

THE INTERPRETATION AND ACQUISITION OF ITALIAN IMPERSONAL *SI*

NINA HYAMS

Department of Linguistics
University of California—Los Angeles
Los Angeles, California 90024

1. INTRODUCTION

In this essay I have two related concerns. The first is to provide an analysis of Italian impersonal *si*, illustrated in the examples in (1). The second is to examine the acquisition of *si* by Italian-speaking children.

- (1) a. *In questa citta si va spesso al cinema*
'In this city one goes (3rd per. sing.) often to the movies'
b. *In questa pasticceria si mangia soltanto i dolci al cioccolato*
'In this pastry shop one eats (3rd per. sing.) only chocolate cookies'
c. *I dolci al cioccolato si mangiano in questa pasticceria*
'Chocolate cookies *si* eat (3rd per. plu.) in this pastry shop'
(= Chocolate cookies are eaten in the pastry shop)
d. *Si mangiano i dolci al cioccolato in questa pasticceria*
'*si* eat (3rd per. plu.) chocolate cookies in this pastry shop'
(= 1c)

As is clear from the English translations of the above sentences, *si* enters into two constructions. The first type of construction, exemplified by the examples in (1a,b), is an active impersonal sentence in which *si* appears to function as a subject with arbitrary reference, roughly equivalent in interpretation to arbitrary

PRO or the English 'one.' In this case *si* may appear with either an intransitive or transitive verb. There is no overt full NP subject and the verb is marked in the "neutral" third-person singular form. Henceforth, I refer to this construction as the 'impersonal *si*.' The sentences in (1c,d) illustrate the second type of construction containing *si*. In this instance *si* does not bear a grammatical function of subject or object. Rather, as noted in many traditional and generative analyses of this construction, *si* appears to function as a passivizing element and the sentence receives a passive (or middle) interpretation. This construction occurs only with transitive verbs; the direct object functions as a subject [either preverbal as in (1c) or postverbal as in (1d)], triggering agreement on the verb. In contrast to the English middle, the Italian middle *si* construction is fully productive; that is, any transitive verb which selects a human subject may undergo the rule. Following Manzini (this volume) I will refer to this construction as the 'middle *si*.'¹

In the Sections 2 and 3 I propose an analysis of the two constructions as they function in the adult grammar of Italian. Anticipating that discussion somewhat, I will adopt a model of morphology in which rules of word formation are free to apply in either the syntax or the lexicon—subject to independent principles of grammar (Borer 1984a). I will argue that *si* may undergo word formation with a predicate in either component: the syntactic process yields the basic properties associated with the impersonal *si* construction, while the lexical process is responsible for the middle construction.

In Section 4 I address the acquisition issue. The acquisition data associated with *si* are of particular interest in that they provide evidence for the hypothesis that the analyses made by children in the course of grammatical development, even when "wrong" from the point of view of the adult system, nevertheless fall within the range of possible analyses predicted by a restrictive theory of grammar. Such results broaden the empirical base of linguistic theory and lend support to the "instantaneous" model of acquisition assumed within Generative Grammar.

2. THE MORPHOLOGY OF *SI*

I begin this section by outlining the basic model of morphology I assume, which provides the theoretical framework within which to analyze *si*.

As noted by Borer (1984a), recent research in grammatical theory centers around the assumption that rules of grammar which map one representation onto

¹ In many traditional and generative analyses, this construction is referred to as the "impersonal passive." Although I have chosen to translate the Italian middle construction as an English passive for ease of exposition, no theoretical claim is intended. While I believe that the middle *si* construction shares certain properties with the passive, it also differs from the latter in a number of respects (see Footnote 16).

another should be allowed to apply freely, subject to conditions of well-formedness which constrain representations at distinct levels, or components. Following in this spirit, Borer proposes that rules of word formation, which map morphemes onto words, may apply freely in either the syntax or the lexicon subject to the principles of well-formedness which constrain lexical (presyntactic) and syntactic representations. Prominent among the syntactic constraints is the Projection Principle (Chomsky 1981), given in (2).

(2) *The Projection Principle*

Lexical features must be represented at every syntactic level.

In effect, the Projection Principle prevents syntactic rules from changing lexical specifications. Such specifications include (minimally) θ -features and Case features associated with a lexical head.²

Thus, given the Projection Principle as a constraint on word-formation rules which apply in the syntax, it is clear that those word-formation rules which destroy (or add) θ -features or Case features will be confined to the lexicon. The proposed system thereby derives the essence of the Lexicalist Hypothesis (Chomsky, 1970), that is, the hypothesis that certain morphological rules must apply in the lexicon. The theory also provides a characterization of the traditional distinction between "derivational" and "inflectional" processes: inflectional processes are those which obey the Projection Principle, while derivational processes are those which would violate the Projection Principle if they applied in the syntax.³

Although the above discussion is sketchy, it is sufficient for our present purposes. Further details of the theory will be discussed as they become relevant. Let us now turn to the analysis of *si*.

We may assume that *si* is lexically specified as having the properties given in (3).⁴

- (3) a. *si* selects a [+V, -N] stem.
b. *si* absorbs a subject θ -role.

² Further relevant lexical features include subcategorization frames, syntactic category features (e.g., *hit* is [+V, -N]), and possibly semantic features which specify the meaning of a particular lexical item. As noted by Borer, these are the 'inherent' lexical features of Chomsky (1965).

³ The correspondence between syntactic word-formation rules and traditional "inflectional" processes, on the one hand, and presyntactic word-formation rules and traditional "derivational" processes, on the other hand, is not, however, one-to-one. While the Projection Principle blocks those "derivational" processes which change inherent features from applying in the syntax, those "derivational" processes which preserve lexical features are free to apply in the syntax. Similarly, "inflectional" rules, which always obey the Projection Principle, are in principle free to apply in either component.

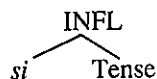
⁴ While there is no doubt a more elegant notational system for specifying the lexical properties of *si*, the list in (3) is adopted for ease of exposition.

- c. *si* absorbs a Case feature.
 d. *si* bears the nominal features [+3rd person, +plural, +masculine].

The specification given in (3d) is the set of inherent features associated with *si*, while the set of properties given in (3a–c) constitute the structural description of the word-formation rule that *si* enters into. I assume—as seems natural—that the requirements of *si* given in (3a–c) must be satisfied; that is, *si* cannot enter into a word-formation process with a predicate which fails to satisfy these requirements. Given that rules of word-formation may apply in either the syntax or the lexicon, these requirements may be satisfied at either level—modulo independent principles which constrain representations at these levels. In the section that follows, I discuss the two word-formation processes which affect *si*; the lexical process and the syntactic process.

2.1. Word-Formation in the Syntax

In this section I consider how the requirements of *si*, listed in (3), can be satisfied in the syntax given the principles which govern that component. Note first that in principle the requirement that *si* select a [+V, –N] stem [see (3a)] can be satisfied in two ways: *si* may affix directly onto the verb or alternatively *si* may appear in INFL as a realization of AG(reement), as suggested in Belletti (1982b). The structure of INFL is given:



In the above case the stem requirement of *si* is satisfied on the mapping from S-structure to PF when the contents of INFL undergo affix hopping. Note, now, that in order to satisfy the thematic requirement given in (3b), *si* will have to appear in INFL. To see why this is so I must diverge briefly to discuss θ -role assignment.

It is typically assumed that the subject θ -role is compositionally determined by the elements in VP.⁵ Thus, the θ -role assigned to *John* in (4a) is agentive while that in (4b) is nonagentive under the most natural interpretation of the sentence, in which *John* did not deliberately or willingly break his arm. (See Marantz 1984 for detailed discussion of this issue.)

- (4) a. John broke the television when the Lakers lost the game
 b. John broke his arm in the automobile accident

⁵ Although the subject θ -role is compositionally determined by the VP in the syntax, I also assume that there is a default or unmarked subject θ -role associated with the verb in its lexical entry (for example, *agentive* for the verb). This will be discussed shortly.

Given the above, it seems reasonable to assume that the subject θ -role is assigned by the VP to the [NP,S] position. Thus in order to receive the subject θ -role, *si* must be somehow associated with the subject position. This association can be established straightforwardly if *si*, by virtue of its position in INFL, identifies a null pronominal (*pro*) in subject position. In fact we may assume that the pair *pro* . . . *si* form a discontinuous pronoun, roughly along the lines proposed in Borer (1983b) for object clitics and the empty category in direct object position. Note, moreover, that *pro* is also assigned nominative Case (which I assume to percolate up from tense to the INFL node) and hence the Case requirement of *si* is also satisfied by virtue of its relation with the [NP,S] position.

Thus, the output of syntactic word-formation results in a subject clitic construction with the following properties.

- (5) a. *si* appears in INFL.
 b. *si* bears a compositional subject θ -role.
 c. *si* bears nominative Case.
 d. *si* identifies *pro* in subject position.

In Belletti (1982b) it is argued that the properties given in (5a–c) are those which characterize the impersonal *si* construction [cf. (1a,b)]. Similarly, Burzio (1981) and Manzini (this volume) have proposed that the impersonal *si* receives nominative Case and a subject θ -role, though on their analyses *si* is generated inside the VP.

Evidence for the clitic status of *si* in the impersonal construction is provided by the fact that it exhibits a range of properties typically associated with pronominal clitics in Romance languages.⁶ First, as illustrated below in (6a), *si* may not be separated from the verb by lexical material; second, it may occur between an object clitic and the verb, as shown in (6b); finally, *si* may not co-occur with a lexical subject, exemplified in (6c). With respect to the latter property, the behavior of *si* is entirely parallel to that of object clitics in Italian, which may not co-occur with a lexical object, as in (6d).

- (6) a. **Si non va al cinema stasera*
 (cf. *non si va al cinema stasera*)
 'One isn't going to the movies tonight'
 b. (*Le montagne*) *le si vede bene da qua*
 '(The mountains) them one sees well from here'
 c. **La gente si va spesso al cinema*
 'People one goes often to the movies'

⁶ The discussion has been greatly informed by various studies of the impersonal construction within the Government-Binding framework, in particular, Belletti (1982b), Burzio (1981), Zubizarreta (1982a), Manzini (1983), and Jaeggli (to appear).

- d. **Lo vedo Gianni*
'(I) him see Gianni'

The impersonal *si* differs from object clitics in that it requires nominative Case. Thus under the hypothesis that *si* is a subject clitic, the ungrammaticality of (6c) follows as an effect of the Case filter, entirely parallel to the situation in (6d). The Case status of *si* is further supported by the fact that it may not appear in infinitivals, with either control or raising verbs, as illustrated below.⁷

- (7) a. **E facile andarsi al cinema in questa città*
'It is easy (for) one to go to the movies in this city'
b. **Giovanni sperava di trovarsi la soluzione*
'Giovanni hoped (for) one to find the solution'
c. **In questo periodo sembra partirsi troppo spesso*
'In this period one seems to leave too often'

Let us now briefly discuss the status of the empty element in the subject position of the impersonal construction. As noted in (5d), the impersonal *si* properly identifies *pro*. I take this instance of *pro* to be 'arbitrary *pro*.' The latter is distinguished from arbitrary PRO by the presence of the nominative Case feature. With respect to the inherent features of person, number, and gender, however, the two arbitrary pronominals are identical, as illustrated by the following examples.

- (8) a. *E difficile [PRO vivere contenti in questa città]*
'(It) is difficult to live happy(ly) in this city'
b. *pro si vive contenti in questa città*
'One lives happy(ly) in this city'

In (8a,b) the adjective *contenti* (=happy) bears the [+plural, +masculine] affix *i*. In (8a) the agreement on the adjective is triggered by the features of arbitrary PRO, while in (8b) the agreement is triggered by the features of arbitrary *pro*.

There is one final aspect of the impersonal morphology worth noting. As mentioned earlier, in the impersonal construction the verb is inflected in a third-person singular form [cf. (1a,b), (8b)]. We may assume that *si* is a spelling out of the AG features in INFL. Because these features are realized as the *si* clitic they are not available to affix onto the verb in the form of verbal agreement. Hence the verb assumes the neutral third-person singular form, as proposed in Belletti (1982b).

While the description of impersonal *si* as a subject clitic is relatively uncontroversial, the status of the middle *si* is considerably less clear. Debate concerning the middle *si* centers around two issues. The first concerns its Case require-

⁷ The sentences in (7a,b) are also impossible under a middle interpretation of *si*. I return to this in the next section.

ments; that is, whether the *si* of the middle construction is associated with nominative Case (Burzio 1981; Manzini, this volume) or whether it bears accusative Case (Belletti 1982b). The second issue concerns the appropriate characterization of the "passivizing" property of middle *si* referred to earlier. Belletti (1982) has proposed that the middle *si* construction is functionally equivalent to the copula passive. On her analysis, the middle *si* absorbs the subject θ -role and accusative Case; these absorption processes being those which define the copula passive as well (Chomsky 1981). Note that in Belletti's analysis *si* is always a clitic generated under INFL which bears the subject θ -role. The middle and impersonal constructions differ only with regard to the Case of *si*. When *si* bears nominative case it is interpreted as an impersonal; when it bears accusative Case the interpretation is "passivelike."

An alternative treatment of the passivizing property of the middle *si* is proposed in Zubizarreta (1982a) and more recently in Manzini (this volume). Within the latter theories *si* is defined as a verbal affix which blocks θ -role assignment to the subject position and accusative Case assignment to the object position.⁸ On the above accounts, then, *si* may function either as a clitic, which bears the subject θ -role and accusative Case (=impersonal, *si*) or as an affix, which blocks features assignment (=middle *si*). While the clitic/affix distinction seems to be essentially correct, the above accounts fail to explain why *si* sometimes functions as a clitic and sometimes as an affix and hence the distinction, while descriptively correct, is stipulative. On these accounts the passivizing *si* "affix" is a syntactic element and the operations associated with *si*—the absorption or blocking of θ -role and Case assignment—are assumed to operate in the syntax.

In the following section, I offer a somewhat different account of the middle *si* construction. Specifically, I will argue that the middle *si* construction [cf. (1c,d)] is the output of the lexical word-formation process. While my analysis incorporates many of the basic insights of the above-mentioned studies, in particular the clitic/affix distinction just discussed, it largely eliminates the need to stipulate these different functions. As we will see, the properties of the two *si* constructions follow as an effect of the level of grammar at which *si* enters into a word-formation process with the predicate.

2.2. Word-Formation in the Lexicon

As a point of departure, let me make explicit certain assumptions concerning the organization of the lexicon. I assume that each lexical entry contains a set of

⁸ The analyses of *si* proposed by Zubizarreta and Manzini differ from each other in several respects. In particular, Manzini proposes that the middle construction has two parallel S-structure representations, one in which *si* is a passivizing affix with the properties noted in the text, and a second in which it appears as the impersonal *si* clitic.

phonological, semantic, and syntactic features, including the category specification of the word—for example, [+V, -N]—and the θ - and Case features associated with lexical heads. With regard to θ -features, although the subject θ -role of a sentence is often compositionally determined by the elements in the VP (as noted in the previous section), I assume that there is a default or unmarked subject θ -role associated with each verb in its lexical entry (provided the verb is one which θ -marks a subject). It is clear that in cases in which there is no verbal complementation (e.g., intransitive verbs), the θ -role is determined by the verb alone. Moreover, there is a strong intuition that there is a basic θ -role associated with most transitive verbs. For example, the default θ -role associated with the verb *break* is *agent*, as in the sentence in (4a). Given this assumption we see that there is a basic asymmetry exhibited in the lexicon with respect to the specification of Case and θ -features: while the subject (default) θ -role assigned by a head is lexically specified, the Case assigned to the subject (i.e., nominative) is not. The latter is assigned by tense (or AG) in INFL. As we will note shortly, this particular asymmetry will have consequences for the analysis of the middle *si* construction.

Given the organization of the lexicon outlined above, it is clear that the requirements of *si* listed in (3a–c) can be satisfied in this component. Property (3a) is satisfied straightforwardly if *si* attaches to a verb; property (3b) is satisfied if the verb to which *si* affixes bears a (default) subject θ -role, specifically a θ -role which can be assigned to a [+human] subject. Thus, *si* cannot affix onto a raising verb, or other predicates which are not specified as assigning a θ -role to subject position.⁹ Notice that there is a clear prediction of this system, namely, that the θ -role absorbed by middle *si* is noncompositional. This issue is discussed further in Section 3.2. Let us turn to property (3c). Note that given the Case/ θ asymmetry discussed above, it is clear that to satisfy property (3c) in the lexicon,

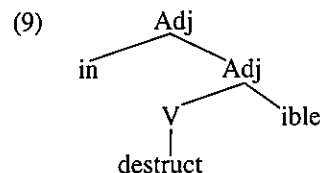
⁹ There is an apparent counterexample to this claim, namely, sentences in which middle *si* is attached to so-called restructuring predicates, for example, *potere* ('can') and *dovere* ('must'), as in the following sentence

- (i) *Questi libri si possono (devono) leggere la sera*
 'These books *si* can (must) read in the evening'
 (=These books can (must) be read in the evening)

I assume that the "modals" *potere* and *dovere* do not assign a subject θ -role (at least in their epistemic interpretation). Moreover, they lack accusative Case features. Thus the question arises as to how the lexical properties of *si* can be satisfied by these predicates in the lexicon. Following Rizzi (1976) and Burzio (1981), I assume that the modal-verb sequence forms a complex predicate. My analysis departs from theirs, however, in that I assume that the complex predicate is not formed in the syntax but rather in the lexicon, where it inherits the lexical features of the main verb. Thus, in (i) the complex predicate *poter leggere* inherits the thematic and accusative Case features of the verb *leggere* ('read') and these features satisfy the lexical requirements of *si*. This analysis follows in the spirit (if not the letter) of Piccolo (1985), who presents some compelling evidence that these complex predicates are base-generated as such and not syntactically derived via restructuring.

si must affix onto a transitive verb. This follows from the fact that the only Case features specified in the lexicon are nonnominative.

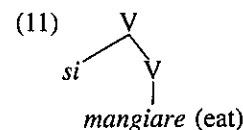
Note that (3b,c) specify that *si* must "absorb" a subject θ -role and Case feature. Following Jaeggli (to appear) I assume that "absorption" is simply the assignment of a feature to a bound morpheme. It is generally assumed that in the syntax the assignment of lexical features is governed by various structural conditions. Minimally, the assigner must c-command the assignee. Thus on the analysis assumed here in the syntax *si* (or more properly, *pro* in the [NP,S] position) is assigned a θ -role by a c-commanding VP, and nominative Case by INFL. We can extend this requirement to the lexicon by making the standard assumption that morphologically complex words also have an internal structure, represented by a tree structure [Lieber (1980) and others], as exemplified in (9). We may further assume that the stem of the complex assigns features to affixes under c-command. [I will assume Reinhart's (1976) definition in which the domain of c-command in the first branching node.] In this regard, consider the structure in (9).



Note that the suffix *-able/ible* is one which absorbs or is assigned the accusative Case feature and the subject θ -role of the verbal stem *destroy*, being c-commanded by the latter [cf. (10a,b)].

- (10) a. John destroyed the television
 b. Cockroaches are indestructible

Similarly, we may assume that lexical affixation of *si* is represented by the structure in (11). The verbal stem satisfies the c-command requirement on assignment (absorption) of features.¹⁰



¹⁰ If the analysis proposed here is correct, there is a curious asymmetry between syntactic and morphological structure in that in the lexicon features are assigned by the stem (i.e., a non-head) while in the syntax features are assigned under government by a lexical head (cf. Chomsky 1981). I have little to say about this asymmetry except to note that if it is the case that the subject θ -role is assigned by the VP in the syntax, the 'assignment-by-head' requirement may be too strong.

While the assignment/absorption of features in the lexicon may be governed by structural conditions similar to those required by syntactic assignment/absorption, the output of the two processes differ in a crucial respect. Recall that the Projection Principle [cf. (2)] states that lexical features must be represented at every syntactic level. This entails that features which are assigned/absorbed by a bound morpheme in the syntax must remain "visible" to syntactic principles, in particular, the Case Filter and the θ -Criterion. Note, however, that feature assignment/absorption in the lexicon is not constrained in this way. Thus lexical features which are assigned/absorbed in the lexicon are arguably "invisible" to principles such as the θ -Criterion and the Case Filter. This invisibility of features can be said to follow from Lexical Integrity Hypothesis, proposed by Lapointe (1981) as an extension of the Lexicalist Hypothesis (Chomsky 1970). The Lexical Integrity Hypothesis is stated below.

The Lexical Integrity Hypothesis

Syntactic rules cannot refer to parts of words.

Following Borer (1984a) (and see also Borer and Grodzinsky, this volume), I assume that lexical integrity may be derived from Allen's (1978) Convention of Bracket Erasure, according to which all internal bracketing in morphologically complex forms is eliminated by the end of the lexical derivation, as exemplified below.

destroy \rightarrow [destruct]ible \rightarrow in[destructible] \rightarrow [indestructible]

Assuming that only those elements that are represented as separate constituents constitute "words" for the syntax, it follows that syntactic rules and principles cannot "see" inside words formed in the lexicon. Returning to the requirements of *si* given in (3b,c), the above formulation entails that when *si* absorbs the accusative Case feature and subject θ -role in the lexicon, these features are effectively eliminated as far as the syntax is concerned.¹¹

Let us sum up the discussion thus far by noting that when *si* undergoes word formation in the lexicon, the resulting construction has the following properties:

- (12) a. *si* appears in VP (on V).
 b. *si* absorbs (=eliminates) the subject θ -role determined by V.
 c. The subject position is nonthematic.
 d. *si* absorbs (=eliminates) accusative Case.

¹¹ We leave open the possibility that thematic features on bound morphemes are visible at some nonsyntactic level of representation. Roeper (1983) argues convincingly that some derivational affixes (e.g., *-able*, *-ing*) license "hidden" or "implicit" arguments, which may show up in an agentive *by* phrase, as in (i) or control into a purposive clause, as in (ii):

- (i) The grammar is learnable (by the child)
 (ii) Doctors recommend meat-eating (to gain weight)

In what follows I will argue that the properties listed in (12), which follow from a lexical process of word formation, are those which characterize the middle *si*. I will consider the empirical effects of the properties given in (12) as compared to those given in (5) for the impersonal construction.

3. ON CERTAIN DIFFERENCES BETWEEN THE IMPERSONAL AND MIDDLE *SI* CONSTRUCTIONS

Let us begin by comparing the Case properties of *si* in the impersonal versus middle construction [cf. (12c) and (5c)].

3.1. The Case of *Si*

As noted by Belletti (1982b) and others, a direct object may be cliticized in the impersonal construction, but not in the middle, as illustrated in (13).

- (13) a. (*Le montagne*) *le si vede bene da qua*
 '(The mountains) them one sees well from here'
 b. *(*Le montagne*) *le si vedono bene da qua*
 '(The mountains) them are seen well from here'

On the assumption that object clitics require accusative Case, the ungrammaticality of (13b) follows from the fact that in this instance *si* has eliminated the accusative Case feature associated with the verb 'see.' In (13a), in contrast, *si* is assigned nominative Case, and hence cliticization of the object is possible.

With the elimination of accusative Case in the middle the direct object is forced to receive nominative Case to escape the effects of the Case filter. A diagnostic for the presence of nominative Case on the object is provided by the fact that a nominative NP triggers agreement on the verb. As illustrated in (14), it is the plural (logical) object (chocolates) which triggers agreement on the verb *mangiare*.¹²

- (14) *Si mangiano i dolci al cioccolato in questa pasticceria*
 'Are eaten (3rd per. plu.) chocolates in this pastry shop' [= (1c)]

Moreover, as is typically the case in free-inversion languages, the nominative NP may also occur in preverbal position, as follows:

- (15) *I dolci al cioccolato si mangiano in questa pasticceria* [= (1d)]

¹² I assume that nominative Case may be directly assigned to postverbal position in Italian and other free-inversion languages.

The fact that the middle *si* co-occurs with a full nominative NP provides prima facie evidence that in this construction *si* does not itself require nominative Case, contrary to the analyses presented in Burzio (1981) and Manzini (this volume). In fact, on the analysis proposed here, in which the Case property of *si* is satisfied in the lexicon, we expect that the middle word-formation rule should be entirely insensitive to the presence or absence of nominative Case given that the latter is not available at this level. Middle *si* should thus be possible in constructions which lack nominative Case. As is illustrated in the following examples, middle *si* may appear under raising verbs (16a,b) while the corresponding impersonal *si* (which requires nominative Case) may not (16c,d).¹³

- (16) a. *Azioni del genere sembrano criticarsi senza ragione*
 'Actions of this kind seem to criticize (to be criticized) without reason'
 b. *Le mele risultano mangiarsi volentieri*
 'Apples turned out to be eaten willingly'
 c. **Sembra criticarsi senza ragione azioni del genere*
 'Seems to criticize without reason actions of this kind'
 (=One seems to criticize. . . .)
 d. **Risulta mangiarsi volentieri le mele*
 'One turned out to eat apples willingly'

Earlier I noted that the impersonal construction is also impossible under control verbs [cf. (7b)]. This follows from the fact that under syntactic word-formation, *si* must be assigned nominative Case, which is unavailable in control structures. Curiously, however, the middle *si* is also impossible in control structures (see Footnote 7).

- (17) a. **I ragazzi sostengono di punirsi senza ragione*
 'The boys claim to have been punished without reason'
 b. **E facile lavarsi di mattina*¹⁴
 '(It) is easy to be washed in the morning'

The embedded verbs in (17) are transitive. Hence, in principle *si* has the option of undergoing word-formation in the lexicon and absorbing accusative Case. The sentences are thus predicted to be grammatical, contrary to fact.

Burzio (1981) and Manzini (this volume) cite examples of the above sort as

¹³ Belletti (1982b) also uses examples like those in (16) to argue that *si* absorbs accusative Case. On her analysis, however, the absorption is a syntactic operation, differing from the impersonal only in that the latter absorbs nominative Case. It is also worth noting that on her analysis the absorption of accusative Case by *si* is a nonlocal operation given that *si* is contained under INFL. As we will see shortly, there is some evidence to suggest that the middle *si* is contained in the VP as I have argued and not in INFL.

¹⁴ Note that this sentence is fully grammatical on the irrelevant reflexive interpretation 'it is easy to wash oneself in the morning.'

evidence for their claim that the middle *si* must be associated with nominative Case (like the impersonal). Belletti, in contrast, has suggested that the ungrammaticality of (17) is due to a basic incompatibility of *si* and PRO. It is interesting to note, however, that similar examples are also ungrammatical in English.¹⁵ Consider the following:

- (18) a. Most bureaucrats bribe easily
 b. *Most bureaucrats try to bribe easily
 c. *In this profession it is necessary to bribe easily
 d. Most bureaucrats seem to bribe easily

The English verb *bribe* undergoes middle formation. However, the middle is impossible in control structures and under adjectives, though possible in the raising case. While I have no immediate explanation for the ungrammaticality of (18b,c), the English examples suggest that the ungrammaticality of the Italian examples in (17) is not directly related to *si*. In particular, it is not related to the Case requirement of *si*. Rather, there seems to be a basic incompatibility between PRO and middle verbs.¹⁶

¹⁵ I am indebted to Osvaldo Jaeggli and Ian Roberts for pointing this out to me.

¹⁶ This is particularly interesting in that PRO is perfectly acceptable in passives embedded under control verbs in both English and Italian, as the following examples illustrate.

- (i) Most bureaucrats hope PRO to be bribed on a regular basis
 (ii) *I ragazzi sostengono di PRO essere stati puniti*
 'The boys claim to have been punished'

Notice that the opposite judgments hold with respect to 'tough' constructions in Italian, that is, middles are possible and passives are not.

- (iii) *Questi libri sono difficile da capirsi*
 'These books are easy to understand (=be understood)'
 (iv) **Questi libri sono difficile da essere capiti*
 'These books are difficult to be understood'

The fact that the middle and passive behave differently both with respect to control structures and tough constructions suggests that they should not be treated as equivalent, as proposed in Belletti's analysis. One further difference between the middle and the passive is that an agentive *by* phrase is possible in the latter but not in the former. The *by* phrase is also impossible in the impersonal construction.

- (v) *I dolci sono stati mangiati da Mario*
 'The sweets were eaten (passive) by Mario'
 (vi) **I dolci si mangiano da tutti*
 'The sweets are eaten (middle) by all'

On Belletti's analysis the above contrast follows from the fact that middle *si* is a clitic which requires a subject θ -role and hence the latter is unavailable for the NP in the *by* phrase. On my analysis, the subject θ -role has been eliminated by *si* and hence is unavailable for the *by* phrase—a θ -Criterion violation in either case.

3.2. The Thematic Properties of *Si*

Let us now consider the thematic properties of *si*. Here we are interested first in the thematic status of the subject position of the impersonal and middle construction [cf. (5d) and (12d)], and second in the nature of the θ -role assigned to *si* in each construction [cf. (5b) and (12b)].

I have argued that the syntactic process of word formation results in a structure in which *si* forms a discontinuous pronoun with an arbitrary *pro* in [NP,S]. The lexical process, in contrast, results in the elimination of the subject θ -role. Thus, in the middle construction the subject position is a nonthematic position.¹⁷ We thus expect that in the middle (but not the impersonal) preposing of the (independently θ -marked) direct object will be grammatical, given that movement is only possible to a non- θ position (cf. Chomsky 1981). As illustrated in example (19) preposing in the middle is perfectly grammatical, though as shown in (20) preposing is not possible in the impersonal construction.

(19) *I dolci al cioccolato si mangiano in questa pasticceria* [= (15)]

(20) **Le montagne si vede da qua*¹⁸
'The mountains one sees from here'

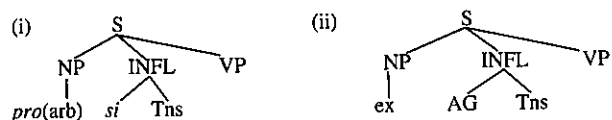
A second prediction is that it should be possible for idiom chunks (possible only in non- θ -positions) to appear in subject position of the middle (examples from Rizzi 1976).

- (21) a. *Mari e monti si sono promessi invano*
'Heaven and earth were promised in vain'
b. *Assistenza si è portata in ritardo ai sinistrati*
'Assistance was brought late to the leftist'

The idiom in (21a) is *promettere mari e monti* (lit. 'to promise seas and mountains') and in (21b) the idiom is *portare assistenza* ('to bring assistance or help').

The second issue, which was discussed briefly in Section 2.2, concerns the status of the θ -role which is assigned to *si*. I suggested earlier that the impersonal

¹⁷ I assume that the subject position of the middle contains a null expletive which is governed by AG in INFL (cf. Safir 1982b for arguments that null expletives must be governed). Thus, the structures I am assuming for the two constructions are as in (i) and (ii); (i) being the impersonal and (ii) the middle.



¹⁸ This sentence is grammatical only if *le montagne* is left-dislocated in which case there is a pause preceding the *si*. I assume that in left-dislocation the object is moved to COMP (or alternatively adjoined to S), not into subject position.

si bears a compositional θ -role determined by the VP [cf. (5b)], while the middle *si* absorbs the default subject θ -role associated with the verb in its lexical entry. In this regard, Jaeggli (to appear) observes that middle formation is impossible where the external θ -role is necessarily compositional as in expressions like *alzare le mani* ('to raise one's hands'), while the latter is fully grammatical in the impersonal construction. Compare the following.

(22) **Le mani si alzano per chiamare l'attenzione del professore*
'Hands are raised to call the teacher's attention'

(23) *In questa aula si alza le mani per chiamare l'attenzione del professore*
'In this classroom one raises one's hands to call the teacher's attention'

In light of the above contrast, Jaeggli proposes that the middle *si* appears inside VP where it absorbs the "pure" external θ -role associated with the verb, what I have called the "default" θ -role, while the impersonal *si* appears in INFL where it absorbs the compositional θ -role determined by the VP. Thus on Jaeggli's analysis the type of θ -role which *si* bears depends on the syntactic position in which it appears. Note that these results follow straightforwardly on the analysis presented here. Recall that when *si* undergoes word formation in the lexicon, it is affixed directly onto the verb. At this level of derivation the only subject θ -role available is the default θ -role associated with the verb. When *si* undergoes word formation in the syntax, in contrast, it appears in INFL where it receives the compositional θ -role determined by VP.

Let us conclude this section by noting that the impersonal and middle constructions differ not only in Case specification, as noted by Belletti (1982b), but also with regard to the content of the empty category in [NP,S] and the nature of the θ -role absorbed by *si*. Both of these results follow from the analysis proposed here.

3.3. The Interaction of Word-Formation and the Syntax

In this section I discuss certain further differences between the middle and impersonal constructions with respect to their interaction with the syntax. I focus on two issues: first, the hypothesis that syntactic rules cannot "see" inside words formed in the lexicon (Lexical Integrity), and second, the interaction of *si* and derived subjects.

Recall that *si* is lexically specified as bearing the 'nominal' features [3p, plural, masculine]. Given Lexical Integrity (or alternatively, the Convention on Bracket Erasure), these features should not be visible to the syntax when *si* undergoes word formation in the lexicon. Hence, we do not expect these features to enter into agreement processes in the middle construction, though agreement should be triggered in the impersonal. As I noted earlier adjectives in the imper-

sonal bear the masculine, plural affix *i*, as in (8b). Past participles, which function like adjectives with respect to agreement, are similarly inflected, as in (24a). Note that there is no such agreement with the middle *si* in (24b).

- (24) a. *Si è andati al cinema ieri sera*
 'One went to the movies last night'
 b. *Questa pasta si è mangiata con piacere*
 'This pasta was eaten with pleasure'
 c. **Questa pasta si è mangiati con piacere*

In (24b) the past participle (which functions like an adjective with respect to agreement) is inflected with the feminine, singular affix *a* in agreement with the subject *la pasta*. Crucially, it cannot agree with *si*, as in (24c).¹⁹ Similarly, the features of middle *si* are not available to license (or identify) an arbitrary *pro* in subject position. Hence, the middle does not receive an arbitrary subject interpretation as does the impersonal. The fact that the features of *si* in the middle construction are inaccessible to the syntactic processes of agreement and identification follows from the hypothesis that *si* undergoes word formation in the lexicon.

Finally, let us consider the interaction of *si* and syntactic rules which yield derived subjects. In Section 3.1 we observed that the middle *si* is insensitive to the presence or absence of nominative Case, illustrated by the fact that it appears in infinitives under raising verbs (and tough constructions; see Footnote 16). This result is expected if the Case requirement of *si* is satisfied at a presyntactic level (i.e., the lexicon), since at this level only accusative Case is available. In a similar vein, we expect that the middle *si* cannot be associated with subjects which are derived via syntactic rules—again because its properties are satisfied at a presyntactic level. This follows straightforwardly given a modular organization of grammar in which components are intrinsically ordered. Recall that the impersonal *si* satisfies its Case and θ requirements by forming a discontinuous pronoun with a null element in [NP,S]. Thus, the impersonal *si*, whose lexical requirements are satisfied in the syntax, can be associated with subjects derived via passive, tough movement, and (marginally) raising, while the middle *si* cannot. Consider the following sentences.

¹⁹ Note that it is not the case that participles invariably agree with the [+nominative] NP. The participles of transitive verbs, for example, are marked with the "neutral" masculine singular *-o* irrespective of the number or gender of the subject, as in (i), or they are inflected to agree with an object clitic if there is one, as in (ii).

- (i) *I gatti hanno graffiato le sedie*
 'The cats [masc. plu.] have scratched [masc. sing.] the chairs'
 (ii) *I gatti le hanno graffiate*
 'The cats them [fem. plu.] have scratched [fem. plu.]'

- (25) a. *Si è stati arrestati dalla polizia*
 'One was arrested by the police'
 b. **I ragazzi si furono arrestati dalla polizia*
 'The boys *si* (middle) were arrested by the police'
 c. *In questo paese si è difficile da accontentare*
 'In this country one is difficult to please'
 d. **I ragazzi si sono difficile da accontentare*
 'The boys *si* (middle) are difficult to please'
 e. ?*Stando così le cose, si risulterebbe costruire le case in questione senza licenza*
 'As things are, one would turn out to build the houses in question without a license'
 f. **Stando così le cose, le case in questione si risulterebbero costruire senza licenza*
 'As things are, the houses in question would be built without a license'

Although I will not attempt a detailed analysis of the above sentences, we may assume that in (25a,c,e) *si* identifies a *pro* in the matrix [NP,S] position from which it obtains a θ -role. In the passive and raising sentences, the *pro* subject has been derived via movement from the embedded clause, while the *pro* subject of the tough construction receives its θ -role via co-indexing with the null element in object position of the embedded clause.²⁰ The ungrammaticality of the examples in (25b,d,f) follows from the fact that there is no external θ -role available for *si* at the presyntactic level.

This concludes our discussion of *si* as it functions in the adult grammar of Italian. In the section that follows, I turn to the acquisition of *si* by Italian-speaking children.

4. THE ACQUISITION OF WORD-FORMATION RULES

Much recent work in language acquisition centers around the hypothesis that the intermediate grammars constructed by the child in the course of grammatical development are constrained by principles of UG. This hypothesis entails that the principles and parameters of UG define not only the limits of variation between adult grammars of different languages, but also the limits within which child grammars may differ from the adult "target" grammar of a particular language. Thus on this view UG also constitutes a theory of "possible error" (misanalysis) by the child and the relationship between language-particular variation and ac-

²⁰ See Chomsky (1981) and Belletti (1982b) for discussion of *tough* constructions.

quisition assumes a particular significance; wherever we find variation with respect a particular rule or principle we have a potential acquisition error.²¹

There is considerable evidence that languages may vary with respect to the level at which various grammatical rules apply. For example, Huang (1982) has argued that English and Chinese differ with respect to the level of grammar at which Move α applies; in English, the rule applies at S-structure and at LF while in Chinese movement is restricted to LF. Similarly, Chomsky (1981) argues that languages may vary with respect to the level at which affix-hopping (Rule R) may apply, the phonology or the syntax.

The theory of morphology presented in the previous sections allows that rules of word formation may apply freely in different components of the grammar, specifically the lexicon and the syntax, subject to the principles which constrain representations at those levels (Borer 1984a). I illustrated the empirical effects of this proposal in the interpretation of Italian *si*. Given such a system, we might expect to find two languages, L and L', which differ in that a particular word-formation rule applies in the syntax in L and in the lexicon in L'. Precisely this situation is found with Italian *si*. In the dialect of Italian discussed in the previous sections, *si* undergoes word formation in both the syntax and the lexicon. There are, however, other dialects—notably Fiorentino and Trentino (cf. Brandi and Cordin 1983)—in which we find only the impersonal *si*, that is, word formation is restricted to the syntax. Alternatively, there are dialects—for example, the Italian spoken in Ferrara and Bergamo—in which the impersonal *si* is possible only with intransitive verbs, that is, word formation is restricted to the lexicon (with transitive verbs).²² To the extent that word-formation rules are subject to this kind of variation, we have an area of potential “error” for the child.

In this section I suggest that such errors do in fact occur in grammatical development. More to the point, I propose that there is an early stage in which the word-formation rule affecting *si* is restricted to the syntactic component and thus the middle construction is absent. I will show, however, that the child’s “misanalysis” of *si* is not an isolated phenomenon. Similar errors occur in the English-speaking child’s acquisition of morphology, in particular in the child’s

²¹ More precisely, principles of UG specify a range of possible “errors.” Further principles are needed to specify why certain “misanalyses” are more likely or more highly valued. The latter may follow from the theory of markedness, various processing considerations, the “strength” or salience of aspects of the input data, and so on.

²² In these dialects, transitive verbs undergo middle formation only, the impersonal *si* construction being restricted to intransitive verbs. The relevant judgments are as follows [cf. (1) in the text]:

- (i) *In questa città si va spesso al cinema* [= (1a)]
- (ii) **In questa pasticceria si mangia soltanto i dolci al cioccolato* [= (1b)]
- (iii) *I dolci al cioccolato si mangiano in questa pasticceria* [= (1c)]

On the analysis of *si* proposed in the text, the appropriate description of these dialects is that *si* necessarily undergoes word formation in the lexicon if the verb is transitive.

interpretation of the nominalizing *-ing* affix and the agentive *-er* and his/her production of novel compounds and “word phrases.” The acquisition data from Italian and English suggest that child grammars may differ from adult grammars with respect to the component in which rules of word formation apply.

4.1. Morphological Inheritance

As a point of departure, let us consider the examples in (26)–(28). In each case, the (a) example is the adult form of the phrase or sentence, while the (b) examples constitute the child’s interpretation of the corresponding phrase or sentence.

- (26) a. the writer with a pen
(accompaniment)
- b. the writer with a pen
(instrument)
- (27) a. Can you show the cow the hitting of the horses
- b. (The cow) is hitting the horses
- (28) a. *Gli occhi non si vedono*
‘The eyes are not seen’
- b. *Non si vede gli occhi*
‘One doesn’t sees the eyes’

Randall (1982) has observed that young children freely interpret the PP complement to a derived *-er* noun as *instrumental* [cf. (26b)]. She argues that, for children, the phrase *writer with a pen* is identical in thematic structure to *write with a pen* or *writing with a pen*. This interpretation is, of course, impossible for adults, for whom *the pen* in (26a) has only an *accompaniment* reading, the *instrument* θ -role having been eliminated by the word-formation rule. On the basis of her experimental results, Randall (1982) proposes that children are obeying a principle of Morphological Inheritance, given in (29).

(29) Morphological Inheritance Principle

A morphologically complex form will inherit the argument structure of the base form.

In the examples in (27), we notice a similar phenomenon. Roeper (1983) notes that there is an early stage at which children tend to interpret the complement of derived *-ing* nominals as the direct object, although for adults the element following the preposition *of* is easily interpretable as the subject of the action nominal. Moreover, these children frequently interpret the matrix object, for example *the cow* in (27a), as the subject of the action nominal. Roeper points out that children assign a simple SVO structure to these clauses, ignoring, as it were,

the determiner preceding the nominal and the preposition following it. As in the example in (26), we can account for the child's "misanalysis" by assuming that the derived action nominal inherits the *agent-patient* argument structure associated with the verbal stem.²³

Finally, in the *si* examples in (28), we see that Italian children systematically interpret the middle construction as an impersonal. This phenomenon is particularly noteworthy insofar as these children have been exposed only to the Ferrarese dialect of Italian (noted previously), in which the impersonal *si* is impossible with transitive verbs. Thus, strictly speaking, sentences like that in (28b), which are taken from spontaneous production, are "inventions" of the child.²⁴ Viewed from the perspective of 'morphological inheritance,' we can describe the phenomenon illustrated in (28a,b) by saying that in the early grammar the morphologically complex form consisting of a verb plus clitic inherits the Case and θ -features associated with the verbal stem. However, we might ask why this should be so or more to the point, why does morphological inheritance persist in the early grammar in instances in which it is barred in the adult grammar?

In the study of adult grammars, it has proved a useful strategy to use the various conditions of well-formedness as diagnostics for determining the level of grammar at which particular rules apply. Thus, if a rule appears to obey the conditions of well-formedness which hold at the level of Logical Form, as opposed to the conditions which hold in the syntax, this is taken as evidence that the rule in question applies at LF. We might employ exactly the same strategy in determining the location of rules in child grammar. Bearing in mind that the strongest acquisition evidence for the operation of a particular rule or principle comes from "mistakes" in which the rules applies where it would not in the adult grammar—for example, *foots* and *goed*—let us reconsider the data in (26)–(28). Note that in each instance the lexical properties of the verbal stem are being preserved under affixation—exactly as would be required by the Projection Principle. Now, given the logic of the system presented in Sections 2 and 3, it follows that if the early word-formation rules obey the Projection Principle, these rules must apply in the syntax. Thus on the assumption that some of the child's word-formation rules are restricted to the syntax and hence are subject to

²³ Action nominals, unlike agentive nouns, can inherit the thematic structure of the verbal stem in the adult grammar as well, as in 'John's hitting of the horses.' The example in the text is of interest in that it illustrates that inheritance is obligatory for children. Thus, in the absence of an overt subject the child assigns the *agent* θ -role to the matrix object.

²⁴ It is possible that the child's use of impersonal *si* with transitive verbs represents an instance of "overgeneralization" based on the impersonal *si* with intransitive verbs, which is permissible in this dialect (cf. Footnote 22). Assuming this to be the case, the question still remains as to why these children systematically fail to also produce the middle construction despite the availability of such constructions in the input data and despite the frequent corrections which they receive when they mistakenly produce the impersonal with transitive verbs. See Hyams (1983) for further examples and discussion of these acquisition data.

the Projection Principle, we derive the effects of morphological inheritance in child grammars.

Note that in the above analysis, children's errors of interpretation provide evidence for the operation of the Projection Principle in early grammars, as is predicted under the hypothesis that grammatical development is constrained by principles of UG. This hypothesis is further supported by the fact that the grammatical "error" which occurs falls at a point at which we find variation in adult grammars.

Before concluding, let us briefly consider some further evidence for the analysis proposed in this section. The examples in (26) and (27) involve the English-speaking child's interpretation of morphologically complex forms. The following examples are production data.

- (30) a pull-things
an open-door
a hiding-thing
a cut-grass
- (31) a. This is *my do it*
b. That's *a flush a toilet*
c. More read dat
d. My see magazine
e. Another turn button

The data in (30), from Clark and Hecht (1982), were obtained while trying to elicit *-er* nominals from children under the age of 3 years, 9 months. Children were asked "what would you call someone [or something] that Vs [e.g., that pulls or pulls something]?" Clark and Hecht analyze these forms as novel compounds. It is clear, however, that the "compounding" rule in the early grammar differs from that of English. In particular, it appears that these forms are created by directly converting the verb and its complement into a nominal, which we might assign roughly the following structure. I will refer to such constructions as "word-phrases," following Borer's (1984a) proposal for similar constructions in the adult language.

- (32) [a [[pull [things]]]]
NP N VP

Notice that in the forms in (30) the internal argument of the verb is preserved, and by hypothesis, Case and θ -marked by the verb. This follows under the assumption that the word-formation process responsible for the novel forms is constrained by the Projection Principle, that is, it applies in the syntax.

With respect to the data in (30), one might argue that the children are not producing innovative complex forms but rather are simply repeating part of the experimenter's previous utterance "What do you call something that *pulls*

things?' As Clark and Hecht note, however, the child's forms are preceded by a determiner, which is not present in the experimenter's question. Moreover, the inflection on the verb is systematically eliminated in the child's response. Both these observations suggest that the responses are not merely imitations. Note, finally, that there are many spontaneous productions (and hence clearly not imitations) which share the basic properties of these elicited forms, for example, the data in (31).

The examples in (31a,b) are from Hamburger and Crain (1982). They suggest that the italic portions of the sentences are "proto-relatives clauses." Thus, the sentence in (31a) is interpreted as (roughly) 'this is my thing such that I do it.' I would like to propose, in contrast, that these expressions are formed by a process of syntactic word-formation, yielding the structure in (33), parallel to that in (32).

(33) [my [[do [it]]]
 NP N VP NP

Note that again the Case and θ -features associated with the verb are maintained as required by the Projection Principle. Thus as in the novel "compounds," the derived noun [which may be preceded by a possessive pronoun as in (31a,d) or quantifier, as in (31c,e)] contains a verb and the NP which it Case and θ -marks.

The utterances in (30) and (31) do appear to violate the Projection Principle in one regard, however. Notice that in most cases there is no external argument. The verb is preceded by a determiner as in 'a pull-things.' The question arises then as to what becomes of the subject θ -role. I will assume, following Chomsky (1981), that an external θ -role must be expressed only if the subject position is obligatory, that is, in clauses. Where the subject position is optional, as in NP, the expression of the external θ -role is correspondingly optional. Notice that in the examples in (31a,d) the genitive pronoun *my* acts as the external argument bearing the subject θ -role. Thus, the "relaxation" of the Projection Principle in NPs can account for the apparent optionality of an external argument in children's word-phrases.

5. CONCLUSION

In this essay I have argued that the basic properties of the impersonal and middle *si* in Italian can be derived from a model of morphology in which word-formation rules are free to apply in the lexicon and in the syntax. I further proposed that a modular system of this sort can account for certain morphological differences which exist between child and adult grammars under the hypothesis that real-time grammatical development is constrained by principles of UG which defines the notion "possible acquisition error."