax under the INFL node, as it has been assumed in Rizzi (forthcoming), for the subject clitics of the Northern Italian dialects.

In this respect, it is possible to show that though impersonal *si* is associated with the nominative Case, hence a *prima facie* candidate for assimilation to the subject clitics of either French or Northern Italian, the distribution of impersonal *si* and of the subject clitics of French and Northern Italian is incompatible, while there is compatibility between the distribution of impersonal *si* and of complement clitics.

Compare first the distribution of impersonal *si* and of complement clitics. A complement clitic, say an accusative clitic, always appears after the negation particle, as in (11); so does impersonal *si*, and what is more, impersonal *si* always appears internally to the string of the complement clitics themselves, as in (12):

- (11) *Non li lavo*
I don't wash them

- (12) *Non li si lava volentieri*
One doesn't gladly wash them

Compare on the other hand the distribution of impersonal *si* and of the subject clitics of the French kind. The French subject clitics have exactly the same distribution as full-subject NPs; so, in particular, contrary to impersonal *si*, they appear before both the negation particle and all complement clitics, as in (13):

- (13) *Je ne les lave pas*
I don't wash them

Similarly, the distribution of impersonal *si* is quite different from the distribution of the subject clitics in Northern Italian. The Northern Italian subject clitics, at least in the variety illustrated here with the Modena dialect, appear before the negation particle, like the French subject clitics and unlike impersonal *si*, as in (14):

- (14) *La n lava brisa i putēin*
She doesn't wash the children

What is more, in Modenese the impersonal element, *s(e)* can and must co-occur with a subject clitic, to be precise the expletive subject clitic *a*, as in (15):

- (15) *A n s lava brisa i putēin in sta manera*
(There) one doesn't wash the children in this way

Obviously, if the impersonal element and the subject clitic co-occur, they cannot occupy the same syntactic position.

Thus we conclude that impersonal *si*, like complement clitics, is generated in the syntax as a clitic on a verb, and is neither in the subject position nor under INFL, unlike the subject clitics of both French and Northern Italian.³

Let us, then, sum up so far. If we are correct, impersonal *si* is associated with the following properties: it is interpreted as a free variable; it is an argument; it has the categorial feature N; it has third-person features and unspecified number and gender features; it is not only phonologically but also syntactically a clitic on

³ Belletti (1982b) explicitly proposes that impersonal *si* is generated under INFL. But it is difficult to see how under such a proposal the Modena dialect could be accounted for. One could say in principle that INFL includes two distinct positions, one for AGR and one for impersonal *si*. But this would amount to introducing a new position in the grammar for the purposes of impersonal *si* only, obviously an undesirable move.

Burzio (1981), on the other hand, explicitly proposes that impersonal *si* can be generated in subject or object position, but is always cliticized from subject position. The different syntax of the French clitics and of impersonal *si* seems however to argue against such a theory. It is true that in the case of French clitics one could assume that the rule which cliticizes them is a rule which applies in the PF component, while Burzio needs to assume that impersonal *si* cliticizes in the syntax. But such analysis only makes matters worse. For, according to Burzio, cliticization of impersonal *si* leaves a trace in subject position; hence he is forced to stipulate that the trace is properly bound by impersonal *si* despite the anti-c-command relation, clearly an ad hoc stipulation.

While the issue of the syntactic position of impersonal *si* may appear to be a relatively trivial one, its implications are potentially quite significant and serve well to illustrate the difference between our theory and others. So it is important to note that by assuming that impersonal *si* is generated under INFL, Belletti (1982b) derives the fact that *si* is assigned a verb's external θ -role if there is one. However, this is at the very least no better than our stipulation that *si* must be nominative, if Belletti's assumptions about the position of impersonal *si* can only stand under ad hoc stipulations about INFL.

Independently of these considerations, as Burzio (1981) points out, Belletti also runs into problems in the movement cases, such as (4) or (5), where the verb assigns no external θ -role; for, in these cases, at D-structure *si* in INFL will not be in a θ -position. What Belletti proposes, then, is that *si* is inserted in S-structure; but this option significantly enriches the grammar. Similar considerations apply to Zubizarreta's (1982a) theory, which in a number of respects derives from Belletti's.

Burzio (1981) argues explicitly against assuming that impersonal *si* must be associated with nominative Case, and in favor of the assumption that *si* must simply be in subject position at some point of the derivation prior to cliticization. His argument is that *si* can be assigned accusative Case, as long as it is otherwise associated with a subject position; and he provides one piece of evidence to support this argument, namely examples of the type of (i):

- (i) *Ritengo essersi speso troppo*
I believe one to have spent too much

(i) however does not show what it purports to show, minimally because Italian *ritenere* does not have the subcategorization properties of English *believe*. On the contrary, as demonstrated in Rizzi (1982), lexical subjects of infinitival sentences embedded under *ritenere* and similar verbs are not accusative, as are the lexical subjects of infinitival sentences embedded under English *believe* and similar verbs, but rather are nominative.

Conversely, then, Burzio's prediction that impersonal *si* can be associated with an accusative position, as long as it is a subject position, is straightforwardly contradicted by the data in (9) and (10).

a verb; and it is obligatorily associated with the nominative Case. In other words, the lexical entry for impersonal *si* looks essentially like (16):

- (16) impersonal *si*: —free variable
 —argument
 —N
 —third-person, unspecified number and gender
 —clitic on a verb
 —nominative

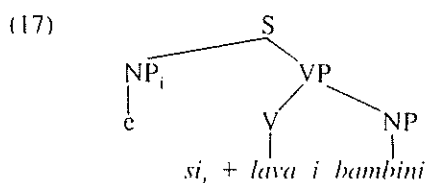
Let us assume that (16) is correct. Obviously, the next question to arise is what syntactic structures are associated with (16). If impersonal *si* is both a clitic on a verb and an argument, and in examples of the type of (1) or (3) it must furthermore be associated with the subject θ -position, then in (1), (3), and similar examples impersonal *si* must be bound to the subject and form a chain with it. What is more, by being bound to the subject and forming a chain with it, impersonal *si* can also be associated with the nominative Case.

Similarly, if in examples of the type of (4) and (5) impersonal *si* must be associated with the nominative Case, it must be bound to and form a chain with the subject in these examples as well. The only difference between examples of the type of (1) or (3) and examples of the type of (4) or (5) is that in (4) or (5) impersonal *si* must be associated with the object θ -position, hence must bind and form a chain with the object as well.

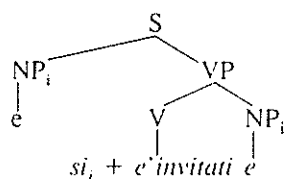
The subject position of all well-formed sentences involving impersonal *si*, on the other hand, including (1), (3), (4), and (5), must be filled by an element capable of binding impersonal *si* and forming a chain with it. Obviously, there is one and only one class of elements in the grammar which can satisfy this description, the class of expletives. Hence in examples like (1), (3), (4), or (5), the subject position must be filled by an expletive, which in Italian will be an empty category.

Furthermore, the object position of sentences like (4) or (5) must be filled by an element which can be bound by impersonal *si* and form a chain with it. But this can be simply a trace, very much as it would be in the familiar case of a chain between the object position and an accusative clitic.

In conclusion then, if the discussion so far is correct, sentences like (1) and (5) will be associated with structures like (17) and (18) respectively, where in both cases impersonal *si* is bound to and forms a chain with an expletive in subject position, and in (18) the chain also includes a trace in object position:



(18)



3. REFLEXIVE *si*

Consider now reflexive *si*. Strictly speaking, reflexive *si* has, as the name implies, the range of interpretation of a reflexive element. However it is also possible to find in Italian a *si* understood as having the same range of interpretation as a reciprocal element. If we call this *si* reciprocal *si*, it is easy to show that reciprocal and reflexive *si* have exactly the same distribution, with the only obvious exception that reciprocal *si*, contrary to reflexive *si*, must have a plural interpretation. For example, in (19) and in (20), both the reflexive and the reciprocal interpretation of *si* are possible: it is true that in (20) only the reciprocal reading is natural, but this has to do with pragmatic considerations which we disregard:

- (19) *I bambini si lavano*
The children wash themselves/each other

- (20) *I bambini si parlano*
The children talk to themselves/each other

If so, we can assume that there is an element *si* which has the range of interpretation of anaphoric elements in general; hence, in particular, it can have the interpretation of a reflexive element or the interpretation of a reciprocal element. This is what we call reflexive *si*, meaning reflexive and reciprocal *si*.

On the other hand, if impersonal *si* is equivalent in its interpretation to PRO_{anb} , reflexive *si* is equivalent, at least in its nonreciprocal part, to control PRO. Thus, if we say that impersonal *si*, as PRO_{anb} , is interpreted as a free variable, we can say symmetrically that reflexive *si* and control PRO, as well as reflexive and reciprocal elements in general, are interpreted as dependent variables. This means that these elements do not have any intrinsically defined interpretation, but are dependent for their interpretation on some antecedent.

Consider then the syntactic properties of reflexive *si*. It is easy to show that reflexive *si* must be an anaphor, and subject as such to Binding Condition A of Chomsky (1981). Indeed, by Binding Condition A, reflexive *si* must be bound in its governing category; if so, and assuming that it is associated with the unmarked definition of governing category (the one which is associated with En-

glish as well), the ungrammaticality of examples like (21) and (22) can be immediately explained⁴:

- (21) **I bambini mi videro lavarsi*
The children saw me wash themselves (each other)
- (22) **Dicono che si lavano i bambini*
They say that they-selves (each other) wash the children

In general, Binding Condition A will exclude along with (22) all examples in which reflexive *si* is in a nominative position, and along with (21), where reflexive *si* is bound over the subject of a small clause, all examples in which reflexive *si* is bound over the subject of any propositional structure.

But saying that reflexive *si* is subject to Binding Condition A and as such is required to be bound in its governing category is not sufficient to capture the restrictions on its antecedent. It is easy to notice that in all of the well-formed examples considered so far, reflexive *si* is associated with a subject antecedent. This is not accidental: on the contrary, as Rizzi (this volume) shows, all the examples in which reflexive *si* is bound by a nonsubject are ill-formed. For example, (23), where reflexive *si* is bound in its governing category by a direct object, is impossible:

- (23) **Si affidai i bambini*
I entrusted the children to themselves (each other)

The problem in (23), however, can be easily solved if the grammar has, in addition to the governing category parameter proposed in Manzini and Wexler (to appear), Wexler and Manzini (to appear), a proper antecedent parameter, under which a proper antecedent for the purposes of the Binding Conditions can either be a subject or any nominal at all.⁵ For if so, and if reflexive *si* is associated with the value *subject* of the parameter, the ungrammaticality of (23) and of all examples in which reflexive *si* is bound by a nonsubject is immediately explained. Furthermore, if we take the value *subject* of the parameter to be the

⁴ Note that we assume that the definition of governing category is parameterized. In particular, we accept the formulation of the governing category parameter in Manzini and Wexler (to appear), Wexler and Manzini (to appear), and the definition of markedness proposed there. The definition of governing category under the unmarked value of the parameter essentially corresponds to the definition of governing category in Chomsky (1981) or Chomsky (1986); the Binding Conditions are substantially unaltered from Chomsky (1981).

⁵ That the binding theory is parameterized with respect to the notion *proper antecedent* in the way indicated is argued in Manzini and Wexler (to appear) and Wexler and Manzini (to appear); that the value *subject* of the proper antecedent parameter is the unmarked value for anaphors also follows from the definition of markedness proposed there.

unmarked value, reflexive *si* will simply be associated with the unmarked values of all of the Binding Theory parameters.

Next, it is easy to show that reflexive *si*, as impersonal *si*, is an argument, and subject as such to the θ -Criterion. Consider for example the well-formed (19) or (20). In (19) and (20), reflexive *si* can be associated with the θ -positions subcategorized by the verb, the direct object, and indirect object position respectively; and unless reflexive *si* is associated with these positions, they cannot be associated with any argument at all. As in the case of impersonal *si*, furthermore, which θ -position reflexive *si* is associated with is immaterial. So in (19) it is associated with a direct object position and in (20) with an indirect object position; but nothing prevents associating it with a subject θ -position, as in (24):

- (24) *Si videro lavare i bambini*
They saw themselves (each other) wash the children

Similarly, it is natural to assume that, as impersonal *si*, reflexive *si* is associated with the categorial feature N, and as a lexical nominal is subject to the Case Filter; or, alternatively, that the chains that it forms at S-structure are subject to the Visibility Condition. In particular, it is easy to see that in all of the well-formed examples considered so far, reflexive *si* can be associated with a Case: it can be associated with the accusative Case in (19) and (24) and with the dative Case in (20). As with impersonal *si*, however, it is easy to notice also that not all Cases can be associated with reflexive *si*. Indeed, the restrictions on impersonal and reflexive *si* are exactly complementary; if impersonal *si* must always be associated with the nominative Case, reflexive *si* can never be associated with it. In the case of reflexive *si*, however, this obviously follows from independent reasons. For, as we saw in (22), all examples where reflexive *si* is associated with the nominative Case are excluded by Binding Condition A.

Next, consider the agreement features of reflexive *si*. Once more it is not difficult to show that the agreement features associated with reflexive *si* are identical to those associated with impersonal *si*. In particular, the third-person feature of impersonal *si* must also be a feature of reflexive *si*, for reflexive *si* can never be bound to a first- or second-person antecedent, as in (25)–(26):

- (25) **Si laviamo*
We wash ourselves (each other)
- (26) **Si lavate*
You wash yourselves (each other)

On the other hand, the number and gender of the antecedent are clearly immaterial; but if we assume that reflexive *si*, as impersonal *si*, is associated with unspecified number and gender features, this also follows.

Finally, given that reflexive *si* is phonologically a clitic on a verb, the problem

of its syntactic position arises. In the case of reflexive *si*, however, there seems to be no alternative but to assume that, as impersonal *si*, it is a clitic on a verb syntactically as well as phonologically. Needless to say, the evidence from the surface distribution of reflexive *si* is fully compatible with this assumption: reflexive *si* surfaces after the negation particle, together with the other complement clitics, as in (27):

- (27) *I genitori non se li affidano*
 The parents do not entrust them to themselves (each other)

Let us sum up so far. If what precedes is correct, reflexive *si* is associated with the following properties: it is interpreted as a dependent variable; it is syntactically an anaphor, an argument, and an N; it is third person; it has unspecified number and gender features; and it is a clitic on a verb. Hence, the lexical entry for reflexive *si* looks essentially like (28):

- (28) reflexive *si*: —dependent variable
 —anaphor
 —argument
 —N
 —third-person, unspecified number and gender
 —clitic on a verb

The next problem to arise is what syntactic structures are associated with (28). Now, if reflexive *si* is an anaphor and is associated with the unmarked values of the Binding Theory parameters, we know that it will have to be bound in the unmarked governing category by a subject. Furthermore, if reflexive *si* is a clitic on a verb and an argument, it will also have to form a chain with a θ -position. If so, the structure of examples of the type of (19) will look like (29), where reflexive *si* is bound by the subject and referentially dependent on it, while it binds and forms a chain with the object position:

- (29)
-
- ```

graph TD
 S --> NP_i1[NP_i]
 S --> VP
 NP_i1 --> i_bambini["i bambini"]
 VP --> V
 VP --> NP_i2[NP_i]
 V --> si_lavano["si_i + lavano"]
 NP_i2 --> e["e"]

```

Having a reasonably complete picture of reflexive *si*, it is possible to explain one last remaining puzzle. As we have seen, reflexive *si* must have a subject antecedent: even this, however, is not quite an adequate description of the facts. The antecedent for reflexive *si* must not only be a subject, but also a nonderived subject; so an example like (30), where reflexive *si* is bound by a derived subject, is obviously ill-formed:

- (30) \**I bambini si furono affidati*  
The children were entrusted to themselves (each other)

For this problem, we can now adopt the solution proposed by Rizzi (this volume). Consider what the structure of (30) looks like, as in (31):

- (31) [*I bambini*<sub>i</sub> [*si*<sub>i</sub> + *furono affidati* *e*<sub>i</sub> *e*<sub>i</sub>]]

In (31), for *i bambini* to be correctly associated with a  $\theta$ -position, it must form a chain with the direct object empty category. The empty category, however, is not locally bound by *i bambini*, but rather is locally bound by reflexive *si*. Hence, because of the lack of local binding, *i bambini* and the empty category cannot form a chain. But if so, the  $\theta$ -Criterion is violated. Hence, obviously, (31) is excluded, and the same is true in general of all examples which involve a derived subject and a reflexive *si* in the same sentence.<sup>6</sup>

#### 4. IMPERSONAL AND REFLEXIVE *SI* UNIFIED

Consider now once more the lexical entries we arrived at for impersonal and reflexive *si*. If these entries are correct, then obviously the two *si*'s share a number of properties. In particular both impersonal and reflexive *si* are arguments and nominals, both are syntactically clitics on a verb, and both have third person, unspecified number and gender features.

Not less obviously, however, the two items differ in at least two separate properties. Although impersonal *si* is interpreted as a free variable, reflexive *si* is interpreted as a dependent variable. Syntactically, furthermore, impersonal *si* is associated with nominative Case, while a seemingly unrelated property of reflexive *si* is that it is an anaphor, subject as such to Binding Condition A.

But consider impersonal *si* again. If (16) is correct, impersonal *si* must be associated with nominative Case. As (17) and (18) show, however, in those instances in which impersonal *si* is, in fact, associated with nominative Case, it

<sup>6</sup> Rizzi (this volume) also proposes a different explanation for the data in (23). Rizzi's explanation starts with the observation that in the examples of the type of (23) not only is *si* bound by its antecedent, but it also binds the antecedent. Hence if the antecedent for *si* is an R-expression, the configuration is excluded by Binding Condition C; if the antecedent is a pronominal, then the configuration is excluded by Binding Condition B.

But consider (i). Under Rizzi's explanation of the *si* data, (i) can only be predicted to be grammatical if the preposition *a* can define a maximal projection, and hence prevent reciprocal c-command between the lexical anaphor and its antecedent:

(i) *Affidai i bambini a se stessi*  
I entrusted the children to themselves

Giorgi (1985), however, argues that the preposition *a* never defines a barrier for c-command. Obviously if she is correct, Rizzi's explanation of the *si* data cannot be right.

necessarily forms a chain with its subject. Hence for all well-formed examples, saying that impersonal *si* is associated with nominative Case or saying that it forms a chain with its subject is perfectly equivalent.

In those instances on the other hand in which impersonal *si* cannot be associated with nominative Case it is easy to show that it cannot form a chain with its subject either. Consider first (8) and (9). In both (8) and (9), the subject of impersonal *si* is an argument; hence a chain containing both impersonal *si* and its subject contains two arguments, in violation of the  $\theta$ -Criterion. Consider then (10) and the structure for (10) in (32):

(32) *Vedo* [ $e_i$  *lavarsi*<sub>*i*</sub> . . .

In (32) it is not difficult to see that the empty category cannot be a trace, since it heads its own chain, ( $e_i$ ,  $si_i$ ); cannot be a PRO, since it is governed; cannot be a variable, since it is not bound by any operator; and finally cannot be a *pro*, since it is not properly identified. Hence the empty category in (32) gives rise to ungrammaticality.

Suppose then that in the lexical entry for impersonal *si* in (16) the property of being associated with nominative Case is substituted with the property of being bound to, and forming a chain with, its subject, as in (33):

(33) impersonal *si*: —free variable  
           —argument  
           —N  
           —third-person, unspecified number and gender  
           —clitic on a verb  
           —bound to and forms a chain with its subject

The lexical entries for impersonal and reflexive *si* appear now to differ syntactically in that while impersonal *si* is bound to, and forms a chain with, its subject, reflexive *si* is an anaphor, associated with the unmarked values of both the governing category and the proper antecedent parameter. But, it seems, the application of Binding Condition A amounts in this instance to saying that reflexive *si* must be bound by its subject.

Suppose then that we revise the lexical entry for reflexive *si* as in (34), so that reflexive *si* is associated with the property of being bound to, and referentially dependent upon, its subject, rather than with the property of being an anaphor:

(34) reflexive *si*: —dependent variable  
           —argument  
           —N  
           —third-person, unspecified number and gender  
           —clitic on a verb  
           —bound to and referentially dependent on its subject

While the lexical entries for impersonal and reflexive *si* in (16) and (29) seemed to reveal an incompatibility of syntactic features, the new lexical entries in (33) and (34) reveal a remarkable similarity. Clearly both impersonal and reflexive *si* are bound to their subjects, and only one difference persists. While impersonal *si*, once bound to its subject, forms a chain with it, reflexive *si*, once bound to its subject, assumes a referential dependence relation to it.

Similarly, it is now easy to see that while impersonal and reflexive *si* differ in their semantic interpretation, the two elements are also remarkably similar in this respect. As the lexical entries in (16) and (29) or in (33) and (34) emphasize, both impersonal and reflexive *si* have a variable-like character, with only one difference; while impersonal *si* is free in its interpretation, reflexive *si* depends in its interpretation upon some antecedent.

Disregarding the remaining differences between impersonal and reflexive *si*, suppose we unify their two lexical entries into one, taking as the properties of the new lexical item all and only the properties that impersonal and reflexive *si* have in common. What we obtain is the unified lexical entry for impersonal/reflexive *si* in (35):

- (35) impersonal/reflexive *si*: —variable  
                                           —argument  
                                           —N  
                                           —third-person, unspecified number and  
                                                                                           gender  
                                           —clitic on a verb  
                                           —bound to its subject

Consider the property of the new impersonal/reflexive *si* of being bound to its subject. One might assume that the relation *binding* subsumes the two different relations of chain formation and referential dependence, and only those two, quite independently of the theory of *si*. If so, it follows quite independently of the theory of *si* that the binding relation between the unified impersonal/reflexive *si* and its subject can be taken either as a chain formation relation or as a referential dependence relation, and only as one of those two relations. So, depending upon the choice of one or the other option, the syntactic properties of impersonal and reflexive *si* respectively can straightforwardly be obtained out of the unique properties of impersonal/reflexive *si*.<sup>7</sup>

<sup>7</sup> Note that in general the account of impersonal and reflexive *si*, as well as of the unified impersonal/reflexive *si* given here, follows closely Manzini (1983). However in Manzini (1983), rather than assuming that the binding relation could be interpreted as either a chain formation or a referential dependence relation, it was assumed that the grammar included two different binding relations notated by cosubscripting and cosuperscripting respectively. In Manzini (1983) it was assumed further that binding by cosuperscripting was restricted to expletive-argument pairs; but the natural step, of course, was to extend it to chain formation relations. If so, however, the binding

Similarly, consider the variable property of impersonal/reflexive *si*. In this respect as well, we can assume that two different options are available, and only two. In particular, one can take the variable to be dependent upon some antecedent or to be a free variable; in the first case one obviously obtains the interpretation typical of reflexive *si*, in the second case the interpretation typical of impersonal *si*. Hence, the different semantic properties of impersonal and reflexive *si*, and only those, can be derived from the properties of the unified impersonal/reflexive *si*.

There is one problem with the approach we are taking. In the lexical entries for impersonal and reflexive *si* in (33) and (34) the free variable interpretation is associated with the property of forming a chain with the subject, the dependent variable interpretation with the property of being referentially dependent upon the subject. Once impersonal and reflexive *si* are unified as in (35) a problem arises of excluding a mismatch of semantic and syntactic properties.

In one case the problem has a straightforward solution. Clearly, if *si* is referentially dependent upon its subject it is necessarily interpreted as a dependent variable. But suppose that *si* forms a chain with its subject; the question is what prevents it from being referentially dependent upon some element other than its subject, and hence having a dependent variable interpretation. Consider concretely an example like (22) again. In (22) *si* obviously forms a chain with its subject; the question is what prevents it from being interpreted as a dependent variable in the way indicated there. Clearly, it is not the fact that the dependent variable interpretation is lexically associated with referential dependence on the subject, as in the entry for reflexive *si* in (34).

Now, in discussing reflexive *si* we treated the property of being an anaphor and the property of being a dependent variable as two different properties, a syntactic and an interpretive property respectively. The natural assumption however is that the syntactic property of being an anaphor reduces to the property of having a dependent variable interpretation. If so, a dependent variable interpretation will only be available for a given element if that element has been subjected to Binding Condition A. But if so, (22) and similar examples are excluded as violations of Binding Condition A. In this case, the values of the governing category and the proper antecedent parameter will have to be taken to be the unmarked values.

In summary, there are no two lexical items impersonal and reflexive *si*, but one single lexical item impersonal/reflexive *si* which subsumes both, and from

---

conditions would have to be reformulated to cover two types of binding, to account for the local binding of traces, clearly not a case of referential dependence. But this amounted to saying that there is one syntactically relevant notion of binding, as seen in the Binding Conditions, though this can have two distinct interpretations, and this is obviously the position adopted here. This position is also explicitly taken in Brody (1984).

On this point I benefited from conversations with H. Borer and M. Brody.

which both can be obtained. The diversification of impersonal and reflexive *si* from the unified lexical entry for impersonal/reflexive *si* originates from two properties, the variable interpretation and the property of being bound to the subject. But differences with respect to these two properties obviously imply a number of other differences. In particular, the complementary Case properties of impersonal and reflexive *si* follow directly from their different binding properties.

Similarly, one can notice that impersonal and reflexive *si* have different positions in the surface string of clitics with respect to other complement clitics: impersonal *si* follows an accusative clitic, as in (12), while reflexive *si* precedes it, as in (27). But this follows in turn from their different Case properties, since one must independently assume that clitics are ordered with respect to their Case.<sup>8</sup>

## 5. THE OTHER *SI*'s

In the sections above we illustrated the properties of impersonal and reflexive *si* and demonstrated that impersonal and reflexive *si* are one single lexical item. However impersonal and reflexive *si* do not exhaust the varieties of *si*'s found in Italian. In particular, beside impersonal and reflexive *si*, at least one third major type of *si* is generally recognized to exist: middle *si*.

Middle *si* has at least one property that neither impersonal nor reflexive *si* has. Consider for example the middle *si* construction in (36):

- (36) *I bambini si lavano volentieri*  
The children wash (middle) gladly

In (36) it is easy to show that *i bambini* is in subject position. For, *i bambini* has all the behaviors of a nominative subject: in particular it agrees with the verb and it can be substituted by a null subject. On the other hand, it is also evident from the interpretation of (36) that *i bambini* is associated with the object  $\theta$ -position. This means then that *i bambini* is in subject position at S-structure and in object position at D-structure. But if so, middle *si* has the same effect on the verb as a

<sup>8</sup> We argued above that Belletti's (1982b) idea that impersonal *si* is generated under INFL has no advantages over ours in terms of the number of stipulations needed to account for the behavior of impersonal *si*, and potential disadvantages in terms of the nature of those stipulations. Now we can add that if impersonal *si* is generated under INFL and reflexive *si* is generated on a verb, the properties of impersonal and reflexive *si* can hardly be unified in a single lexical entry. Similarly, we saw above a number of problems with Burzio's (1981) analysis in terms of cliticization from the subject position. Now of course we can add that it is difficult to see under Burzio's analysis how impersonal and reflexive *si* could be unified at all.



passive morphology element, inducing movement from the object into the subject position.

Apart from its passivizer property, the properties of middle *si* seem essentially identical to those of impersonal *si*. Suppose then that all the properties of impersonal *si* are also properties of middle *si*. If so, the lexical entries for impersonal and middle *si* can be straightforwardly unified. In particular, the new lexical entry for impersonal/middle *si* will have all the properties of the entry for impersonal *si* and an additional optional passivizer property, as in (37); impersonal *si* will then be obtained when the passivizer property is not realized, middle *si* when the passivizer property is realized:

- (37) impersonal/middle *si*: —free variable  
                                   —argument  
                                   —N  
                                   —third-person, unspecified number and gender  
                                   —clitic on a verb  
                                   —bound to, and forms a chain with, its subject  
                                   —(passivizer)

Suppose that what precedes is on the right track. The next question to arise is whether the partially unified lexical entries for impersonal/middle *si*, as in (37), and for impersonal/reflexive *si*, as in (35), can be further unified into one lexical entry for all of the three types of *si*. There are two obvious ways to go about doing this: to generalize the variable and binding properties of impersonal/middle *si*, as in (37), on the model of impersonal/reflexive *si*, or equivalently to add to impersonal reflexive *si*, as in (35), an optional passivizer property, on the model of impersonal/middle *si*. In both cases the generalized lexical entry for *si* in (38) is obtained:

- (38) *si*: —variable  
                   —argument  
                   —N  
                   —third-person, unspecified number and gender  
                   —clitic on a verb  
                   —bound to its subject  
                   —(passivizer)

From (38) impersonal and reflexive *si* can be obtained if the passivizer property is not realized, middle *si* if the passivizer property is realized. Impersonal and middle *si* are obtained if *si* is taken to be a free variable and to form a chain with its subject, and reflexive *si* if *si* is taken to be a dependent variable and to be referentially dependent upon its subject.

But if the generalized lexical entry for *si* in (38) is correct, a fourth type of *si* is

predicted to exist, a *si* which on the one hand is a dependent variable and referentially dependent on its subject, as reflexive *si* is, and on the other hand is a passivizer, as middle *si* is. Obviously, if we prove that this fourth type of *si*, call it middle-reflexive *si*, does exist, our theory will receive strong confirmation.

Consider, then, an example like (39), noticed in Burzio (1981):

- (39) *Gli unici bambini lavatisi*  
The only children (who) washed themselves

In (39) *si* is embedded in an expression headed by a *-en* type participle which is interpreted essentially as a restrictive relative. Now, this type of expressions, which we will call a *small clause relative*, is restricted to a narrow class of verbs. In particular, as noticed in Burzio (1981), a small clause relative can be headed by a passive participle, as in (40), and by an ergative participle, as in (41). But a small clause relative can apparently never be headed by the participle of an intransitive verb, as in (42), nor by the participle of a transitive verb, as in (43); that the object of the transitive verb is made into a clitic, as in (44), or into a reflexive, as in (45), seems to be equally immaterial:

- (40) *Gli unici bambini lavati dai genitori*  
The only children washed by the parents
- (41) *Gli unici bambini arrivati in tempo*  
The only children arrived on time
- (42) \**Gli unici bambini telefonato in tempo*  
The only children telephoned on time
- (43) \**Gli unici genitori lavati i bambini*  
The only parents washed the children
- (44) \**Gli unici genitori lavatili*  
The only parents washed them
- (45) \**Gli unici bambini lavati se stessi*  
The only children washed themselves

The generalization with small clause relatives seems to be that their participle heads must be associated with movement from the object into the subject position; so it is indeed for the passive in (40) and the ergative in (41), but neither for the intransitive in (42) nor for the transitives in (43)–(45).

Consider however (39) again. *Si* in (39) is interpreted as a dependent variable; hence the possibility that *si* in (39) can belong to the impersonal or the middle type is excluded. But if the *si* in (39) belongs to the reflexive type, then it is impossible to explain why (39) is well-formed rather than ill-formed, as (45) is. Suppose, however, that the *si* in (39) belongs to the middle-reflexive type. If so,

the well-formedness of (39) can straightforwardly be explained, for the *si* in (39) can both have a dependent variable interpretation, and as a passivizer, make the verb into a NP-movement verb, compatible as such with the small clause relative construction. Thus examples of the type of (39) provide evidence in favor of the existence of middle-reflexive *si*.

A similar point can be made with a different class of examples, also from Burzio (1981). The genitive clitic *ne* of Italian can only cliticize from within an object nominal. Now, consider the counterparts to our standard examples of middle and (middle-)reflexive *si* with the subject in postverbal rather than preverbal position. In the case of middle *si*, we expect the subject to be able to occupy the object position; hence we expect *ne* cliticization to be possible from the postverbal subject. As it turns out, however, *ne* cliticization is possible from the postverbal subject also in constructions with reflexive interpretation, as in (46):

- (46) *Se ne lavano molti*  
Of them wash ('are washed'/'wash themselves') many

Obviously this can only be explained under the assumption that the *si* in (46) and similar examples is a middle-reflexive *si*; if so, indeed, we expect the postverbal subject to be in object position, and *ne* cliticization to be possible from it.<sup>9</sup>

Summing up, there is one single lexical item *si* from which four different types of *si* can be obtained, impersonal, reflexive, middle, and middle-reflexive *si*. If *si* is taken to be a free variable and the optional passivizer property is not realized, impersonal *si* is obtained; if the passivizer property is not realized and *si* is taken to be a dependent variable, reflexive *si* is obtained. If on the other hand *si* is associated with both the free variable property and the optional passivizer property, middle *si* is obtained; and finally if *si* is associated with both the optional passivizer property and the dependent variable property, middle-reflexive *si* is obtained. This situation is summarized in (47)

(47)

|               | Free variable        | Dependent variable         |
|---------------|----------------------|----------------------------|
| Nonpassivizer | Impersonal <i>si</i> | Reflexive <i>si</i>        |
| Passivizer    | Middle <i>si</i>     | Middle-reflexive <i>si</i> |

<sup>9</sup> In fact, Burzio (1981) introduces a distinction like our distinction between reflexive and middle-reflexive *si*, while speaking exclusively about reflexive *si*. Burzio's idea is that in one reflexive construction *si* spells-out the accusative Case features and receives the object  $\theta$ -role through a chain with an empty category in object position; in another reflexive construction *si*, while still spelling-out accusative Case, also absorbs the subject  $\theta$ -role, so that a chain must form between the object and the subject position. In this respect the similarity between Burzio's theory and ours is apparent.

But the types of *si* which we have taken into account so far do not yet exhaust the types of *si* found in Italian. There is however one major difference between the types of *si* that we have taken into account and the ones that we have disregarded. All the types of *si* that we have taken into account are fully productive; the other *si*'s are semiproductive at best. In particular a sizable number of transitive verbs have their ergative counterpart formed by adding what Burzio (1981) terms ergative *si*, for example *spegnere/spegnersi* in (48a) and (48b), respectively; however, an equally sizable number of transitive verbs have phonologically identical ergative counterparts, as for example *affondare/affondare* in (49);

- (48) a. *Spensi la luce*  
I turned off the light  
b. *La luce si spense*  
The light went off
- (49) a. *Affondammo la nave*  
We sank the ship  
b. *La nave affondo'*  
The ship sank

From this we conclude that ergative *si* differs from the types of *si* investigated so far in that it is attached to a verb not in the syntax but in the lexicon. Borer (1984a) has suggested that affixes can be attached either in the morphology or in the syntax, but because the lexicon and the syntax are subject to different laws, the set of properties associated with the affix will be realized in different ways in the lexicon and in the syntax. Suppose then that, pursuing the logic of this essay, we assume that ergative *si* is nothing but the unified *si* in (38) as attached in the lexicon rather than in the syntax.

Plausibly, what establishes the affix-like character of *si* is its passivizer property; hence we can assume that *si* is always a passivizer when it is attached in the lexicon. Of the other properties of *si*, on the other hand, we can assume that its semantic properties are meaningless when it is attached in the lexicon, as are its argument and categorial properties. If so, when attached in the lexicon *si* is left with its agreement and clitic properties and with the property of being bound to its subject. We might assume that binding properties are meaningfully associated with an affix as well as with an independent lexical element; however, while binding can be interpreted as either a chain formation relation or a referential dependence relation in the case of independent lexical elements, the interpretation available for binding in the case of morphology elements like ergative *si* is realized simply as agreement.

The properties of ergative *si* predicted above can be argued to correspond to its observed properties. So ergative *si* is only compatible with a third-person subject: this, of course, is explained if ergative *si* agrees with its subject and is furthermore associated with a third-person feature. Furthermore, according to

Borer (1984a), an element with passivizer properties inserted in the syntax has the effect of blocking the assignment of the external  $\theta$ -role in the subject position; an element with passivizer properties inserted in the the lexicon, on the other hand, has the effect of eliminating the subject  $\theta$ -role altogether. The latter is clearly the effect of ergative *si*.

At this point, one last remark is in order. We have shown that different types of *si* can be unified into one single lexical item. But does the theory that we propose for Italian *si* extend to other languages as well? It appears to extend to languages which have elements with exactly the same range of constructions as Italian *si*, for example, Spanish. But in a number of languages the situation is rather more complicated in that an element of the language seems to have some, but not all, the constructions attested by Italian *si*. For example, French *se* clearly has a middle construction and a middle-reflexive construction, but not an impersonal construction. Icelandic *-st* seems to have a middle and a middle-reflexive construction but neither the impersonal nor the reflexive one.

Consider French. Given our theory, it is difficult to see how the lack of the impersonal construction could be accounted for independently of some difference between the lexical entries for *si* and *se*. If the two lexical entries do differ, however, it is easy to see that there is a quite straightforward parameter with respect to which they could differ and produce exactly the desired results. Suppose that the optional passivizer property in the entry for Italian *si* is an obligatory property in the entry for French *se*. This means, in terms of the table in (47), that French *se* can have neither an impersonal nor a reflexive construction, while it can have a middle and a middle-reflexive one; but this, in turn, is exactly what the data seem to indicate. Thus the lexical entry we propose for Italian *si* can be straightforwardly parameterized to extend to French *se*. In the same way, we assume that our theory of Italian *si* can be extended to *si*-like elements with the same range of constructions as French *se*, for example, Icelandic *-st*.<sup>10</sup>

<sup>10</sup> The parameterization issue also offers an opportunity to refer to a number of recent works which are concerned, rather than with Italian or similar languages, with French, Icelandic, and so on. These include Grimshaw's (1982) account of French *se* and Andrews' (1982) account of Icelandic *-st*, both in a Lexicalist framework; Wehrli's (this volume) account of French based on argument absorption, similar in a number of respects to Grimshaw's but not necessarily dependent on a Lexicalist framework; and Marantz's (1984) account of a number of languages, including French and Icelandic, in a Chomsky-type framework. Minimally, while the approach we take easily allows for parametric variation, it is not entirely clear how the approach these works take could be extended from French or Icelandic to Italian without abandoning the unification of the different types of *si*. Indeed, that impersonal *si* is unmanageable in terms of affixation is explicitly recognized in Grimshaw (1982); and a treatment of *si*-like elements in terms of affixation is at the core of all of the proposals mentioned. Otherwise, for what we consider to be definitive criticisms of Lexicalist approaches, the reader is referred to Marantz (1984).

Finally, our own treatment of middles, ergatives, and parameterization here is only partial. For an essentially complete account compatible with the theory in this essay, the reader is referred to Manzini (1983).

## ACKNOWLEDGMENTS

This essay developed over the past few years as part of a more general project on Italian *si*, middle constructions, and restructuring and reanalysis in Romance syntax, and was influenced in its development by a substantial number of people. In particular, the material presented here was discussed in courses at UC Irvine and at UCLA in the winter and spring 1984, and in seminars at University College London and SOAS in the autumn of the same year. Among the members of the audiences there, I would like to thank especially H. Borer, N. Elliott, O. Jaeggli, and M. Brody. I would also like to thank the Program in Romance Literatures and Linguistics at UCLA, and especially C. Otero, the Cognitive Science Program at UC Irvine, and P. Culicover and K. Wexler, for making both the Irvine and the UCLA courses possible. The material presented here is mostly an elaboration of parts of Manzini (1983).