

5. THE VARIABILITY OF IMPERSONAL SUBJECTS

1. INTRODUCTION

My goal in the present paper is to analyze the quantificational and anaphoric properties of so called impersonal *si* constructions in Italian, exemplified in (1).

- (1) In Italia, *si* beve molto vino
 In Italy *si* drinks a lot of wine
 In Italy people drink a lot of wine

Si here is a subject clitic and it expresses a non specific, generally plural, human subject (something similar to the English bare plural *people*, which however, unlike *si*, is not restricted to the subject position). I think that what I have to say about *si* extends to impersonal subjects in Romance (e.g. French *on*) and, possibly, to impersonal subjects in other language families as well (e.g. German *man*), but I will limit myself here to a consideration of Italian data. The syntax of impersonal *si* has been extensively studied (for important work within the Principles and Parameters framework, see e.g. Burzio (1986) and Cinque (1988)). However, to the best of my knowledge, no truth conditionally explicit logical forms for constructions involving *si* have been proposed, and this is a gap that the present paper seeks to fill. As we shall see, to fill such a gap is not a trivial task. In the remainder of this section, I will list a cluster of six empirical properties that impersonal *si* appears to have and that a semantically explicit theory of such a construction should take into consideration. I will argue that a version of Discourse Representation Theory (to be specified in Section 2)¹ can offer an interesting account of the properties listed below.

The first property I would like to consider has to do with the observation that impersonal *si* constructions are usually understood as expressing a quantification of some kind, but their quantificational force can vary considerably.

- (2) Property A: quantificational variability in generic vs. episodic sentences.

- a. In Italia si beve molto vino
In Italy si drinks a lot of wine
In Italy, everybody/people drink a lot of wine
- b. In Italia ieri si è giocato male
yesterday in Italy si played poorly
yesterday people in Italy played poorly

Sentence (2a) is understood as a general property of people in Italy. Sentence (2b) is likely to be understood as being about some particular group of people (say, the Italian soccer team), although the possibility of it being about all Italians is not ruled out. Sentence (2a) is generic, while sentence (2b) is episodic. As (2) illustrates, generic sentences appear to trigger a quasi universal reading of *si*, while episodic sentences favour an existential reading. This behavior seems parallel to that of *bare* plurals in English (cf. Carlson, 1977). Consider for example the following cases.

- (3) a. Bears like honey
b. Bears are hibernating in this area

Sentence (3a), which is generic, triggers a quasi universal reading of the subject. In contrast with this, sentence (3b) is not so constrained. While the parallelism between *si* constructions and bare plurals is revealing, it should also immediately be noted where it breaks down. Bare plurals in English happily go with kind-level predicates (cf. 4a), but impersonal *si* constructions do not (cf. 4b).

- (4) a. On normal hiking trails, one meets a lot of people. But in the desert people are pretty rare
- b.*Nel deserto si è piuttosto rari
In the desert si is rather rare

This entails that Carlson's idea about bare plurals (viz. that they are kind denoting NP's) cannot be directly applied to *si*.

A second property of *si* has to do with its capacity of acting as antecedent for other pronominal elements. It turns out that *si* can antecede a rather limited class of elements. In particular, overt non reflexive

pronominals and the null subject *pro* of Italian cannot be anaphoric to *si*. This is shown by the ungrammaticality of the following sentences (on the indicated readings).

- (5) Property B: *si* cannot antecede overt pronominals and *pro*.
- a.**si_j è detto che loro_j hanno sbagliato*
si said that they were wrong
people_j said that they_j were wrong
- b.**si_j è detto che pro_j vinceranno*
si_j said that pro_j will win.

On the other hand, *si* can antecede both clitic and non clitic reflexives:

- (6) Property C: *si* can antecede reflexives.
- a. *si_j è troppo spesso ingiustificatamente orgogliosi de se' stessi_j*
people are too often proud of themselves for no reason
- b. *Ci si è lavati*
self-cl. *si* aux washed
People washed themselves

A further anaphoric property of *si* is that it can control, i.e. it can act as the antecedent of the null subject PRO of infinitives. This is illustrated by sentences of the following kind:

- (7) Property D: *si* can control
- si_j è cercato di [PRO_j vincere]*
people tried to win

Interestingly, *si* can also antecede other occurrences of *si*. This means that *si* can, as it were, be anaphoric to itself (and this is why it is regarded as a pronoun):

- (8) Property E: *si* can antecede itself
- a. *Se si gioca male, si perde*
If people play badly, they loose
- b. *Ieri, si è giocato male e si è perso.*
Yesterday, people played badly and they/people lost

Structures such as those in (8a), where *si* occurs in an if-clause, bring

up a further property of *si*. In much recent work, if-clauses are generally regarded as restrictions of quantificational structures whose scope is the main clause. The quantifier in constructions of this sort is adverbial, i.e. an adverb of quantification. In sentences like (8a), such a quantifier is taken to be phonologically null, and to have a universal or generic force. It is interesting now to observe what happens to structures like (8a) when an overt quantificational adverb is present.

(9) Property F: variability with if/when clauses.

- a. Se si è alti si è sempre belli
Everyone who is tall is beautiful
- b. Se si è alti, si è talvolta anche belli
Some people who are tall are also handsome

As is evident from the glosses, the quantificational force of *si* appears to be determined by the quantificational adverb. This parallels what happens with indefinites in conditional sentences like:

(10) If an Italian is tall, he is always/often/rarely blond

The phenomenon in (10) has been studied extensively, especially with DRT. It might be useful to recall how classical DRT accounts for the pattern in (10) to see how readily such an account can possibly extend to the pattern in (9). The main assumptions of classical DRT are the following.

(11) Classical DRT

- a. Indefinites are free variables, bound by the closest binder.
- b. If/when clauses form tripartite structures of the form $Q[A][B]$, where Q is a possibly phonologically null adverb of quantification, A is the if-clause and B the main clause.
- c. In certain environments (e.g. in the scope of a quantifier, the text level) indefinites undergo a rule of existential closure.
- d. Quantifiers bind unselectively all the variables in their scope.
- e. Indefinites are subject to a novelty condition, which forces their index to be new.

According to these assumptions, sentences like (10) wind up receiving the following interpretation:

- (12) $Q_x[\text{Italian}(x) \wedge \text{tall}(x)] [\text{blond}(x)]$
where Q can be *always* ($\approx \forall$), *often* ($\approx$ many), *sometimes* ($\approx \exists$), etc.

This accounts nicely for the variability of indefinites under quantificational adverbs. The rule of existential closure (11c) is necessary to give indefinites the existential force they have in ordinary episodic sentences (e.g. *a man walked in*). The novelty condition (11e) is necessary to prevent indefinites from being anaphoric to some other NP. Without the novelty condition, in classical DRT a sentence like (13a) under the given indexing (which is allowed under standard syntactic assumptions) would be interpreted as sentence (13b).

- (13) a. If an Italian_i is dark haired, an American_i is blond
- b. If an Italian_i is dark haired, he_i is an American and he_i is blond

The novelty condition rules out the indexing in (13a), since the index on *an American* is not new in that text.

Classical DRT has also been extended to dealing with the contrast between generic vs episodic sentences, illustrated by examples (2)–(3) above (see on this e.g. Krifka, 1988). The idea here is that the generic operator *Gn* is just a null universal adverb of quantification with a special modal character (which enables it to tolerate exceptions). When an overt if-clause is missing, the restriction for the *Gn*-operator is provided by the context (possibly integrated by material from the main clause). To use Diesing's (1992) terms, in such cases the main clause is partitioned into a restriction and a scope. Thus for example, a sentence like (14a) is interpreted as (14a'), while (14b) is interpreted as (14b'):

- (14) a. A man with taste and money drives a Porsche
- a'. $Gn\ x [\text{man with taste and money}(x)] \exists y [\text{porsche}(y) \wedge \text{drive}(x, y)]$
- b. A computer routes a modern plane
- b'. $Gn\ x [\text{modern plane}(x)] \exists y [\text{computer}(y) \wedge \text{route}(y, x)]$

The primed versions of the sentences in (14) give the most salient interpretations of their unprimed counterparts. In particular, in sentence (14a) the subject is incorporated into the restriction of *Gn* and the object into its scope, while in sentence (14b) the opposite is the case. By assuming that bare plurals are plural indefinites, a similar treatment can be readily

extended to them. Thus, for example, the following sentences would be interpreted in a manner parallel to those in (14):

- (15) a. Men with taste and money drive Porsches
b. Computers route modern planes

Given the quantificational variability of *si* (property A) and its sensitivity to the presence of quantificational adverbs (property F), it is of course tempting to regard *si* as an indefinite and to extend to it a DRT-like approach. This, however, does not per se account for the anaphoric properties of *si* noted in B through E. Moreover, there is also a rather puzzling feature that would remain in an analysis of conditional sentences such as (9) and (10) above along the lines of classical DRT. The puzzle is the following. If *si* is an indefinite, it should be subject to the novelty condition and hence the occurrence of *si* in the main clause of (9) and (10) should not be able to be anaphoric to the occurrence of *si* in the if-clause. In other words, while *si* truly acts as an indefinite in the restriction of quantificational structures, it seems to act as a pronoun (i.e. as a definite) in the scope of such structures. This is a behavior one would like to understand better.

Thus, even if it appears that some version of DRT might be of use in looking at the behavior of impersonal *si* constructions, there clearly remains a number of questions that have to be dealt with. To this task, I now turn.

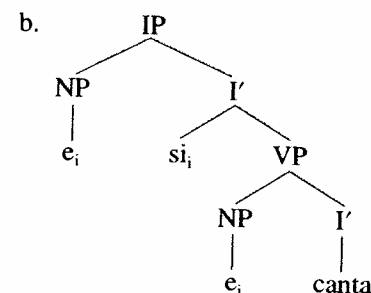
2. BACKGROUND

As a first step towards accounting for the behavior of *si*, I will spell out here some basic assumptions that form the background to my proposal. As far as syntax is concerned, I adopt a version of the Principles and Parameters framework, according to which the grammar of a language is viewed as a device that generates pairs of Phonetic Forms and Logical Forms. I assume that Logical Forms are then truth-conditionally interpreted, by means of a procedure that maps them into a logic (say, a version of Montague's Intensional Logic). Such a logic incorporates certain features of the ideas put forth with DRT. In what follows, I outline some more specific aspects of my background assumptions.

2.1. Syntactic Preliminaries

The syntactic properties of *si* form an interesting and complex area of research in their own right. What I have to say about the quantificational and anaphoric characteristics of *si* is largely neutral, as far as I can tell, with respect to the specifics of its syntax. Following the work of Burzio (1986) and Cinque (1992), I will adopt a structure of roughly the following kind:

- (16) a. *Si canta*
People sing



The basic idea is that *si* occurs somewhere in Infl and is coindexed with an empty expletive in subject position (i.e. in Spec of IP). Since I also adopt the hypothesis that subjects originate in Spec of VP, *si* will be coindexed also with this position (so that it can receive a theta role from the verb). In much recent work, the Infl-node has been further analyzed into a series of functional heads, where information about tense, aspect, and negation is located. I will ignore here the inner articulation of Infl (except when directly relevant to our purposes) and will not try to settle in which functional domain between VP and IP *si* is exactly located. I will also try to stay neutral on other important issues such as whether *si* occurs in a Spec or in a Head position, whether it moves, and the like. I assume simply that the interpretive procedure takes something isomorphic to (16b) as its input.

It has been pointed out in the literature that there are several *si*'s. For example, when *si* is in construction with transitive verbs, we have two options. Either the verb shows singular agreement (i.e. it agrees with *si*). Or else the verb agrees with the object (in these cases, the object tends to be preposed). These options are illustrated in (17a) and (17b) respectively:

- (17) a. Da qui si vede le montagne
 From here si sees (SING) the mountains (PL)
 From here, one can see the mountains
- b. Da qui le montagne si vedono bene
 From here the mountains (PL) si see (PL) well
 From here the mountains can be seen well

Sentence (17a) is only marginally grammatical in my dialect. The form in (17b), which for me is the more standard one, bears many similarities to passive and is sometimes called “*si passivante*”. Cinque (1988) has also argued that besides an argumental *si*, there is also a non-argumental one. In the present paper, I will ignore these distinctions. All the various *si*-constructions appear to have a common core of quantificational and anaphoric properties (namely those presented in the introduction) and it is this common core that constitutes the focus of present theory. My hope is that (at the appropriate level of abstraction) *si*-constructions can be regarded as a semantically unitary phenomenon.

2.2. Semantic Preliminaries

Since the early times of DRT several problems have emerged, that have lead to modifications of various kinds. I want to present briefly the two I will adopt. They concern the form of tripartite quantificational structure and the rule of existential closure.

In classical DRT, a set of construction specific construal rules are assigned to (clause initial) conditionals structures of the following kind:

- (18) a. If a poet_i is inspired, she_i usually writes a poem_j
- b.
- ```

 IP
 / | \
 usuallyi [if a poeti is inspired] [∃j[shei writes a poemj]]

```

In these structures the adverb is sister adjoined to the left of IP so that its restriction is formally identified as its first sister, and its scope as its second sister. Now structures of this kind are inconsistent with the restriction to binary branching structures advocated in most current versions of X'-theory. To make DRT consistent with current tenets of X'-theory, I have proposed in Chierchia (1992b, 1994) to adopt instead the following kind of tripartite structures:

- (19)
- ```

      IP
     /  \
    CP   IP
    |    /  \
    |  usuallyi IP
    |             |
[if a poeti is inspired] [shei writes a poemj]
  
```

These structures can be obtained by a straightforward application of move α from surface structures like (18a) and are readily amenable to a compositional semantics. The scope of the adverb is its C-command domain, while the restriction is the part of its M-command domain which is not also C-commanded (i.e. the material in the same local environment as the adverb, which is not included in the scope).² Besides being in line with current assumptions concerning the syntax of LF, these structures have a further advantage, namely they allow us to do away with the novelty condition, thereby simplifying the construal component further. Consider the problematic cases exemplified in (13a) above repeated here.

- (20) a. If an Italian_i is dark haired, an American_i is usually_i blond

- b.
- ```

 IP
 / \
 CP IP
 | / \
 | usuallyi IP
 | |
[if an Italiani is dark haired] [an Americani is blond]

```

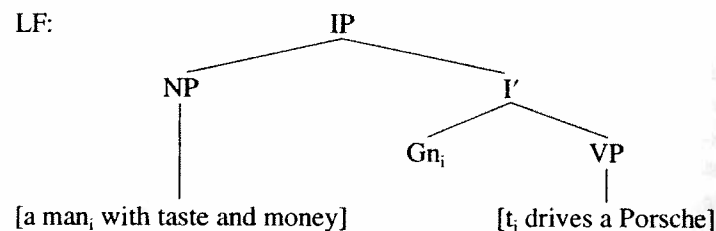
By comparing (20b) with the well formed (19) it is easy to see what is wrong with the former. In (20b) we see an R-expression, namely *an American<sub>i</sub>*, which is bound by the quantificational adverb and its restriction, while no such violation occurs in (19). Thus, we can maintain that (20b) is ruled by the Binding Theory (and, more specifically, by principle C). By assuming the tripartite quantificational structures I am proposing, the effects of the novelty condition can be derived from an independently needed principle of grammar.<sup>3</sup>

The LF for tripartite structures I am proposing also extends naturally to cases where we have an adverb of quantification without an overt *if*-clause. Here too, the adverb can select its scope by adjoining to it,

and its restriction is what is left out of the scope. To illustrate how this works, let us go back to the examples in (14) above. The Logical Forms associated with the intended readings of the sentences in (14) would be as follows:

- (21) a. A man with taste and money drives a Porsche

b. LF:



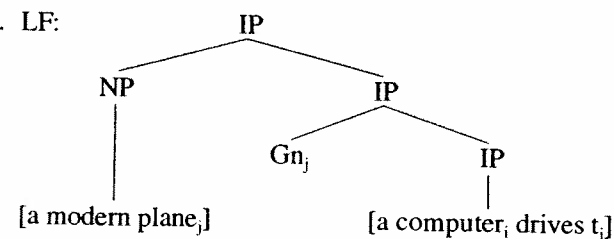
c. Interpretation:

$\text{Gn } x, s[\text{man with taste and money}(x) \wedge C(x, s)] \exists y[\text{Porsche}(y) \wedge \text{drive}(s, x, y)]$

Here the quantificational adverb is the phonologically null generic operator. I assume it is adjoined to a position where it C-commands the VP and is C-commanded by the subject. I also assume a strong version of the Davidsonian hypothesis, whereby every verb has an extra argument ranging over eventualities/situations. Such an argument typically will be bound by any quantificational adverbs present. Moreover, it should be born in mind that on top of the structurally identified portion of the restriction (in (21a), the subject), the context will typically supply further ways of restricting the range of the quantifier. As (21c) shows, I assume that this takes the form of a variable *C*, whose value depends on the context. Formula (21c) is to be understood as follows: for every *x* and every situation *s*, if *x* is a man with money and taste and *s* is a situation containing *x* of the appropriate kind (i.e. the factors that typically trigger driving must be present), then *x* drives a Porsche in such a situation. All this is still an oversimplification, but for our purposes it should do. Let us consider next another example:

- (22) a. A computer routes a modern plane

b. LF:



c. Interpretation:

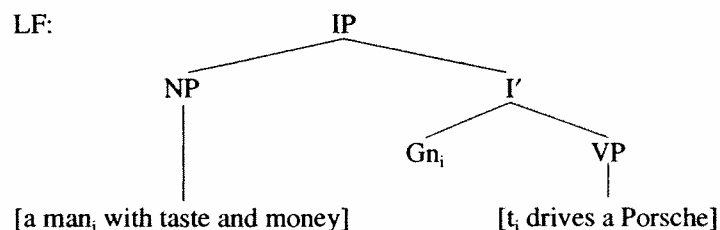
$\text{Gn } x, s[\text{modern plane}(x) \wedge C(x, s)] \exists y[\text{computer}(y) \wedge \text{route}(s, y, x)]$

Here the *Gn*-operator gets adjoined at LF to IP (where it C-commands the subject). The object is then extracted, and thereby winds up in the restriction. The interpretation of the resulting LF is as shown in (22c). I assume that this process of partitioning the clause (that, like Quantifier Raising, is just an instance of move  $\alpha$ ) is basically free, subject to standard conditions on movement and issues of pragmatic plausibility.

This concludes my review of the first modification of classical DRT that I will adopt. The second has to do with the claim that indefinites are essentially free variables. To make this claim stick, it is necessary to assume a rule of existential closure that gives indefinites existential force in the relevant environments. In the early days of DRT it was assumed that such environments were limited to the scope of operators and to the text. But it has subsequently become clear that further environments must be added to these original ones (for example, the restriction of quantificational determiners like *every*, *most*, etc. must also undergo existential closure – cf. below, Section 4).<sup>4</sup> In view of this situation I have proposed in Chierchia (1992a, 1994) to turn the assumptions of classical DRT on indefinites upside down. Rather than viewing indefinites as free variables that get existentially closed in a number of diverse contexts, let us go back to the Fregean approach that treats them as existentially quantified terms. Let us assume, moreover, that adverbs of quantification can, as it were, turn indefinites into free variables, by means of a rule of “existential disclosure” (Dekker, 1990). To see what I mean by this, let us go back to example (21), repeated here:

(23) a. A man with taste and money drives a Porsche

b. LF:



In standard DRT, the indefinites *a man* and *a Porsche* are free variables. *A man* in the restriction of the tripartite structure in (23b) is bound by  $Gn$ , while the existential force of *a Porsche* is obtained by a rule that inserts an existential quantifier coindexed with it adjoined to the scope of (23b). I claim, instead, that *a man* and *a Porsche* each have their own existential quantifier. The operator  $Gn$  (and in general, any adverb of quantification) does two things at once: (i) it wipes out the quantifier from the NP it is coindexed with, turning it into a free variable, and (ii) it binds the resulting variable.

If we think of existential disclosure as a syntactic operation on logical forms, it would amount to deleting some information from such logical forms. Such an operation wouldn't be any less or any more ad hoc than an operation that inserts existential quantifiers, such as existential closure as originally proposed by Heim (1982). However, it turns out that if we adopt a view of meaning somewhat different from the standard one, the operation of existential disclosure need not be thought of as a syntactic operation on logical forms, but becomes a rather natural semantic operation closely related to  $\lambda$ -abstraction. Let me elaborate on this a bit.

The standard view of meaning is based on the assumption that sentences have a certain informational content (the truth conditions they are associated with, or the proposition they express). However, it is also possible to think of sentence meaning in a more dynamic way, as proposed, for example, by Stalnaker (1978). The idea is the following. When illocutionary agents engage in a conversational exchange, they must share a certain body of knowledge. Such a body of knowledge constitutes the information state on which a communication act operates. The utterance of a sentence by an agent (who draws on his or her system of knowledge) modifies the common background. As a result of the utterance of a sentence, the information state originally shared by the

participants is turned into a new one. From this perspective, it is natural to think of sentence meaning as a function that takes as input a common informational background, and turns it into a new one, which successive moves in the communication game can further modify. Sentence meaning can, accordingly, be regarded as a function from common backgrounds into new common backgrounds or, more generally, as a function from contexts into new contexts. This view of meaning can be made precise. It is possible to build up a recursive semantics that rather than assigning to sentences truth conditions or propositions, assigns them context changing functions. In such a semantic system, the operation of existential disclosure constitutes a move as natural as ordinary  $\lambda$ -abstraction in a canonical Tarskian semantics.

I realize that readers not familiar with these ideas will remain unconvinced by my claim. Readers should feel free to think, for the purposes of this paper, of existential disclosure as an operation on logical forms that literally erases existential quantifiers from indefinite NP's (perhaps, by removing the indefinite article *a*). Even so, there is an empirical advantage to this way of proceeding over the alternative based on existential closure. The advantage is the following. Existential closure has to apply to a range of diversified environments that have nothing in common with each other. In contrast, existential disclosure applies to just one environment, namely the domain (= restriction + scope) of quantificational adverbs. Thus at the very least, it affords us a more economical way of describing the facts. My further claim is that adopting a somewhat more abstract, but arguably principled view of meaning enables us to clearly see that the relevant generalization is a natural one. I ask the reader unfamiliar with dynamic semantics to suspend judgement on this latter claim, until having had a chance to check whether it has some basis.

To summarize, the modified version of DRT that I will adopt is based on the following assumptions:

- (24) a. The scope of a quantificational adverb is what it is adjoined to at LF. Its restriction is what is external to the scope (and in the same local environment as the adverb).
- b. Indefinites are existentially quantified.
- c. Adverbs of quantification "disclose" the indefinites they are coindexed with (where indices on quantificational adverbs are assigned freely, subject to standard conditions on binding).

This leads to a radical simplification of the construal component needed to deal with the anaphoric properties of indefinites. More specifically, it enables us to dispense with the novelty condition and with the rule of existential closure. As far as the latter is concerned, the advantage is the fact that while existential closure has to apply to a seemingly arbitrary class of environments, existential disclosure applies to just one environment. In terms of this apparatus, which, to the extent that my claims can be defended, has a certain amount of independent justification, I will now couch my proposal on the semantics of impersonal *si* constructions.

### 3. THE SEMANTICS OF *SI* AND ITS EMPIRICAL CONSEQUENCES

The hypothesis I wish to put forth is that *si* is interpreted as an operation that takes a property and does two things to it: (i) it closes existentially the argument corresponding to the subject, and (ii) it restricts the range of such an argument to groups of humans (perhaps drawn from a contextually specified set). I assume, moreover, that the latter restriction is implemented by choosing a distinguished variable (which, for convenience I will notate as  $x_{arb}$ ) and by stipulating that the range of such a variable is restricted to groups of humans (cf. the Appendix for technical details). There are many other grammatical constructions that involve a restriction to the feature "human", and I do not claim that they all realize this restriction as I am proposing *si* does. I am also aware of the fact that the intended sortal restriction into the semantics of *si* can be implemented in many other conceivable ways. However, using a distinguished index, besides being simple, has interesting consequences concerning the anaphoric properties of *si*, as we shall see.

According to the hypothesis just sketched, the interpretation of a simple sentence like (25a) proceeds as indicated in (25b).

- (25) a. *Si canta*  
       People sing  
       b.  $\exists x_{arb}[\text{dance}(x_{arb})]$   
            $x_{arb}$  = a variable restricted to ranging over groups of humans

In the simplest cases we can regard *si* as a function from properties (of type  $\langle e, t \rangle$ , in Montague's system) into formulae (of type  $t$ ). We can thus define a general functor *SI* of the appropriate type, which will constitute the interpretation of *si*, along the following lines:

$$(26) \quad SI(P) = \exists x_{arb}[P(x_{arb})]^5$$

On the present view, the interpretation of *si* (i.e. the functor *SI*) bears a strong resemblance to what has been proposed for the interpretation of passive (cf. e.g. Dowty, 1978), with two main differences. First, while passive applies to transitive verbs (i.e., at some level, it is an operation on 2-place relations), *si* is not so restricted and can apply to any VP (i.e. it is an operation on 1-place properties). The second difference is that *SI* introduces a sortal restriction to humans that is both syntactically and semantically projected in the form of a distinguished index, while passive does not. This similarity with the semantics of passive comes of course as no surprise, given the many syntactic similarities between the two constructions often discussed in the literature. It should also be noticed that *si* is not assigned an ordinary NP meaning (i.e. a generalized quantifier) but is rather regarded as an operation on properties. This fits rather well with the fact that *si* is a clitic and, consequently, does not have the syntactic properties of ordinary NP's.

This hypothesis on the semantics of *si* is rather simple. Yet, embedded within the background assumptions outlined in Section 2, it enables us to derive in a rather straightforward way all of the six properties identified in the introduction. In what follows, I sketch how. The first group of consequences (discussed in Section 3.1) are derived from the hypothesis that *si* is interpreted as an existentially quantified construction. This part does not make crucial use of the fact that *si* is associated with a distinguished variable. The second group of consequences (discussed in Section 3.2) are centered on the properties of *si* as an antecedent. They are crucially based on the idea that *si* has a distinguished index, and do not hinge so much on its indefinite character. Finally, in Section 3.3, I address the pronominal (i.e. anaphoric) behavior of *si*. The account proposed there crucially exploits the interaction between the indefinite character of *si* and its distinguished index. I hope that dividing things up this way will help the reader to follow my proposal and choose what to leave and what, possibly, to take.

#### 3.1. Quantificational Variability in Generic vs. Episodic Sentences

A sentence like (25a) above is expected to be ambiguous depending on whether it is understood generically or episodically. Formally, this ambiguity is represented in terms of the presence vs. absence of a generic

operator in the LF of the relevant sentence. Formula (25b) represents the case where *Gn* is absent. Such readings are prominent in contexts of the following kind:

- (27) a. Q: Che sta succedendo qui? A: si canta  
Q: What is going on here? A: people are singing
- b. Q: Cosa è successo ieri in campeggio? A: Si è cantato  
Q: What happened yesterday in the campground? A: People sang

These question-answer pairs bring out the episodic reading of *si*. The truth conditions we get, for example in case (27b), is that there is a group of people, from a contextually specified set, (which may or may not include the speaker or the hearer and may or may not extend to the totality of the relevant people) that is engaged in singing. These truth conditions appear to be roughly adequate.

Let us now turn to the case where (25a) is interpreted generically. This will happen when there is a *Gn*-operator in the Logical Form of (25a). There are two possible structures of this kind, depending on how the scope of *Gn* is selected:

- (28) a.
- ```

      IP
     / \
    NP  I'
    |   / \
    |  I  VP
    |   / \
    |  Gn  VP
    |   |
[eexpl] [siarb] [earb canta]
  
```
- b.
- ```

 IP
 / \
 Gn IP
 |
 [siarb canta]

```

In (28a) *si* is construed as having scope over the *Gn*-operator. In (28b) the *Gn*-operator has scope at the IP-level and thus has scope over *si*. Let us consider now the readings that our theory assigns to such logical forms, starting with (28b). In structure (28b), *Gn* lacks a structurally

projected restriction. Moreover, since the argument of *cantare* is existentially closed by *si*, the only possibility is for *Gn* to bind the event/situation variable. We thus get an interpretation of the following kind:

- (29)  $Gn\ s[C(s)]\ [\exists x_{arb}\ sing(s, x_{arb})]$

Formula (29) says that in every contextually relevant situation, there is singing going on. This means that unless a suitable value for *C* is recoverable from the context, we will not be able to interpret (28b). Cases in which we will be able to interpret (28b) are, for example, the following:

- (30) a. Q: Cosa si fa' di solito in campeggio? A: si canta  
what si does usually while camping? A: si sings  
What do people do usually while camping? A: they sing.
- b. In campeggio si canta  
while camping, si sings  
While camping, people sing.

In (30a), the previous question makes it clear how the variable *C* is to be interpreted. In (30b) we actually have a suitable syntactically overt restrictor in the form of a locative. In either case, the *si* construction winds up being interpreted as saying: take every situation where one is camping (and where the right triggering factors for singing are present), in any such situation there is singing going on. This looks like a plausible account for the quasi universal force of generic *si* constructions.

Let us turn now, to the LF (28a). In this case, we first form a generic property that then gets existentially closed by *si*. The corresponding truth-conditions are the following:

- (31)  $\exists x_{arb}\ Gn\ s[C(s, x_{arb})]\ [sing(s, x_{arb})]$

This formula says that there is some group of people (perhaps drawn from a salient set) that has the habit of singing (i.e. such that in every situation involving them where the right triggering factors are present, they sing). In isolation, this reading is hard to get. It corresponds to the "specific indefinite" reading of a sentence like (32a).

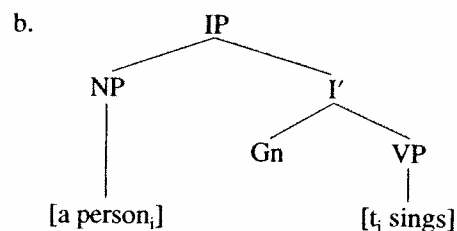
- (32) a. A person I know sings  
b. A person sings

My approach has only pragmatic reasons to offer as for the oddity of a specific indefinite (generic) reading of *si* constructions taken in isolation.

The thing to note is that overt indefinites display a similar behavior. For a specific indefinite reading to be possible with overt indefinites, the restricting property has itself to be fairly specific as in (32a). When the property is not itself sufficiently specific as in (32b), it is very hard to get a specific indefinite reading: (32b) cannot readily be interpreted as (32a). So in this regard, the behavior of *si* is fully parallel to that of indefinites, with one crucial difference: there is no way to overtly narrow down the restricting property for *si* as one can do for overt indefinites (like 32a), which makes the specific indefinite reading for generic *si* constructions in isolation hard to get. Thus the line I take is that while (28a) is well formed and results in interpretation (31), when sentence (25a) is uttered in isolation this reading is hard to get for pragmatic reasons.

There is a further related consequence of the present view. A sentence like (32b) can be understood as being about people in general. This is evidenced by contexts of the following kind.

(33) a. A cat meows. A person sings



c.  $Gn\ s, x[\text{person}(x) \wedge C(s, x)]$  [sing(s, x)]

In discourses like (33a), the natural interpretation of the second sentence is the one given in (33b, c). In such an interpretation, the property *person* is incorporated into the restriction of *Gn*, reflecting the fact that we are talking about people in general. Our theory predicts that generic *si*-sentences cannot be interpreted in a parallel manner. *Si* does not have a syntactically projected restricting property; it is associated with a sortally restricted variable (i.e. a variable that comes with a built-in domain restriction). Hence, the restriction of *si* (while being similar to that of *a person*) stands no chance of being used to restrict *Gn* or some other quantificational adverb. This means that a sentence like *si canta* cannot be understood as attributing a generic property to the kind 'human'. It always needs to be further contextualized. As far as I can tell, this is empirically correct.<sup>6</sup>

To summarize so far, we have considered in the present section the interpretation of sentences with *si* that occur without an if/when clause (property A, sec. 1). Our theory, together with the version of DRT sketched in Section 2, predicts that the quantificational force of *si* in these cases should correlate with the presence vs absence at LF of the Gn-operator (or of analogous overt quantificational adverbs). In this regard, *si* behaves like an ordinary indefinite. However, unlike overt indefinites, *si* lacks a syntactically projected restricting property. Consequently the interpretive possibilities of sentences with *si* are more restricted than those of overt indefinites in parallel constructions. This seems to be borne out empirically.

### 3.2. *Si* as an Antecedent

It was noted in Section 1 that *si* cannot bind overt non-reflexive pronouns or small *pro* (property B). On the present view, this follows immediately from the fact that binding requires coindexing. But the index on *si* is a distinguished one, i.e. one which is different from the indices that get freely instantiated on NP's. This was justified in part as a way of implementing formally the sortal restrictions on *si*, which cannot be coindexed with other NP's. It follows, as a special case, that *si* cannot bind non-reflexive pronouns or small *pro*'s. The following examples illustrate.

(34) a.  $Si_{arb}$  è sostenuto che Gianni  $li_i$  volesse provocare  
 Si aux-3rd-SING claim-PAST-PART that Gianni them-CL  
 wanted-SUBJ provoke

People<sub>i</sub> have claimed that Gianni wanted to provoke them<sub>i</sub>

b.  $Si_{arb}$  è sostenuto che  $pro_j$  volessero provocare Gianni  
 Si aux-3rd-SING claim-PAST-PART that  $pro_j$  wanted-SUBJ  
 provoke Gianni

People<sub>i</sub> have claimed that they<sub>j</sub> wanted to provoke Gianni

In (34a, b) the clitic *li* and *pro* cannot be interpreted as anaphoric to *si*. This follows simply from the fact that the variable associated with *si* carries an index (notated as 'arb') which is distinct from the indices of other variables.

Now an interesting problem is raised by the fact (discussed in e.g. Suárez (1982)) that *pro* in certain contexts does admit so called arbitrary or indefinite interpretations. For example:

- (35) *pro* ti hanno cercato  
 pro you-CL aux-PAST look for-PAST-PART  
 somebody was looking for you

The natural interpretation of *pro* in (35) is as an indefinite, as the gloss shows. Yet, even under this interpretation *pro* cannot be understood as anaphoric to *si*, as (34b) illustrates. This is puzzling because while not identical,<sup>7</sup> the behavior of *si* and indefinite *pro* is in several key aspects similar. For example, both *si* and indefinite *pro* appear to be restricted to groups of humans, both receive an existential interpretation in episodic sentences and a quasi universal one in generic sentences, and so on. In light of this similarities, one could claim that *pro* can also get the index *arb*, with the associated semantic effects of restricting the interpretation to sets of humans. But then one would expect that *pro* under the arb-interpretation can be anaphoric to *si* (and vice versa), contrary to fact.

I do not have a completely non-stipulative account as to why this is so, other than insisting on the distinctiveness of *si*'s index, even with respect to indefinite interpretations of *pro*. It should be noticed, however, that there are independent reasons for maintaining that the indices associated with *si* and indefinite *pro* must be distinct. They have to do with a contrast in person features between *si* and indefinite *pro*. Compare (35) with its counterpart with *si* given below:

- (36) ti si è cercato  
 you-ACC-CI si AUX-PAST look for  
 People/we looked for you

The interpretation of *pro* in (35) is that of a 3rd person pronoun and clearly excludes the speaker. The interpretation of *si*, on the other hand, not only does not exclude the speaker, but generally favours a speaker-oriented interpretation (as the gloss in (36) suggests). Whatever the correct analysis of this contrast turns out to be, it is clear that indefinite *pro* and *si* must differ in  $\phi$ -features. Now it is generally maintained that if two items carry the same index, they must share the same  $\phi$ -features (i.e. they must agree). It follows then, by contraposition, that if two items don't have the same  $\phi$ -features, then they cannot have the same index. This suffices to prevent *si* from anteceding *pro* even on its indefinite interpretation and corroborates our assumption about the distinctness of *si*'s index.

We can now address the question as to why *si* can act as the antecedent of reflexives (property C). As evidenced especially within the semantic literature on reflexivization (cf. e.g. Bach and Partee (1980)) reflexive marking signals that two slots of a relation have been identified. So, for example, if a two place relation *R* is reflexively marked (whether by the clitic or by the non clitic reflexive pronoun), its interpretation will be along the following lines.

- (37)  $\text{refl}(R) = \lambda x[R(x, x)]$

A reflexive property such as (37) then gets predicated of the subject. It follows that the relationship between a reflexive and its antecedent is not the same as the relation between a non-reflexive pronoun and its antecedent. The antecedent of a reflexive does not directly bind it. Hence the distinctness of *si*'s index plays no role, as binding is not relevant. So we wind up with structures of the following kind:

- (38) a. *si* è orgogliosi di se stessi  
           *si* is proud of one self  
       b.  $\exists x_{arb}[\text{self}(\text{proud-of})(x_{arb})]$

In (38b) a reflexive property is attributed to a group of humans, which yields the correct interpretation for (38a).

What I just said can be restated in slightly different terms. We can think of the relation between a reflexive and its antecedent as articulated in two parts, schematically indicated in what follows:

- (39) John<sup>i</sup> O<sub>i</sub><sup>j</sup> [likes himself<sub>i</sub>]

The first component of the structure in (39) is a relation between the reflexive and an operator (perhaps to be identified with the reflexive itself, raised to an A'-position). The operator is interpreted as a  $\lambda$ -abstractor whose semantic role is that of identifying two slots of the predicate, as illustrated in (37). The second component of (39) is a relation between the operator and the subject (indicated in (39) by cosuperscripting only for convenience). This is essentially a predication relation and does not involve referential indices. Hence an NP can antecede a reflexive regardless of whether the former shares its referential index with the latter.

That this view (which can be implemented in many ways – see e.g. Reinhart and Reuland (1993) for a recent contribution) might be on the

right track is corroborated by the behavior of reflexives in VP-ellipsis constructions such as:

- (40) John likes himself and Mary does too.

The VP-proform *does* in (40) must be interpreted as the reflexive property of self-liking, even though *John* and *Mary* have distinct  $\phi$ -features and hence distinct referential indices. As Williams (1979) and Sag (1976) have shown, under the view that there is a reflexive operator associated with a reflexivized VP, defining a notion of structural identity between the VP anaphor and its antecedent presents no particular problem. Under the view that the antecedent directly binds the reflexive, such a task appears to be more problematic.<sup>8</sup>

Let us now turn to the case of control (property D). Infinitives are generally taken to be clausal structures with a PRO subject. It has been argued, however, in a number of works that semantically infinitives do not denote propositional creatures (like sets of worlds or states of affairs) but property-like creatures.<sup>9</sup> On the basis of this view, a sentence like *John tries to win* is to be interpreted as something like "John tries to have the property of winning", and the control relation must be understood as a kind of predication relation along the lines suggested in Williams (1980). To exemplify:

- (41) a. John<sub>i</sub> tries O<sub>j</sub> [PRO<sub>j</sub> to win]  
b. try (john,  $\lambda x$  win(x))

PRO in (41a) is abstracted over by an operator (perhaps to be thought of as PRO itself raised to an A'-position, as in the case of reflexives), giving rise to a derived predicate. This derived predicate is then indirectly attributed to the matrix subject, as proposed by Williams.<sup>10</sup> The semantics of (41a) is (41b), which says that John bears the try-relation to the property of winning.

I will not try to defend here this position, which was dubbed in Higginbotham (1992) the "attributive view of control". What I am interested in pointing out are the consequences of this view for the anaphoric potential of *si*. The attributive view of control assimilates the latter to a form of indirect predication, rather than to quantifier binding, much like what was argued for in connection with reflexives. In a control structure, a derived predicate is formed, which is then indirectly predicated of the controller. Hence, as with reflexives, the

referential index of the controller plays no direct role, and cannot interfere with the control relationship. In the cases where *si* is involved, we have structures of the following kind:

- (42) a.  $si_{arb}^i$  tries O<sub>j</sub> [PRO<sub>j</sub> to win]  
b.  $\exists x_{arb}$  [try ( $x_{arb}$ ,  $\lambda x$  win(x))]

The interesting thing in this connection is that under the predicative view absolutely nothing special needs to be said as to the ability of *si* to control.

There is a clear generalization that emerges from the cases of reflexives and control just discussed. Whenever a dependency between an NP and a pronominal element is achieved by forming a predicate and attributing it directly or indirectly to an argument, *si* will be able to antecede that pronominal element. The reason why this is so is that *si*'s distinguished index prevents it from binding but does not limit in any way its capacity of being attributed a property. If this is correct, it ought to generalize to *any* relationship where an anaphoric dependency is mediated by an operator (arguably interpreted as an abstractor and hence as a predicate-forming operator). To check whether this holds we need to look at the major construction types that have been argued in the literature to involve operators.

One such major construction type involves parasitic gaps. However, our prediction cannot be tested with structures of this sort, as parasitic gaps involve A'-binding, while *si* can only be an A-binder (i.e. it cannot head a relative clause, nor does it have a wh-counterpart). Besides parasitic gaps, there is at least another important class of cases that has been argued in the literature to involve operator-binding and does allow for antecedents in A-position, namely the so called tough-movement construction. Since at least Chomsky (1981) and in much subsequent work, the structure of tough-movement constructions has been taken to be some variant of the following (I give directly the Italian case):

- (43) a. Gianni è difficile da accontentare.  
Gianni is hard to please.  
b. Gianni<sub>i</sub> è difficile O<sub>j</sub> [PRO da accontentare t<sub>j</sub>]

In (43b), the non subject gap is bound by an operator. The resulting derived predicate is then indirectly predicated of the matrix subject. If our generalization is correct, we would expect it to be possible for *si* to be construed with the operator in (43b). Empirically this means

that *si* should be able to act as the antecedent of the non-subject gap in constructions of this sort. This expectation is borne out:

- (44) A quindici anni, *si* è difficili da accontentare.  
 At fifteen years, *si* is difficult to please  
 At fifteen, one is difficult to please.

In (44) *si* is obligatorily construed with the object gap of the infinitival.

To wrap up this section, we noted in the introduction that *si* is rather choosy as an antecedent, allowing certain pronominal elements to be anaphoric to it and disallowing others. The hypothesis that *si* is associated with a distinguished index (into which its peculiar composition in terms of  $\phi$ -features and its sortal restrictions are encoded), in combination with an independently motivated hypotheses on the nature of operator-variable dependencies, provide a rather general account of this aspect of the behavior of *si*. *Si* cannot bind because binding involves coindexing and *si*'s index is peculiar to it. However, nothing prevents *si* from entering into predication relations (mediated by an operator), for in the latter case referential indices are not directly relevant.

### 3.3. *Si* as an Anaphoric Element

We noted in the introduction that *si* can, as it were, antecede itself (property E) and that when this happens in if/when clauses, *si* displays strong quantificational variability effects under quantificational adverbs (property F). Some relevant examples are repeated here:

- (45) a. Se *si* è alti, *si* è sempre anche biondi.  
 If one is tall, one is also always blond.  
 b. Ieri, *si* è giocato male *si* è perso.  
 Yesterday, people played badly and they/people lost.  
 c. *Si* è detto che *si* sarebbe intervenuti.  
 People said that they/people would intervene.

Prima facie these cases all look alike: the first occurrence of *si* sets up a discourse referent to which subsequent occurrences are anaphoric. However, the present theory makes sharply different predictions about

them. To see what they are, consider (45b) first. (The case of (45c) is analogous to (45b)). Sentence (45b) involves two occurrences of *si* in coordinated episodic sentences. We are assuming that each occurrence of *si* closes existentially the VP. Hence the interpretation of (45b) is going to be:

- (46)  $\exists x_{arb} [\text{play badly}(x_{arb})] \wedge \exists x_{arb} [\text{loose}(x_{arb})]$

As (46) makes clear, under the present approach the second occurrence of *si* is not, in fact cannot be, anaphoric to the first one. Yet the most natural interpretation of (45b) is one according to which those who played badly lost. Clearly, the present theory forces us to regard this intuition as a purely pragmatic phenomenon. More specifically, one can claim that the apparent anaphoric dependence in (45b) is rooted in two things. First, the contextually supplied group of humans is presumably the same for both occurrences of  $x_{arb}$ . Second, the nature of the predicates in (45b) favors an interpretation whereby the value of  $x_{arb}$  that verifies the first disjunct also verifies the second. If in structures like (45b) we switch to other predicates, this second effect vanishes:

- (47) Ieri *si* è vinto e *si* è perso.  
 Yesterday, people won and people lost.

This sentence has a non-contradictory reading, and its most natural interpretation is one according to which some subgroup of a specified set of humans won and a different subgroup lost (though an interpretation where the subgroup is the same is also possible).

We might modify our theory so as to provide a semantic account for the anaphoric behaviour of *si* in (45b). Since we are dealing with indefinites, one might be tempted to extend existential disclosure to cases like (45b), turning the second occurrence of *si* into a variable, which then can be made anaphoric to the first occurrence of *si*. Essentially, one might set things up in such a way that *si* automatically discloses any subsequent "active" occurrence of *si*. There are various conceivable ways in which this could be done rather elegantly in a dynamic setting. However, this strategy would also lead to a somewhat more complex theory. In particular, one would have to relax the strong claim that only quantificational adverbs can turn indefinites into variables. As far as examples like (45b, c) are concerned, the present and stronger version of the theory, coupled with simple pragmatic considerations, appears to be viable and

hence should be preferred. So far, I have not encountered examples that would force one to resort to a less constrained approach.

Things are different for if/when clauses such as the one in (45a). On the most prominent reading of (45a), a single instance of a tall person that is not blond (from a contextually salient class) suffices to falsify it. Moreover, how many humans that satisfy the antecedent must also satisfy the consequent varies with the choice of quantificational adverb, as we saw. Thus it is clear that there is no alternative to treating both occurrences of *si* in (45a) as variables bound by the quantificational adverb, in the spirit of DRT. Within the present framework, this is what we would expect. Existential disclosure, which turns indefinites into free variables, applies freely in the domain of quantificational adverbs. Thus, in particular the quantificational adverb *always* in (45a) can first disclose and then bind the indefinite *si*, yielding the desired reading.<sup>11</sup> The LF structure of (45a) is given in (48a) and its interpretation in (48b):

- (48) a.
- 
- b.  $\forall x_{arb} [\text{tall}(x_{arb})] [\text{blond}(x_{arb})]$

Different choices of quantificational adverbs in (48a) would correspond to different quantifiers in (48b). Sentence (45a) ought to be compared with its counterpart containing overt indefinites, which is ungrammatical on the intended reading:

- (49) a.\*Se un italiano<sub>i</sub> è alto, un italiano<sub>i</sub> è sempre anche biondo  
if an Italian<sub>i</sub> is tall, an Italian<sub>i</sub> is always also blond

- b.
- 

On the basis of the hypothesis set forth in Section 2, the ungrammaticality of (49a) is due to principle C of the binding theory. As (49b) shows, at the level where the scope and restriction of the quantificational adverb are structurally identified (i.e. LF), the second occurrence of the non-pronominal NP *un italiano* is coindexed with the C-commanding operator *sempre*, in violation of principle C that requires such NP's to be free (and subsumes the effects of Heim's Novelty Condition, along the lines discussed in Section 2). In contrast with this, *si*, while being an indefinite (and hence interpreted as an existentially closed variable), is syntactically a pronominal element and thus exempt from principle C. Consequently, structure (48b) is well formed.

In the present theory, adverbs of quantification are *free* to disclose the indefinites in their domain. This means that *si* doesn't have to be disclosed. Cases where *si* does not appear to be disclosed are the following:

- (50) a. Se *si* prepara un compito<sub>i</sub> così', di solito, non ci<sub>i</sub> *si* capisce niente.  
If *si* prepares an assignment this way, usually not in it- CL *si* understands anything  
When an assignment is prepared this way, people are unable to understand it.
- b. Most<sub>i</sub> [assignment(y)  $\wedge$   $\exists x_{arb}$  prepare ( $x_{arb}$ , y)] [ $\neg \exists x_{arb}$  understand( $x_{arb}$ , y)]

Sentence (50a) is about assignments. On the most natural reading, the quantificational adverb *di solito* binds the variable associated with the NP *un compito*, not the one associated with *si*. In fact, it is intuitively clear that for (50a) to be true the people who prepare the assignment and those who fail to understand it need not (and typically will not) be the same. Thus the indexing that yields the most natural reading of (50a) is the one given there, which corresponds to the interpretation in (50b).

Another example is the following:

- (51) a. Se un televisore nuovo è difettoso, lo *si* aggiusta.  
If a new TV set is defective, it-CL *si* fixes  
If a new TV set is defective, one fixes it.
- b.  $\text{Gn } y [\text{new TV set}(y) \wedge \text{defective}(y)] \exists x_{arb} [\text{fix}(x_{arb}, y)]$

Here *si* occurs only in the scope of the (phonologically null) quantifi-

cational adverb. Hence it cannot be disclosed, for otherwise the Gn-operator would be binding something which is not in its restriction, in violation of a natural principle of well formedness on quantificational structures. So, it is expected that in these cases *si* can only receive an existential interpretation, such as the one given in (51b), which seems to be right.

There is a further interesting consequence of the present theory. Since our assumption is that indefinites are disclosed only by adverbs of quantification and not also by, say, quantificational determiners like *every* or *most*, we would not expect donkey-type dependencies between occurrences of *si* such as the ones illustrated in (45a) to be possible with relative clauses. This seems to be correct. Consider:

- (52) Ogni persona che *si* inviti per una conferenza *si* aspetta che la *si* tratti bene.  
 Every person that *si* invites for a talk expects that *si* treats him or her well  
 Every person that gets involved for a talk expects to be treated well.

The sentence is chosen so as to pragmatically maximize the possibility of interpreting the second occurrence of *si* as anaphoric to the first one. And yet, the interpretation of (52) does not require that the people that do the inviting and those who do the actual hosting be the same. This follows from the present approach, as each occurrence of *si* is existentially closed. Hence different occurrences are not semantically forced to covary.

To summarize, the anaphoric behavior of *si* in if-clauses and in a number of related constructions seems to follow in a rather direct way from the present approach. This concludes our derivation of the properties of *si* observed in the introduction.

#### 4. SOME COMPARISONS AND THEORETICAL CONSEQUENCES.

The approach I have developed is based on the following assumptions: (i) indefinites are existentially closed; (ii) they can be disclosed by adverbs of quantification; and (iii) they are not subject to the novelty condition but to the binding theory. These assumptions have been motivated on the basis of evidence that has nothing to do with impersonal subjects (cf. Chierchia, 1992b, 1994). Yet they enable us to derive the

bulk of the properties of impersonal subjects with the help of one construction-specific assumption, namely that the index of *si* is a distinguished one. The key factor is that indefiniteness is a property of both pronominal and non pronominal NP's. A pronominal indefinite can in the appropriate configuration be anaphoric to a suitable antecedent. A non-pronominal indefinite cannot.

It might be useful to discuss my proposal with respect to some alternatives that might come to mind. This will help bring out some general theoretical consequences of the present theory. Suppose, to start, that we tried to analyse *si* just as a free variable over groups of humans. According to this hypothesis, prototypical *si*-constructions would have logical forms of the following kind:

- (53) a. Ieri *si* è bevuto molto.  
 Yesterday *si* drank a lot  
 b. drank ( $x_{arb}$ )  
 (54) a. In Italia *si* beve molto.  
 In Italy *si* drink a lot  
 b. in Italy drink a lot ( $x_{arb}$ )

Sentence (53) is episodic. The referent of  $x_{arb}$  is whichever group of people is salient in the context. Sentence (54) is generic; it attributes the property of having the disposition to drink to its argument. If we utter (54) out of the blue, the only salient group of people would presumably be the Italians. Thus (54a) can be construed as attributing to italians the disposition to drink, which seems adequate. Thus sentences like (53) and (54) appear to have a viable analysis on the hypothesis that *si* is simply a free variable. Moreover, by maintaining that such a variable has a distinguished index, one could account for the behavior of *si* as an antecedent as developed in Section 3.

There are various problems, however, that I see with this alternative. The main ones have to do with sentences like (45a), repeated here as (55a):

- (55) a. Se *si* è alti *si* è sempre anche biondi.  
 If one is tall one is always also blond  
 b. ??If they are tall they are always also blond.

The problem here is that if *si* was simply a free variable, it should act as a definite in contexts like (55a) and hence it should not be able to

be quantified over by adverbs of quantification. We would expect, in other words, sentence (55a) to have a status similar to (55b). But while sentence (55a) is perfect, sentence (55b) is rather marginal. Its marginality can be imputed to the fact that the adverb of quantification *always* has no variable to bind (since the value of *they* has to be fixed by the external context). In so far as I can see, the hypothesis that *si* is simply a free variable is incompatible with its behavior in *if/when* clauses, where *si* appears to get bound by quantificational adverbs, just like overt indefinites. One can of course stipulate that *si* is interpreted as a free variable except when an adverb of quantification is around. But this would be completely ad hoc.

A different alternative would be to try to stick as much as possible to classical DRT by making the following moves. In classical DRT NP's are marked as + or - DEF(inite). NP's that are - DEF, are subject to the novelty condition while definite NP's must instead be interpreted as "old" or "familiar" variables. Now, one could assume that certain NP's are actually lexically unmarked with respect to the feature +/-DEF. This can be taken to imply that *si* is freely interpreted as either definite or indefinite depending on the context. Put in other terms, we could say that *si* is ambiguous between a definite and an indefinite interpretation. It is after all certainly not inconceivable that some NP's might be ambiguous in exactly this way.

This approach would get us the intended readings for the examples we have considered. For example, the prominent reading of (55a) would be obtained by assuming that the first occurrence of *si* in (55a) is an indefinite (and hence liable to being bound by quantificational adverbs), while the second occurrence is a definite and hence anaphoric to the first. However, it is easy to see that such an approach will massively over-generate. For example, we ought to be able to find readings of (55a) where just the opposite happens: the first occurrence of *si* is interpreted as a definite and the second as an indefinite. I.e. by manipulating the context, we should be able to come up with a reading where the first occurrence of *si* is anaphoric to some suitable antecedent previously established in discourse, and the second occurrence is indefinite. However, no amount of context manipulation seems able to deliver such a reading, as examples of the following kind illustrate:

- (56) Ieri si è giocato male. Se si gioca male, spesso si perde.  
Yesterday people played poorly. If one plays poorly one loses.

In spite of the presence of a possible antecedent for the second occurrence of *si* in (56), it is just impossible to interpret it as anaphoric. If *si* is simply ambiguous between a definite and an indefinite interpretation, the impossibility of this reading would be mysterious.

One could also try to recast my proposal in straight DRT along the following lines. Assume first the approach to tripartite structures sketched in Section 2, whereby the adverb of quantification selects its scope by adjoining to it at LF. This enables us to dispense with the novelty condition in favour of a binding theoretic account of the relevant phenomenon, just as I have proposed. Assume then that indefinites are treated as free variables that get existentially closed when no adverb of quantification is around and in the scope of operators. Assume finally that *si* carries a distinguished index. An approach along such lines would be roughly equivalent to the one explored in the present paper, and choosing between the former and the latter becomes a kind of "in house" dispute among different variants of the same basic insight. There is however, an area where the two approaches in question do differ. It has to do with sentences like (52), repeated here:

- (57) Ogni persona che si inviti per una conferenza si aspetta che  
la si tratti bene.  
Every person that gets invited for a talk expects to be treated  
well.

Here *si* doesn't occur in the scope of an operator, nor is it bound by the text-level existential closure. Yet it is interpreted as existentially closed. On the theory put forth in Sections 2 and 3, this is just what we would expect. On the DRT variant we are presently considering one must add a further environment to the rule of existential closure, namely the restriction of quantificational determiners. This is in fact an instance of a general problem hinted at in Section 2 and having to do with proportional quantification. If we start out by treating indefinites as free variables and close them existentially, we end up with a set of disjoint environments to which existential closure must apply. If we treat indefinites as existentially closed, we can simply say that they can be "reopened" by quantificational adverbs. The latter move is a natural one only in a dynamic semantic setting.

There is one further interesting point to make in this connection. One of the theoretical lines on anaphora that is being studied most exten-

sively is the so called E-type strategy. This is based on the idea that pronouns can go proxy for definite descriptions whose content is retrieved from the context. This strategy is viewed sometimes as complementary and sometimes as an alternative to DRT. Let me give the flavor of the basic insight by means of an example modelled on a proposal by Heim (1990). On an E-type analysis a sentence like (58a) would be interpreted roughly as indicated in (58b).

- (58) a. If someone is in Rhodes he is not in Athens.  
 b. Any minimal situation *s* in which someone is in Rhodes is part of a situation *s'* in which the person who is in Rhodes in *s* is not in Athens.

There are many conceivable variants on this strategy. However, the example should make it clear what the general line of approach is. Conditionals involve quantification over situations (or events). Pronouns in the consequent are interpreted as definite descriptions relativized to the situation introduced in the antecedent. This strategy can be developed in a precise and compositional manner. In whichever way its details are going to be worked out, it is pretty clear, however, what problems are going to arise in trying to extend it to impersonal subjects. Consider again:

- (59) a. Se *si* è alti *si* è anche biondi.  
 b. Any minimal situation *s* in which *someone* is tall is part of a situation *s'* in which *the person who is tall in s* is also blond.

The problem with (59a) is that the first occurrence of *si* has to be interpreted as being existentially quantified (*someone* in (59b)), while the second has to be interpreted as a definite description (*the person who is tall in s* in (59b)). It is difficult to see how a theory based solely on the E-type strategy is going to be able to predict where *si* has to be interpreted as an indefinite and where as a definite.

In spite of the cursory character of these remarks, what seems to emerge from them is that to come up with an alternative to the present theory that has fewer construction-specific assumptions appears to be a non trivial task.

## 5. CONCLUDING REMARKS.

We have adopted a version of DRT based on the following assumptions:

- (i) Adverbs of quantification select their scope by adjoining to it at LF. They are freely coindexed with indefinites in their restriction.
- (ii) Non pronominal indefinite NP's coindexed with a C-commanding adverb are ruled out by principle C. Pronominal indefinite NP's are allowed instead.
- (iii) Indefinites are interpreted as existentially closed. Adverbs of quantification turn them into variables.
- (iv) Impersonal *si* is a pronominal indefinite with a distinguished index.

We have seen that the bulk of the quantificational and anaphoric properties of impersonal subjects follows from these assumptions (of which only (iv) is construction-specific). We have also seen that it appears to be difficult to derive such properties with fewer stipulations. This seems to lend support to a DRT based approach to indefinites. More specifically, it seems to support a version of such a theory that employs existential disclosure, an operation which can be most naturally couched in a dynamic semantic setting.

Independently of how convincing these conclusions may be, it is clear that the behavior of impersonal pronouns such as *si* are not only interesting in their own right but also offer an important testing ground for any theory of quantification and anaphora. Only by extending current debates to new structures such as those we have began to consider here can one hope to make some progress in understanding the grammar of anaphoric dependencies.

## APPENDIX

Intensional Logic with distinguished variables.

Let  $x_{arb}$  be a distinguished variable of type *e* of IL. Let us assume models for IL of the following form:

DEFINITION 1. A model for IL is a tuple *M* such that

$M = \langle U, W, I, U^*, F \rangle$  where

- (i) *U* is the domain of individuals, which include also groups or

plural entities (cf. e.g. Landman (1989) or Schwarzschild (1991), among many others) and eventualities (cf. e.g. Parsons 1990).

- (ii)  $W$  is the set of worlds.
- (iii)  $I$  is the set of instants.
- (iv)  $U^*$  (a subset of  $U$ ) is a set of groups of humans.
- (v)  $F$  is an interpretation function (defined as in IL).

The definition of the set  $D_a$  of possible denotations of type  $a$  proceeds as for standard IL (cf. e.g. Dowty, Wall and Peters (1981)). The notion of assignment to variables is modified as follows:

**DEFINITION 2.** An assignment is a function  $g$  from IL's variables into values of the appropriate type, such that for any  $v_{n,a}$  if  $n \neq \text{arb}$ , then  $g(v_{n,a}) \in D_a$  and if  $n = \text{arb}$ ,  $g(x_{\text{arb}}) \in U^*$ .

The definition of  $\llbracket \cdot \rrbracket$  remains the same as the one for standard IL. However, given the restriction on assignments embodied in Def. 2, when quantifiers (or the  $\lambda$ -abstractor) bind  $x_{\text{arb}}$ , the values that  $x_{\text{arb}}$  can take are restricted to  $U^*$ . In a version of IL with indexicals, the specification of the domain  $U^*$  should be incorporated as part of the specification of the context, as is commonly done for indexicals.

I now turn to defining the semantics for *si*, taking into account that VP's can be individual-taking functions or quantified NP-taking functions. I will limit myself to a consideration of the extensional case, the intensional one being a routine modification of the former. Assuming that NP's can be of type  $e$  or of type  $gq = \langle \langle e, t \rangle, t \rangle$  (where "gq" stands for "generalized quantifiers"), VP's can be of type  $\langle e, t \rangle$  or of type  $\langle gq, t \rangle$ . We now define a polymorphic operator SI of type  $\langle a, t \rangle$ , where  $a = e$  or  $a = gq$  as follows.

**DEFINITION 3**

$$SI(P) = \begin{cases} \exists x_{\text{arb}} P(x_{\text{arb}}), & \text{if } P \text{ is of type } \langle e, t \rangle \\ \exists x_{\text{arb}} P(\lambda Q Q(x_{\text{arb}})), & \text{if } P \text{ is of type } \langle \langle \langle e, t \rangle, t \rangle, t \rangle \end{cases}$$

#### NOTES

<sup>1</sup> See Kamp (1982) and Heim (1982). Since these works, DRT has developed in many different directions. The version I will adopt employs a dynamic semantics, cf. Groenendijk and Stokhof (1991), and is based on the ideas articulated in Chierchia (1992a, 1994).

<sup>2</sup> The definitions of C-command and M-command are fairly standard:

- (i)  $\alpha$  C-commands  $\beta$  iff the first branching node that dominates  $\alpha$  dominates  $\beta$ .
  - (ii)  $\alpha$  M-commands  $\beta$  iff the first maximal projection that dominates  $\alpha$  dominates  $\beta$ .
- The notion of M-command can be replaced with the notion of being in the same checking domain, along the lines proposed by Chomsky (1992).

<sup>3</sup> See Chierchia (1994, ch. 3) for further details. Notice that this line is not available on the assumption that tripartite structures are as in (18b), for here restriction and scope are symmetrically positioned with respect to the quantificational adverb, and hence there would be no way of distinguishing the well formed (19a) from the ill formed (20a).

<sup>4</sup> This has to do with the so called "proportion problem". The proportion problem affects if-clauses also and requires a modification of the assumption of classical DRT also in this domain. For discussion, cf. Kadmon (1987) and Chierchia (1992a, 1994, ch. 2), among many others.

<sup>5</sup> This operation might have to be articulated further. For example, one would have to spell out exactly the semantics of VP's with an inner subject. I try to do so in Chierchia (1992b) where I regard traces in Spec VP as property-forming operators. Moreover, properties might be "intensional", i.e. they might apply not to ordinary individuals but to NP-intensions. In Chierchia (1992b) I argue that the generic operator creates properties of exactly this kind. Thus, the definition of SI has to be modified so that it can apply also to intensional properties. In the appendix I give an indication of how the definition of SI can be modified to cope with such cases. Further complications may come from the treatment of expletives (as in the subject position of *si* constructions there is a null expletive), which I will ignore for the time being.

<sup>6</sup> In our discussion, we limited ourselves to a consideration of so-called stage level predicates (cf. Carlson, 1977). Since in general the interaction of individual level predicates with genericity is more narrowly constrained than for stage level predicates, one would expect fewer options to be available for *si* constructions with individual level predicates. I think that this expectation is clearly borne out. But to show it properly would take us too far afield. For relevant discussion, see Diesing (1992), Kratzer (1988), and Chierchia (1992b), among many others.

<sup>7</sup> For example, while *si* shows quantificational variability under adverbs of quantification, *pro* does not:

- (a) Se *si* è alti *si* è sempre/di solito/raramente anche belli  
If *si* is tall, *si* is always/usually/rarely also handsome
- (b) se *pro* sono alti, *pro* sono sempre/di solito/raramente anche belli  
If *pro* are tall, *pro* are always/usually/rarely also handsome

The *pro*-subject of sentence (b) cannot be interpreted indefinitely. It must be understood as being about a specific group of people or about a specific kind.

<sup>8</sup> *si* can also antecede long distance reflexives such as *proprio*, to which similar considerations apply, *mutatis mutandis*. Cf.:

- (a) Si ritiene spesso che i propri errori siano piu' giustificati di quelli degli altri.  
Si believes often that SELF mistakes are more justified than those of others  
People<sub>i</sub> often believe that their<sub>i</sub> own mistakes are more justified than those of others.

<sup>9</sup> See e.g. Dowty (1985), Chierchia (1989) and references therein.

<sup>10</sup> Williams's own implementation is somewhat different, but very close in spirit to the one sketched in the text.

<sup>11</sup> In fact, under the assumption that within individual-level predicates either there is no Davidsonian argument (Kratzer, 1988) or that such argument is for independent reasons inaccessible (Chierchia, 1992b), the only possibility for the quantificational adverb *always* in a sentence like (45a) is to bind the variable associated with *si*.

#### REFERENCES

- Bach, E. and Partee, B. H. (1980) Anaphora and Semantic Structure. In K. J. Kreiman and A. Ojeda (eds.), *Papers from the Parasession on Pronouns and Anaphora*, Chicago Linguistics Society, 10.
- Burzio, L. (1986) *Italian Syntax*, Reidel, Dordrecht.
- Carlson, G. (1977) *Reference to Kinds in English*. Ph.D. Dissertation, UMass Amherst, distributed by IULC. Also published by Garland, New York.
- Chierchia, G. (1989) Anaphora and Attitudes *De Se*. In R. Bartsch, J. van Benthem, and P. van Emde Boas (eds.), *Semantics and Contextual Expressions*, Foris, Dordrecht.
- Chierchia, G. (1992a) Anaphora and Dynamic Binding, *Linguistics and Philosophy* 15, 111–183.
- Chierchia, G. (1992b) Individual Level Predicates as Inherent Generics, to appear in G. Carlson and F. J. Pelletier (eds.), *The Generic Book*, University of Chicago Press.
- Chierchia, G. (1994) *Dynamics of Meaning. Anaphora, Presuppositions and the Theory of Grammar*, Chicago University Press, forthcoming.
- Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.
- Chomsky, N. (1992) *A Minimalist Program for Linguistic Theory*, MIT Working Papers in Linguistics, Cambridge, Mass.
- Cinque, G. (1988) On *Si* Constructions and the Theory of *Arb*, *Linguistic Inquiry* 19, 521–582.
- Dekker, P. (1990) *Existential Disclosure*. ITLI Prepublication Series, University of Amsterdam.
- Diesing, M. (1992) *Indefinites*, MIT Press, Cambridge.
- Dowty, D. (1978) Governed Transformations as Lexical Rules in a Montague Grammar, *Linguistic Inquiry* 9, 393–426.
- Dowty, D. (1985) On Some Recent Analyses of Control, *Linguistics and Philosophy* 8, 1–41.
- Dowty, D., R. Wall and S. Peters (1981) *Introduction to Montague Semantics*, Kluwer, Dordrecht.
- Groenendijk, J. and Stokhof, M. (1991) Dynamic Predicate Logic, *Linguistics and Philosophy* 14, 38–100.
- Heim, I. (1982) *The Semantics of Definite and Indefinite NPs*, Ph.D. Dissertation, UMass Amherst. Published in 1989 by Garland, New York.
- Heim, I. (1990) E-Type Pronouns and Donkey Anaphora, *Linguistics and Philosophy* 13, 137–178.
- Higginbotham, J. (1992) Reference and Control. In R. Larson, S. Iatridou, U. Lahiri, and J. Higginbotham (eds.), *Control and Grammar*, Kluwer, Dordrecht.
- Kadmon, N. (1987) *On Unique and Non-unique Reference and Asymmetric Quantification*, Ph.D. Diss., University of Massachusetts at Amherst.
- Kamp, H. (1982) A Theory of Truth and Discourse Representation. In J. Groenendijk, T. Janssen, and M. Stokhof (eds.), *Formal Methods in the Study of Language*, Mathematisch Centrum, Amsterdam.
- Kratzer, A. (1988) Stage Level and Individual Level Predicates. In M. Krifka (ed.), *Genericity in Natural Language*, University of Tuebingen, Technical Report. An extended version to appear in G. Carlson and F. J. Pelletier (eds.), *The Generic Book*, University of Chicago Press.
- Krifka, M. (1988) The Relational Theory of Generics. In M. Krifka (ed.), *Genericity in Natural Language*, University of Tuebingen, Technical Report. An extended version to appear in G. Carlson and F. J. Pelletier (eds.), *The Generic Book*, University of Chicago Press.
- Landman, F. (1989) Groups I, *Linguistics and Philosophy* 12(5), 559–605.
- Parson, T. (1990) *Events in the Semantics of English*, MIT Press, Cambridge, Mass.
- Reinhart, T. and Reuland, E. (1993) Reflexivity, *Linguistic Inquiry* 24, 657–720.
- Schwarzchild, R. (1991) *On the Meaning of Definite Plural NP's*, Ph.D. Dissertation, University of Massachusetts, Amherst.
- Stalnaker, R. (1978) Assertion. In P. Cole (ed.), *Syntax and Semantics* 9, Academic Press, New York.
- Sag, I. (1976) *Deletion and Logical Form*. Ph.D. Dissertation, MIT. Published by Garland, N.Y.
- Suñer, M. (1982) Big PRO and Little Pro, manuscript, Cornell University.
- Williams, E. (1977) Discourse and Logical Form, *Linguistic Inquiry* 8, 101–139.
- Williams, E. (1980) Predication, *Linguistic Inquiry* 11, 203–238.