

Chiara Dal Farra¹, Silvia Silleresi¹, Artemis Alexiadou², Maria Teresa Guasti¹ & Uli Sauerland³

¹ Università degli Studi di Milano-Bicocca

² Humboldt-Universität zu Berlin

³ Leibniz-Centre General Linguistics (ZAS)

The uses of clitic *si* in Child Italian: a corpus study

Introduction The exact role and interpretation of the clitic *se/si* is heavily debated in Romance languages, particularly in Italian (Belletti 1982; Manzini 1986; Cinque 1988; D’Alessandro 2007; Pescarini 2015, a.o.). Italian *si* appears in a variety of contexts, listed with examples in Table 1. Despite extensive theoretical research, (a) there is no consensus as to whether there is a single lexical item *si* or two different ones in the grammar of Italian, and (b) there are no detailed acquisitional studies dedicated to *si* in Italian. Studies on some other Romance languages suggest that there is a developmental path in the acquisition of *si*’s different roles: the reflexive/anticausative functions seem to be acquired first (see Teomiro & Escobar 2013 for Spanish; Barrière et al. 2000 for French). However, the use of impersonal *si* is often not considered in this literature. The current study contributes to the theoretical debate through the lens of acquisition by addressing three questions: (i) do children use *si* adult-like?, (ii) given the several different functions of *si*, is there a difference in their development in child Italian?, and (iii) can children data help us decide between competing theories of *si*?

Method A corpora analysis in the CHILDES database (MacWhinney 2000) was run through an examination of all productions of *si* followed by a verb (*si+V*) by 16 typically-developing Italian-speaking children aged 1;8 to 3;4. We annotated all *si+V* occurrences ($n = 283$) according to the following criteria: (a) identification of *si*-type constructions (following the annotation in Table 1), (b) appropriate versus inappropriate uses of *si* (grammatical vs ungrammatical constructions in the adult grammar), (c) verb type used (transitive, unaccusative, unergative, alternating). As for the criteria in (a), plural verbal agreement was not available as a distinguishing factor between impersonals and passivizing *si* (Pescarini, 2015), as in most of the utterances the verb appeared in its singular form.

We will integrate this first analysis with more specific data regarding these occurrences of *si*, namely (i) preverbal or postverbal realization of DP (when present), (ii) cases of DP drop, and (iii) reference of *si* to an animate or inanimate entity. Moreover, we will consider (i) instances of 1st and 2nd person reflexive clitics (to check whether the low number of reflexives was due to a corpus effect), and (ii) cases of omission of *si* in anticausative verbs which mark the alternation. Finally, the result will be compared to corpora analysis of child-directed speech in the CHILDES database.

Results Our preliminary results show that children aged 1;8-3;4 use *si*-type constructions productively and adult-like. Figure 1 reports the rate of production of the four main *si*-type constructions attested in the dataset: impersonal (54%), anticausative (29.3%), true reflexive (6.3%) and inherent reflexive (5.4%). Concerning the other possible functions of *si*, we found no instances of passive or middle *si*. Statistical analysis revealed a significant difference in the mean rate of production of impersonal *si* ($p < .001$) w.r.t. the other attested *si* constructions. Among the impersonal *si* constructions ($n = 153$), 56% have a generic interpretation ($n = 86$) and 15.6% a speaker-inclusive ($n = 24$) (the remaining 28% was ambiguous between the two readings). An effect of verb type emerged on impersonal *si*, which was significantly ($p < .001$) more likely to occur with transitive verbs than with all other verb types (Figure 2). In particular, almost the totality of impersonal generics occurred with transitive verbs (93%). In addition, as expected, anticausative *si* occurred exclusively with alternating verbs, inherent reflexive *si* only with unaccusatives and true reflexive *si* only with transitives.

Analysis Three main results emerged from our analysis: (i) four types of *si* constructions (impersonal - generic and existential -, anticausative, true reflexive, inherent reflexive) were attested and produced grammatically by Italian children from very early on; (ii) passive and middle *si* constructions were not produced; (iii) impersonal *si*, and in particular generic impersonal *si*, was by far the most frequent production in the dataset. Focusing on (iii), we note that in our data *si* was realized (almost) exclusively with transitive verbs, even though they can occur with unergatives and unaccusatives in adult language. From the preliminary results we obtained, it is not entirely clear whether our data help decide between two competing theories of *si*: according to theory 1, there is a single lexical item *si* across different contexts (Manzini 1986); according to theory 2, there are different types of *si*'s. Notably, impersonal *si* is different from the other *si* constructions (d'Alessandro 2007; Pescarini 2015 a.o.). While the production of impersonals, reflexives and anticausatives seem in line with theory 1, it is not clear why middle and passive *si* are not produced. We propose that impersonal *si* is analyzed by children as an impersonal pronoun, a pronominal element in Spec, VoiceP, saturating the external argument variable (see Schäfer 2017). We suggest that impersonals delay the production of passives and middles, since they seem to be the more efficient structure to express a passive interpretation, which would be more complex, given (i) additional layer of structure above VoiceP (see Bruening 2013; Alexiadou et al. 2015), and (ii) the requirement of movement (see smuggling approaches in Collins 2005; Belletti 2019). This explains the transitivity restriction we identified in our data and receives cross-linguistic support from languages that express passive structures via impersonals (see Keenan & Dryer 2007; Koopman 2021). Finally, the fact that impersonal *si* were produced both with animate and inanimate references, suggests that they cannot be interpreted as reflexives.

Table 1. Italian *si*: functions and examples

<p>True reflexive (1) Maria <i>si</i> critica. Maria <i>si</i> criticizes.</p>	<p>Inherent reflexive (2) Gianni <i>si</i> addormenta. Gianni <i>si</i> falls asleep.</p>	<p>Anti-causative (3) Il vaso <i>si</i> è rotto. The vase <i>si</i> is broken.</p>
<p>Middle (4) Quel libro <i>si</i> legge facilmente. That book <i>si</i> reads easily.</p>	<p>Passive (5) <i>Si</i> sono vendute delle auto. <i>Si</i> was sold some cars.</p>	<p>Impersonal generic (6a) <i>Li</i> <i>si</i> spende molto. There <i>si</i> spend a lot.</p> <p>Impersonal speaker inclusive (6b) Domani <i>si</i> va al mare. Tomorrow <i>si</i> go to the sea.</p>

Figure 1. Proportion of production per *si*-types

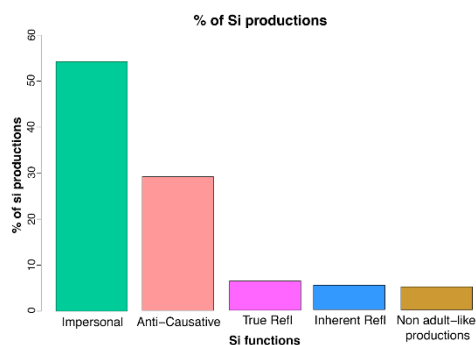
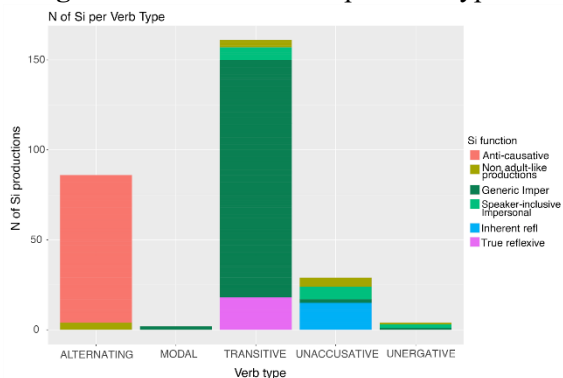


Figure 2. Production of *si* per verb-type



(Selected) References

Alexiadou, A., Anagnostopoulou, E., & Schäfer, F. (2015). *External arguments in transitivity alternations: A layering approach*. Belletti, A. (1982). "Morphological" passive and pro-drop:

the impersonal construction in Italian. *Journal of Linguistic Research* 2: 1-34. **Bruening**, B. (2013). By Phrases in Passives and Nominals. *Syntax* 16: 1-41. **Cinque**, G. (1988). On Si Constructions and the Theory of arb. *Linguistic Inquiry* 19: 521-582. **Manzini**, M.R. (1986). On Italian si. In *The Syntax of Pronominal Clitics*, 241-262. **Pescarini** (2015). *Le costruzioni con si. Italiano, dialetti e lingue romanze*. Roma: Carocci Editore. **Schäfer**, F. (2017). Romance and Greek medio-passives and the typology of Voice. *The verbal domain*, 129-152.