Italian Middle si: Evidence from non-finite clauses

Leonardo Russo Cardona
University of Siena

1 Introduction

Italian, like other Romance languages, has a system of clitic pronouns. One of those, si, has a special status in that it can be used also for impersonal constructions. In matrix clauses, these structures are found with any kind of verb, but in non-finite clauses they are licensed only with transitive (and unergative) verbs and in a small number of contexts, suggesting there are some syntactic constraints severely limiting their distribution. In this squib, after introducing the basic data about impersonal si constructions (henceforth ISCs), I first critically evaluate some previous analyses of this puzzle (Cinque 1988, Dobrovie-Sorin 1998) through the lenses of an up-to-date framework on the syntax of si (D’Alessandro 2008), discussing the different categorisations and characterisations of si-licensing contexts in non-finite clauses. What emerges from this discussion is that ISCs are found in those non-finite contexts where nominative is assigned. However, there are some exceptions to this generalisation: here I mainly focus on one of them, namely the tough construction (exemplified in 1), which has rarely been considered in analyses of si constructions, and outline some of its interesting properties. This also includes a comparison with similar constructions which accept si in some cases: I present the main data about infinitival relatives under this light. Finally, I make a hypothesis about the special status of si constructions found in this type of structures, based on some common properties observed in these seemingly exceptional contexts, since they pass all the tests that detect middle constructions.

(1) Questo libro è difficile da leggersi.
   this book is tough da to.read=si
   \approx ‘This book is tough to read.’

2 Impersonal si Constructions: Finite vs non-finite contexts

Before introducing the relevant data for a comparison between si constructions in finite and non-finite contexts, it is important to highlight that this has been a very intensely debated topic in the history of Italian syntax, essentially starting from Napoli (1976). The main characteristics that the literature has attributed to impersonal si since then mostly relate to the fact that it is comparable to an arbitrary subject (e.g., Hyams 1986, Manzini 1986) and it is similar to passive constructions.
in some respects (e.g., Belletti 1982). One of the undesirable consequences of there being so many different views has been inconsistency in the names used for the different constructions involving si: so, a clarification is in order. In this squib, I am only concerned with ISCs, defined as those in which the subject is semantically undetermined (following D’Alessandro 2008). Furthermore, I will only focus on transitive verbs, as this is the only verbal class that allows me to compare finite and non-finite clauses (including tough constructions).

2.1 Finite clauses

Transitive verbs in finite clauses show the following two agreement patterns in ISCs:

(2) a. Qui *si* leggono i libri.
   here *si* read.3pl. the books
   ‘Books are read here.’

   b. Qui *si* legge (*i*) libri.
   here *si* read.3sg the books
   ‘What one does here is read books.’

According to D’Alessandro (2008), the slight interpretative difference between (2a) and (2b) is caused by a difference in where si first merges (using the structure in Figure 1):

![Figure 1](assumed-syntactic-structure-for-iscs-adapted-from-dale-rcardona.png)

- In (2a), si is merged as the Specifier of a head encoding telicity (res), which has the internal argument as its Complement. resP then merges with v: thus,

---

1 Unergative verbs are found in non-finite ISCs but not in tough constructions, which rely on the presence of an explicit internal argument.
Italian Middle *si*

*si* blocks accusative assignment to the internal argument (in XP), which can get nominative from T.

- In (2b), *si* is merged as the external argument of the verb in Spec, vP. So, it gets nominative from T and the internal argument is assigned accusative.

In both cases, *si* is considered to be an argument. This type of analysis will be the starting point of this squib.

Similarly to (2a, b), sentences involving adverbs like *facilmente* (‘easily’) also accept both options, with an even clearer interpretive difference.

(3) a. *Questi libri* *si leggono facilmente.*
   *these books* *read.3PL easily*  
   ‘These books read easily.’

   b. *Si legge facilmente (i) libri.*
   *1sg read.3SG easily the books*  
   ‘One easily reads books.’

In the next sections, the role of such lexical items will be studied in relation with some apparently exceptional non-finite contexts where ISCs are found.

2.2 Non-finite clauses

Cinque (1988) offered one of the first extensive accounts of why non-finite ISCs only involve transitive and unergative verbs, and why they seem to be (mostly) limited to Aux-to-Comp and raising structures (as analysed by Rizzi 1982). I report here some examples similar to those used by Cinque (1988):

(4) a. *Essendosi scoperti* *i veri colpevoli,* ...
   *being=si found.MASC.PL the.MASC.PL real.MASC.PL culprits*

   b. *Ritengo essersi scoperti* *i veri colpevoli.*
   *Believe.1SG to.be=si found.MASC.PL the.MASC.PL real.MASC.PL culprits*
   ‘I believe the real culprits to have been found.’

   c. *Sembrano essersi scoperti* *i veri colpevoli.*
   *seem.3PL to.be=si found.MASC.PL the.MASC.PL real.MASC.PL culprits*
   ‘The real culprits seem to have been found.’

All these structures realise a nominative subject, which the past participle of the non-finite verb agrees with (and also *sembrare* ‘to seem’ in 4c). Notably, the options without agreement are ungrammatical; so, it seems that only the pattern in (2a) is available in these non-finite constructions:
Cinque (1988) starts exactly from this asymmetry (which is also responsible for the impossibility of ISCs with other verb classes in these non-finite clauses) to posit two different types of nominative *si* differing in their argumental status: only one of them is compatible with the absence of person agreement, while they are both excluded in control configurations. Conversely, Dobrovie-Sorin (1998) posits the existence of a nominative *si* and an accusative *si*, the former being disallowed in non-finite contexts.

On the other hand, if we analyse this phenomenon with the derivation of ISCs I assumed, in (4a-c) *si* would be assigned accusative (which is the same conclusion reached by Dobrovie-Sorin 1998) since the internal argument takes nominative Case. If we let go of Cinque’s multiple versions of *si* then, why isn’t the other option (the one observed in 2b) viable? There could be either a semantic constraint banning atelic predicates from this kind of clauses or some syntactic reason that makes the derivation crash if *si* is merged in Spec,vP. The latter explanation is probably to be favoured since it is consistent with Dobrovie-Sorin’s (1998) well-grounded generalisation that ‘Nominative clitics must be identified by overt subject agreement morphemes’ (p.415).

The other important question pertains the impossibility of *si* in configurations like (6a, b):

(6) a. *Sarebbe meglio valorizzarsi i giovani.
‘It would be better to appreciate the youth.’

b. *I giovani preferirebbero essersi valorizzati.
‘The youth would prefer to have been appreciated.’
Italian Middle *si*

and the derivation crashes. In (6b), which is also an obligatory control structure, the matrix subject should be coreferential with the internal argument of the infinitive, just like in its equivalent with a passive structure:

(7)  

\[
\text{I giovani preferirebbero essere stati valorizzati.}
\]

\[
\text{the .masc.pl young .masc.pl would .pref .3pl to .be been .masc.pl}
\]

\[
\text{valorizzati.}
\]

\[
\text{appreciated .masc .pl}
\]

'The youth would prefer to have been appreciated.'

It is unclear why this is not allowed with ISCs in configurations like (6b) but is acceptable in other control structures like (8) with no nominative assignment, seemingly exceptional as already noticed by both Cinque (1988) and Dobrovie-Sorin (1998), who proposes that an unspecified 'semantic constraint related to control' (p.421) could be at play:

(8)  

\[
\text{Questo tessuto ha la proprietà di lavarsi * (facilmente).}
\]

\[
\text{this fabric has the property di to .wash = si easily}
\]

'This fabric had the property of washing easily.'

Here, the adverb *facilmente* ('easily') seems to impose a middle reading (defined as a non-agentive, property interpretation)\(^2\) on the *si* construction, making it acceptable even in a control configuration. Then, we are left with two asymmetries in need of an explanation:

\[
\alpha \text{ Between (6b) and (7)};
\]

\[
\beta \text{ Between (6b) and (8)}.
\]

For this reason, Cinque (1988) argues for the existence of three versions of *si* in total: one can never appear in non-finite clauses, one is incompatible with PRO and needs a nominative-assigning head, explaining (\(\alpha\)); and one is found in middle constructions (regardless of PRO), explaining (\(\beta\)). In D’Alessandro’s (2008) terms, these distinctions cannot be accounted for: accepting the fact that *si* can never be assigned nominative in non-finite cases, in all such contexts *si* must then first merge in Spec, res\(^2\), receive accusative and then cliticise on the verb; however, the key

\(^2\) Notably, under this definition not all ISCs patterning like (2a) are necessarily middle constructions. For instance, a progressive construction would be acceptable (which is incompatible with an obligatory property reading, see Keyser & Roeper 1984: 385):

(1)  

\[
\text{Qui si stanno leggendo dei libri. here si are reading some books}
\]

\[
\text{'Some books are being read here.'}
\]
to resolving this puzzle syntactically could be found in the movements to the TP domain undergone by *si* after it is first merged.

Another direction which might be worth exploring could be focusing on what makes ISCs acceptable in sentences like (8) by studying other similar cases, assuming that normally *si* is not compatible with lack of nominative assignment to a DP: this is the topic of the next section.

3 TOUGH CONSTRUCTIONS

Constructions like (9) accept impersonal *si*, even if they disallow nominative, similarly to (8):

(9) *Questo libro è difficile da leggere /leggersi.*

this book is tough to.read to.read=si

≈’This book is tough to read.’

To my knowledge, no study about *si* constructions has considered these structures since Belletti (1982). According to her derivation, the embedded infinitival is reanalysed along with the *tough* item as a complex adjective and does not involve control PRO but an empty NP governed by *si*. In a minimalist perspective, some elements of this account can still be useful (e.g. the idea of *si* ‘absorbing’ accusative Case and the external theta-role) while others are no longer tenable. Here, I am particularly interested in noting the similarities and differences with regard to (8) and assessing the role of the *tough* items in licensing a special kind of ISC.

3.1 ISCs in tough constructions?

Firstly, it should be mentioned that there seems to be some degree of optionality in (9): in both cases the external theta-role of the infinitival is taken up by an arbitrary subject and the internal argument appears to have been promoted to (or at least be coindexed with) the matrix subject position. Conversely, in constructions like (8), the absence of *si* would cause an active reading of the infinitival, as shown by (10), which does not require a *tough* adverb:

(10) *Questo sapone ha la proprietà di uccidere tutti i batteri.*

this soap has the property di to.kill all the bacteria

‘This soap has the property of killing all bacteria.’

However, both (9) and (8) rely on the presence of a *tough* item: in one case this is realised as an adjective introducing an infinitival (*difficile* ‘difficult’ in 9), in the other case as an adverb modifying the verb (*facilmente* ‘easily’ in 8). So, these items might have a functional status in licensing a middle *si* construction.

A preliminary hypothesis could be that tough adverbs like *facilmente* (‘easily’) are hosted by some specific functional projection in the TP à la Cinque (1999), present in both (8) within the infinitival clause and in finite clauses like (3a). For instance, this
would be consistent with Roberts’s (1987) generalisation about middles: ‘Middles must be governed by an “appropriate” Infl or a modifier’ (p.233). Also Benincà & Tortora (2009) propose that some specific phrases in the functional domain interact with si and determine its different uses as well as the asymmetries between finite and non-finite clauses. The adverbs I have highlighted could thus serve to impose a property reading and in turn license an interpretation of the verb as having an implicit, backgrounded agent, arbitrary in reference. What is the status of the adjective in tough constructions, then?

First, it can be tested whether tough constructions pattern in the same way as (8) in terms of having a mandatory middle reading. I will adapt the tests Cinque (1988: 562–563) applies to constructions like (8), involving agentive modifiers, purpose clauses and adjunct small clauses, also drawing on Fellbaum & Zribi-Hertz (1989) and Keyser & Roeper (1984):

(11) a. Questo vetro è facile da danneggiarsi (??involontariamente).
    this glass is easy da to.damage=si unwillingly
    ≈’This glass is easy to unwillingly damage.’

b. Questi politici sono difficili da corrompersi (*per far vincere 
       these politicians are tough da to.bribe=si to make to.win 
       il proprio partito).
       the own party
    ≈’These politicians are tough to bribe to make one’s own party win.’

c. Questo problema è difficile da affrontarsi (*anche uniti).
    this problem is tough da to.tackle=si even united.masc.pl
    ≈’This problem is tough to tackle even united.’

So, tough constructions are subject to the same constraints as (8): this suggests they are indeed an instance of a syntactic middle construction and that the tough items are necessary for it to be licensed. Moreover, the role of si appears to be fundamental in this respect, as the options equivalent to (11a–c) but with no si are acceptable:

(12) a. Questo vetro è facile da danneggiare (involontariamente).
    this glass is easy da to.damage unwillingly
    ‘This glass is easy to unwillingly damage.’

b. Questi politici sono difficili da corrompere (per far vincere il
    these politicians are tough da to.bribe to make to.win the
    proprio partito).
    own party
    ‘These politicians are tough to bribe to make one’s own party win.’
c. Questo problema è difficile da affrontare (anche uniti).
this problem is tough DA to.tackle even united.masc.pl
‘This problem is tough to tackle even united.’

Therefore, there is no true optionality between the options with si and the ones without: only the former have a compulsory non-agentive reading imposed by si in combination with a tough item. However, some aspects of middle constructions are also shared by tough constructions with no si since they nevertheless describe a property rather than an event, as shown by the fact they do not admit progressive constructions (Keyser & Roeper 1984: 385):

(13) * Questo libro è facile da stare/starsi leggendo.
this book is easy DA to.be/to.be=si reading
‘This problem is tough to tackle even united.’

In my analysis of ISCs in non-finite clauses so far, I have identified two contexts where si is licensed despite the unavailability of nominative assignment to a DP. The tests I applied, mostly following Cinque (1988), showed that these contexts are similar in that they have the same constraints with regard to agentive adverbs, purpose clauses and adjunct small clauses. These facts are consistent with the definition of middle constructions as forcing a non-agentive (the agent theta-role is not realised), property reading (non-eventive, stative interpretation). However, ISCs in finite clauses can sometimes differ in these properties:

(14) Questi libri si usano per imparare l’inglese.
these books si use.3pl per to.learn the=English
‘These books are used to learn English.’

Thus, some other properties of the structure (e.g. lack of a nominative-assigning head) could add some further constraints to the agentivity of ISCs (‘subject suppression’, see Hoekstra & Roberts 1993) found in tough constructions, by making the external argument unable to be a controller.

The mandatory property reading appears to be found also in tough constructions with no si and is probably linked to the properties of the tough adjectives selecting such infinitival clauses, as well. Therefore, it is the combination of these two factors (tough adjectives and presence of si) that yields a middle reading.

3.2 Syntactic peculiarities of tough constructions

The possible presence of si is not the only interesting aspect of tough constructions. In fact, they have been a heatedly debated topics cross-linguistically due to their peculiar syntactic properties at least since Lasnik & Fiengo (1974) and Chomsky (1981). Both in English and in Italian, these structures establish a dependency between the matrix subject and the infinitival internal argument (Cinque 1996, Roberts 1997):
Italian Middle *si*

(15)  a. [Questo libro]_i_ è difficile da leggere/leggersi t_

    b. [This book]_i_ is tough to read t_

This is evident from their equivalence with, respectively:

(16)  a. È difficile leggere questo libro.

    b. It is tough to read this book.

The extraction of the object from the infinitival clause and its promotion to the matrix subject position has been the object of discussion in modern generative syntax because it appears to be an instance of A-A’-A movement (see Hicks 2009, Boskovic 2021). However, the issue mainly stems from the fact that in English the dependency may in principle hold across multiple embedded clauses. This is not true in Italian:

(17)  a. [This book]_i_ is tough to convince a kid to read t_

    b. *[Questo libro]_i_ è difficile da convincere un bambino a leggere t_

This difference between Italian and English could lie in different properties of the CP layer of the *tough* infinitivals and/or different movement patterns: consequently, the analyses that have been offered so far (e.g., Hicks 2009) to account for English *tough* constructions might not be fully applicable to Italian. To shed more light on their syntax, then, it might be worth studying some similar constructions involving *da* in Italian, like infinitival relatives (Cinque 2020), as shown in the next subsection.

3.3 Comparison with infinitival relatives

Italian infinitival relatives allow *da* + *si* in some cases, and seem to overlap with tough constructions in some respects:

(18)  a. *Questo libro* è *da* leggere /leggersi.

    this book is *da* to.read to.read=*si.

    ‘This book has to be read.’

    b. *Questo è un libro* *da* leggere /leggersi.

    this is a book *da* to.read to.read=*si

    ‘This is a book to read.’

    c. *Ho trovato un libro* *da* leggere /??leggersi.

    have.1sg found a book *da* to.read to.read=*si

    ‘I found a book to read.’

In (18a), there is no overt adjective introducing the infinitival clause, which has a mandatory deontic interpretation, while in (18b, c) this reading alternates with a possibility interpretation (Cinque 2020: 196–198). Let us assume that the deontic
meaning is conveyed by a null adjective, similarly to Kayne’s (2014) proposal for English ‘IS TO’ constructions: this adjective, in the case of Italian, would have the same syntactic properties as those belonging to the tough class as it supports a da + si option. In (18c), which is ambiguous between a deontic and possibility interpretation, da leggersi (‘to read’) may only be grammatical to the extent that it is interpreted non-restrictively and deontically, as some reduction of the ‘IS TO’ construction, even if this is not the preferred reading. Cinque (2020: 191–198) noticed that a non-restrictive reading of the relative clause necessarily implies a deontic reading (both conditions seem necessary to license si). Constructions like (18a) are also similar to tough constructions since their matrix subject is non-thematic, as shown by the fact that they have the same interpretation as (19):

(19) Bisogna leggere questo libro.
It is necessary to read this book

‘It is necessary to read this book.’

These similarities suggest that it would be useful to integrate accounts of tough constructions and infinitival relatives in Italian to shed some light on their puzzling syntax. In particular, ISCs could play a diagnostic role to this end: their acceptability in relative clauses signals a deontic interpretation, as opposed to a possibility reading that seems to be strongly preferred in sentences like (18c), which express a referential subject and hardly accept a non-restrictive interpretation. The reason why this is incompatible with si could be that, as I have proposed for tough construction, the only type of ISC that can be found in infinitival relatives is a middle construction.

3.4 Middle si constructions and PRO

Going back to the exceptional status of tough constructions with regard to the presence of si in an environment where nominative cannot be assigned, infinitival relatives appear to share this characteristic as well. Normally, these structures are assumed to contain a PRO in subject position, obligatorily controlled by the implicit Experiencer of the tough predicate (Hoekstra & Roberts 1993: 186).

According to the analysis offered by Kayne (2014), Romance languages like Italian can also accept an overt Experiencer in these constructions, which would thus control PRO in sentences like (20). Remarkably, this is not possible when si is there:

(20) Questo libro è difficile per Mario da leggere /*leggersi.
this book is tough for Mario to.read to.read=si

‘This book is tough for Mario to read.’

Hence, si seems to force a quasi-universal interpretation on the infinitival subject, and to be in complementary distribution with PRO (which is consistent with Belletti’s 1982 claims). This could also explain why the da + si option in (18c), which is very marginally grammatical, only accepts a non-restrictive reading: in a restrictive
interpretation, the matrix subject (which has a specific reference in this case) should control the subject of the infinitival clause. This is not possible with *si* there.

It is important to notice that, conversely, *si* in sentences (4a-c) can be interpreted quasi-existentially, as it is ‘compatible with specific time reference’ and ‘compatible with the existence of a single individual satisfying the description’ (Cinque 1988: 546). For instance, the following sentences are all acceptable (using one of tests found in Cinque 1988 and in Hoekstra & Roberts 1993):

(21) a. Essendosi scoperti i veri colpevoli
being=si found.masc.pl the.masc.pl real.masc.pl culprits

  *stasera, …

  ‘The real culprits having been found tonight, …’

b. Ritengo essersi scoperti i veri colpevoli
Believe.1sg to.be=si found.masc.pl the.masc.pl real.masc.pl culprits

  *stasera.

  ‘I believe the real culprits to have been found tonight.’

c. Sembrano essersi scoperti i veri colpevoli
seem.3pl to.be=si found.masc.pl the.masc.pl real.masc.pl culprits

  *stasera.

  ‘The real culprits seem to have been found tonight.’

Using Cinque’s (1988) categorisation, then, the obligatory quasi-universal reading would be a property of ‘middle’ *si* as it is shared by ISCs in tough constructions, infinitival relatives and sentences like (8): all of them are ‘incompatible with specific time reference’ and ‘incompatible with the existence of a single individual satisfying the description’ (p.546). For example, repeating the test for (22a-c), these sentences are not grammatical (keeping in mind that *stasera* (‘tonight’) could be more acceptable if interpreted as modifying the matrix clause):

(22) a. Questo tessuto ha la proprietà di lavarsi facilmente (*stasera).
this fabric has the property di to.wash=st easily tonight

  ‘This fabric had the property of washing easily tonight.’

b. Questo libro è difficile da leggersi (*stasera).
this book is tough da to.read=st tonight

  ≈‘This book is tough to read tonight.’

c. Questo libro è da leggersi (*stasera).
this book is da to.read=st tonight

  ‘This book has to be read tonight.’
Rather than positing the existence of multiple versions of *si* however, I would like to highlight that the presence of *tough* items seems to correlate with the cluster of properties that I have identified for these *si* constructions:

- Non-agentive, property interpretation (middle);
- Quasi-universal reading of *si*.

Moreover, (22b, c) show a dependency between the matrix subject and the infinitival object.

Therefore, this type of middle *si* constructions would be a subset of ISCs, assuming they involve the same syntactic derivation within the vP layer as (2a) and (4a-c), with some additional constraints. The *tough* items then appear to play a fundamental role in licensing an ISC when nominative is not available, and also to impose these constraints on the syntax and semantics of the infinitivals they introduce.

4 Conclusion

In this squib, I first reviewed some classical accounts of ISCs in non-finite clauses, attempting to critically re-evaluate them and the evidence they employed through a minimalist perspective by suggesting some directions to apply D’Alessandro’s (2008) framework to non-finite clauses. Initially, I focused on non-finite contexts where nominative can be assigned to an overt subject, as in Aux-to-Comp and raising structures: here, ISCs can be regularly found, with *si* receiving accusative Case (as nominative clitics are banned from non-finite clauses). To fully understand the relationship between the availability of nominative and presence of impersonal *si* I suggested that the properties of the infinitival functional spine of such constructions should be studied more in depth. This is especially important since in some cases ISCs of a special kind are licensed in contexts with no nominative assignment: in particular, I devoted most of my attention to *tough* constructions and showed that the presence of *si*, which in this case has an obligatory quasi-universal reading, imposes a non-agentive interpretation. This, combined with the fact that these structures already accept only a property reading, causes a middle interpretation.

The situation is quite similar for structures like (8) and for infinitival relatives that seem to employ a null adjective with a deontic meaning and which has roughly the same behaviour as *tough* adjectives. In general, *tough* items seem to provide some syntactic grounds for ISCs to appear even in non-finite clauses lacking nominative, but require additional constraints in the interpretation and in the syntax of such constructions. In these cases, *si* is in complementary distribution with PRO.

In conclusion, I have provided abundant data about how different kinds of ISCs are licensed in non-finite clauses and have outlined several properties pointing to the existence of middle constructions within ISCs. This has shown why the study of *si* could be useful to understand their peculiarities and could serve as a background for future research on the issues I have touched upon and to update existing theories.
References


Leonardo Russo Cardona
The University of Siena
l.russocardona@student.unisi.it