Overt Subject Distribution in Early Italian Children.

Paolo Lorusso [1], Claudia Caprin [2] & Maria Teresa Guasti [2]
[1] Universitat Autònoma de Barcelona & [2] Università di Milano-Bicocca

1. Introduction

It is a well-known fact that children drop subjects from their early sentences. The phenomenon has been well-documented in many languages where this is not permitted in the adult grammar like English, Dutch, German, French (Gerken, 1991; 1996; Radford, 1990; Rizzi, 2000; Bloom, 1990, 1993; Hamann & Plunkett, 1998), but rarely investigated in null subject languages (but see Serratrice, 2002, Serratrice e Sorace, 2003), like Italian, Catalan and Portuguese, except for comparative purposes (Bates, 1976; Hyams, 1986; Valian, 1991; Pizzuto e Caselli, 1992). Cross-linguistic research has shown a great influence of the target language on the early subject omissions: children learning to speak a non-null subject language drop the sentential subject less than children learning to speak a null-subject language. In fact American learners drop subjects less than their Italian peers. Valian (1991) comparing the general percentage of null subjects in a *non-pro-drop* language, i.e., English, and in a *pro-drop* language, i.e., Italian, found that subjects are omitted at different rates by learners of the two languages: American English learners show a lower percentage of omission (31% in sentences with MLU < 2.0 and 11% in sentences with MLU 2.0-3.0) than Italian learners (almost 70 %) (see also Valian & Eisenberg, 1996 for a comparison between English and Brazilian Portuguese). These data were used as a reply to Hyams' (1986) proposal according to which English learners misset the *pro-drop* parameter assigning a positive value to it. As Valian argued, the different percentages of omission suggest that subject omission in early English and Italian are two distinct phenomena, a line fully developed by Rizzi (1993/4) and supported by the fact that the distribution of early null subjects in Italian and English is different. For example, Italian learners frequently omit subject in finite interrogative clauses (see Guasti, 1996), while English learners do not. Rizzi (1993/4) proposed that early null subjects in early non-pro-drop languages are a root phenomenon, limited to main declarative clauses and governed by a new parameter, the root null subject parameter (Rizzi, 2000). This parameter operates in adult languages and is responsible for subject omission in languages like Levantin Arabic, Brazilian Portuguese (where null subjects are allowed only from main declarative clauses) and is employed by children under processing pressure.

Influence of the target language was also confirmed by longitudinal studies with bilingual children who were simultaneously acquiring an overt and a null-subject language. Both Serratrice (2002), who studied an English-Italian bilingual child, and Juan-Garau & Perez-Vidal (2000), who studied an English-Catalan bilingual child, found a great discrepancy in subject dropping depending on the language.

Despite these results, a detailed study about the way children learning null subject languages use the sentential subject has not been conducted yet. This study aims at starting to fill this lacuna by looking at early Italian. In particular, we analyze the realization and the distribution of subjects with different classes of verbs (transitive, unergative and unaccusative verbs) and show that children use subjects in different ways depending on the verb class. These findings have two consequences. On the one hand, they show that children distinguish these different classes of verbs and are aware of their different argument structure; on the other, we argue, they shed new light into the debate concerning the acquisition of passives, providing clear evidence that children distinguish between unaccusative and unergative verbs.

2. Methods

2.1. Linguistic Corpora

Our study is based on the analysis of two corpora of early Italian: one collected through a longitudinal method and the other through a cross-sectional method.

a. Longitudinal corpus

The longitudinal corpus consisted of the productions of four Italian children aged between 18 and 36 months (Calambrone corpus, Diana Martina, Raffaello, Rosa CHILDES database MacWhinney & Snow 1985; Cipriani et al., 1989). The child corpora consisted in 17573 utterances of which 4733 include a verb; of these we examined only declarative sentences for a total of 2838 utterances. We also examined the speech of adults

on a subset of the files that were used in the analysis of children speech. In this work, 2095 adult declarative sentences with verbal constructions were examined (Lorusso, 2003).

b. *Cross Sectional corpus*

The participants to the cross-sectional study were 59 children (25 males and 34 females), attending a public day care center, ranging in age from 22 to 35 months (Mean age=28.92 sd=2.8). Each child was videotaped and audiotaped in a natural conversational setting while playing with a caregiver at her nursery school. The toys used were the same for all the children. Each recording session lasted about 20-25 minutes. Each tape was transcribed by one listener and checked by another. All the doubtful portions were listened again, and when the listeners did not reach an agreement, the portions were labelled as unintelligible. Conventional Italian orthography was used. CHAT format of CHILDES coding system was adopted for all the transcriptions (MacWhinney 1997). In order to compare the linguistic development of children, they were grouped on the basis of their MLU computed in words (Bates, 1976; Pizzuto & Caselli, 1992). So we considered three groups based on MLU-w values: from 1.0 to 1.5 (G1 N=15); from 1.5 to 2.0 (G2 N=19); from 2.0 to 3.1 (G3 N=25). The corpus consisted in 10621 utterances 2577 of which included a verb and entered the analysis. We also analyzed the speech of 9 adults during their interaction with children (Caprin, 2003).

2.2. *Analysis*

For each child we examined the use-omission of subject as a function of the syntactic frame of the verb: transitive, unaccusativeand inergative verbs. When the subject was present, we examined whether it occurred in preverbal or postverbal position. We separately considered whether the verb was in a simple or compound tense and, in the last case, we kept trace of the auxiliary that was used. We distinguished among declarative, imperative and interrogative sentences, but in the analysis we retained only declarative sentences.[1] The same analysis was performed on the speech of the caregivers during their interaction with children. The same analysis was performed on the cross-sectional data.

3. Results

The results we obtained from the analysis of the two corpora were quite similar. We report them separately, starting from the results of the longitudinal study. Table 1 summarizes the percentages of overt and null subjects in the overall production of the four children and of the adults interacting with them.

Tab.1 Percentage of overt/null subjects for children and adults in the longitudinal corpus.

	Overt Subjects	*Null Subjects*
Diana	27%	73%
Martina	33%	67%
Raffaello	21%	79%
Rosa	23%	77%
Total of Children	**25%**	**75%**
Total of Adults	**26%**	**74%**

[1] Interrogative sentences, especially with first and second person verbs, generally lack an overt subject. In addition, the subject of interrogative sentences cannot lay in the preverbal position. Because of these specific structural properties, interrogatives were not analysed. Similarly, we did not analyze imperative clauses, since the subject in these clauses is absent in non-pro-drop languages as well.

We can observe that there is a strong similarity between children and adult productions. If we examine the percentage of overt subjects as a function of the verb they occur with (tab.2), we notice a difference between adults and children which is not evident in the overall results reported in table 1.

Tab.2 General Percentage of overt subjects distribution for each children and for adults depending on the class of the verb in the longitudinal corpus

Overt Subject Distribution along Verb Classes (Percentage)			
	Unaccusatives	Unergatives	Transitives
Diana	36%	22%	26%
Martina	41%	25%	32%
Raffaello	37%	23%	15%
Rosa	32%	30%	19%
Children	**36%**	**25%**	**22%**
Adults	**41%**	**40%**	**20%**

We can notice that all the children produce more overt subjects with unaccusatives than with other verb classes. Adults differ, in that they use overt subject more or less at the same rate with unergatives and unaccusatives. This different behaviour of subjects with unaccusative verbs with respect to the other two classes of verbs is also evident in the distribution of overt subjects in table 3.

Tab.3 Distribution of overt subjects with the three classes of verbs: unaccusatives, intransitives, transitives in the speech of children and adults in the longitudinal corpus

Overt Subject Position along Verb Classes (Percentage)						
	Unaccusatives		Unergatives		Transitives	
	Preverbal Position	***Postverbal Position***	***Preverbal Position***	***Postverbal Position***	***Preverbal Position***	***Postverbal Position***
Diana	31%	69%	60%	40%	73%	27%
Martina	32%	68%	80%	20%	75%	25%
Raffaello	35%	65%	95%	5%	67%	33%
Rosa	37%	63%	73%	27%	74%	26%
Children	**34%**	**66%**	**79%**	**21%**	**72%**	**28%**
Adults	**43%**	**57%**	**83%**	**17%**	**63%**	**37%**

We observe a significant difference in the distribution of subjects as a function of the verb type in child speech (X^2= 37.39 p<.005). With unaccusative verbs, subjects tend to occur in postverbal position, with transitive and unergative verbs they tend to be placed in preverbal position. The same distribution is observed in adult's speech. In summary, subjects are omitted less with unaccusative verbs than with the other two classes of verbs and overt subjects of unaccusative verbs occur more frequently in postverbal position.

Let us now turn to the results of the cross-sectional study. First, we observe that the use of overt subject differs among the three groups (Kruskal-Wallis χ^2 = 9.750 df=2 p=0.008) and increases with MLU-w (Jonckeheere J-T=2.851 p=0.004).

Second, the comparison of overt subject use between transitive verbs followed or not by complements (licit and illicit omissions) showed no difference. This suggests that whether the verb has a complement or not does not influence the omission of subjects, contrary to what was found for English (Bloom 1990, 1993; Valian 1991).

Third, overt subject use was greater with unaccusative than with unergative verbs (Wilcoxon z=-2.571 p=0.01), a phenomenon that is also observed in the input (Wilcoxon z=-2.429 p=0.01).

Tab. 4 Overt subjects distribution for each children group and for adults in the cross-sectional corpus. The first value is the number of subjects who produced the contexts, the second value is the percentage of overt subjects.

Overt Subject Distribution (Percentage)		
	Number	Percentage
G1	14	9.3%
G2	18	26.5%
G3	25	24.5%
Children	**57**	**21.5%**
Adults	**9**	**36.3%**

Tab. 5 Overt subjects distribution for each verb class in the cross-sectional corpus. The first value is the number of subjects who produced the contexts, the second value is the percentage of overt subjects.

Overt Subject Distribution (Percentage)		
	Number	Percentage
Lexical verbs	53	15.3%
Transitives with complement	50	16.2%
Transitives with licit omission of complement	45	12.5%
Transitives with illicit omission of complement	36	11.5%
Unaccusatives	39	28.7%
Unergatives	39	12%

The same trend is observed when we look at the distribution of overt subjects. The order SV was preferred with transitive verbs (Binomial p<.001) and with unergative verbs (Binomial p=.035), while overt subject in utterances with unaccusative verbs occupy either pre-verbal (SV) or post-verbal position (VS).

Tab.6 Distribution of overt subjects with the three classes of verbs: unaccusatives, intransitives, transitives in the speech of children in the cross-sectional corpus

Overt Subject Position along Verb Classes (Percentage)		
	Preverbal Position	Postverbal Position
Transitives	73.62%	26.38%
Unaccusatives	46.29%	53.71%
Unergatives	73.91%	26.09%

In summary, we observe that there is an increase in subject use as a function of the MLU and, more interestingly, that the pattern of expression of overt subjects is influenced by the class of the verb. More omissions are observed with transitive and unergative verbs than with unaccusative verbs. This discrepancy is also observed in the distribution of overt subjects. Both studies have shown that transitive and unergative verbs tend to have subjects in preverbal position, while unaccusative verbs tend either to have it preferentially in postverbal position or equally frequently in preverbal or postverbal position. There are similarities between the expression of subject between children and adults, but there are also differences. For example, in the

longitudinal study, adults tend to use more overt subjects with intransitive verbs (unaccusatives and unergatives) than with transitive ones.

3. Discussion

The main result of this study is that children use subjects in different ways depending on the verb classes: they use more subjects with unaccusative verbs and they do not have a preference for the preverbal position with the same class of verbs. In so doing, their production of overt subjects seems to be influenced by an ergative pattern in that they treat subjects of unaccusatives differently from subjects of unergatives and transitives. This pattern is to be linked to the informational and syntactic characteristics of the arguments of unaccusative verbs. While unergative and transitive verbs project their subject in a position external to the VP, unaccusative verbs project their subject in a position internal to the VP (possibly adjoined to the VP, as in Koopman and Sportiche, 1991). This difference is reflected in our data in that subjects of unaccusative verbs remain in the postverbal position more frequently than subjects of unergative and transitive verbs. Assuming that verbs move from V to I, we argue that subjects of unaccusative verbs can remain in the internal argument position, because they can receive partitive case there, as proposed by Belletti (1988). Subjects of unergative and transitive verbs, instead, cannot lay in the base generated position located in the VP area, because they can't be assigned case there. Then, they must move to the preverbal position to get nominative case. The different case properties associated to unaccusative versus unergative and transitive verbs account for the different distributional pattern of overt subjects. We observed that subjects of unaccusative verbs can occur in the preverbal position. This is obtained by having them to move from the internal position to the specifier of IP, where all subjects move, a movement that occurs also in other constructions, such as passives. In this respect, subjects of unaccusative verbs behave as subjects of unergative and transitive verbs. The crucial difference is that they have an additional option, that of staying in their base generated position. Similar data are reported for other *pro drop* languages. For example, Catalan learners display a similar pattern of distribution of overt subjects. Cabré Sans (2004) analyses the productions of two Catalan monolinguals aged between 14 and 37 months (Serra – Solé: Pep, Alvar. CHILDES database MacWhinney & Snow 1985). She found that overt subjects of unaccusative verbs are expressed in post-verbal position in 65% of the cases, while with unergatives this happens in 29% of the cases and with transitive verbs only in 22% of cases.

The specific properties of the internal argument of unaccusative verbs are also responsible for the pattern of subject omission. It is generally assumed that the omitted subject is a null pronoun located in the specifier of IP where it is licensed and identified by agreement features. Subjects of unaccusative verbs can be null only if they move to Spec IP. But since they also have the option of remaining in the internal position, there will be less opportunities for omission compared to subjects of unergative and transitive verbs, which need to move to Spec IP. While omission of subjects is a frequently observed phenomenon in the earliest stages of language acquisition, even in non-null subject languages, omission of internal arguments is a rare phenomenon. The fact that subject omission with unaccusative verbs is less frequent than subject omission with unergative and transitive verbs indicates that children are aware of the peculiar grammatical status of the subject of unaccusative verbs: omitting it, they treat it as a subject, but in doing this in a limited way, they are treating it as an internal argument. Our findings are reminiscent of the distributional pattern of arguments in the home signed languages described by Zheng and Goldin-Meadow (2002). These authors found American English and Chinese deaf children not exposed to a conventional language omit arguments according to an ergative pattern in that they omit the subjects of transitive sentences, but much less the subject of unaccusative sentences and the object of transitive sentences. They also found that the same pattern is present in the speech of Chinese hearing children, in spite of the fact that Chinese is not an ergative language. On this basis, they suggest that, when not forced by a model, language learners fall back to an ergative pattern that is a sort of default. Our data on early Italian go in recisely the same direction. Although we have not looked at the omission of objects of transitive verbs in this study, we know that Italian learners do not frequently omit them. In the cross-sectional corpus employed in this study, Caprin (2003) found few omission of objects. At the same time, we have found less omissions of subjects on unaccusative sentences than of subjects of transitive and unergative sentences. Our subjects, like the Chinese and American English deaf children arrange their omission of arguments around an ergative pattern.

Although these findings are interesting in themselves, their relevance extends beyond what we have discussed so far. First of all, these data show that children distinguish among different classes of verbs. This fact is at odds with Tomasello's verb island hypothesis (1992). According to this hypothesis, children learn each single instances of a verb form with its morphological properties and its argument structure. In constrast, what we found suggests that children have generalized properties of the argument structure of classes of verbs. As a matter of fact, children in our studies use the subject differently depending on the class of verbs it occurs with. In order to further support this conclusion, we looked at the unaccusative verbs used by one child (Diana)

that participated in the longitudianal study and at the unaccusativ verbs used by her caregiver. Out of 44 different forms of unaccusative verbs produced by the child, we found 11 that were used also by the adult either in the same file or in a previous file; of these, 7 were used in the same syntactic frameworks by both the adult and the child (with null subject, VS or SV order) and thus could not be considered as a sign of generalization. Four verbs, instead, were used by children in frameworks that were not employed before by the adult and thus these could be considered as the result of a generalization. In addition, 33 of the unaccusative verbs that were used by the child, were not previously used by the adults. Although these data come from a preliminary analysis, they show that children are not using the unaccusative verbs exclusively in the same syntactic frames in which they have heard them being used by the adults. Thus, very soon we observe generalizations with respect to the syntactic contexts in which verbs are used.

Our data are also relevant for they can shed a new light on the controversy concerning passive constructions in child language. Borer and Wexler (1987, 1992) argue that children have difficulties in forming verbal passives, because they do not have access to the operation of A-movement of the internal argument to Spec IP or are not able to merge the internal argument to the functional projections responsible for the assignment of case and agreement with the verb, IP in the Government and Binding (Chomsky 1981), IP in the Minimalist framework (Chomsky 1995). In order to express passives, it is claimed that children systematically employ adjectival passive constructions, that, unlike verbal passives, do not involve movement. One prediction of this proposal is that children should be unable to move the internal argument of unaccusative verbs to the preverbal position, because this is the kind of movement that children are claimed not to be able to perform in the case of passives. However, it is well known that children produce and comprehend unaccusative verbs from an early stage (Déprez & Pierce 1992; Snyder & Stromswold 1997) and that A-chains with unaccusative verbs can be formed, as children use the order SV with unaccusative verbs in several early languages (Guasti, 2002). To maintain Borer and Wexler's proposal and at the same time to make sense of the SV order with unaccusative verbs, and Babyonyshev et al. (2001) propose that children assign to clauses including unaccusative verbs the same structural analysis that they assign to clauses including unergative verbs, i.e., a structure in which the subject is generated externally to the VP. Under this account, we would expect children to treat unaccusative and unergative verbs in the same way, an expectation that is clearly not fulfilled by our data. The idea that children consider unergative and unaccusative verbs as members of the same class is in contrast with our results, that provide evidence for a clearcut distinction between these two classes of verbs in child Italian. In addition, our data show that children can either employ the order SV or VS with unaccusative verbs, i.e., they can optionally move the internal argument to Spec IP. Thus, there is no longer any reason to assume that they should not be able to do the same with passives. Consequently, the problems that children may have with passives cannot be located in their inability to form A-chains, but elsewhere.

5. Conclusions

Children learning Italian use overt subjects differently depending on the verb class: overt subjects are more frequent with unaccusative than with inergative and transitive verbs. Moreover children seem to prefer to place overt subjects preverbally with transitive and unergative verbs, while they place subjects after the verb with unaccusative verbs (or they show no preference for preverbal or postverbal position). Our data makes the hypothesis that children learn verbs on an island basis very unlikely, as children treat uniformly verbs with the same argument structures and distinguish them from verbs that have different argument structures, that is, they have generalized the argument structures across classes of verbs. More specifically, our data indicate that children do not analyze unaccusative and unergative verbs in the same way. This has consequences for the proper characterization of children problems with passives. Unaccusative verbs have properties in common with passives, in particular both involve movement of the internal argument to the surface subject position. It has been claimed that this movement cannot be performed by children in passive constructions. This hypothesis can no longer be maintained since the same movement is clearly attested in children sentences with unaccusative verbs.

References

Babyonyshev, M., Ganger, J., Pesetsky, D. &. Wexler, K. (2001). The maturation of Grammatical Principles: Evidence from Russian Unaccusatives. *Linguistic Inquiry*, 32, 1-44.

Belletti, A. (1988). The case of Unaccusatives. Linguistic Inquiry, 19, 1-34.

Bates, E. (1976). *Language and Contex: the acquisition of pragmatics. New York: Academic Press.*

Bloom, P. (1990). Subjectlesss sentences in child language. Linguistic Inquiry. 21(4),491-504.

Bloom, P. (1993). Grammatical continuity in language development: the case of subjectless sentences. *Linguistic Inquiry.* 24(4), 721-734

Borer, H., Wexler, K. (1987). The maturation of syntax. In T. Roeper and E. Williams, eds., *Parameter setting*. Dordrecht: Reidel.
Borer, H., Wexler, K. (1992). Bi-uniqe relations and the maturation of grammatical principles. *Natural Language & Linguistic Theory,* 10, 147-189.
Cabré Sans, Y. (2004) *L'adquisició de la posició del subjecte en català*, Ms, Universitat Autònoma de Barcelona.
Caprin, C. (2003). *La grammatica nei primi discorsi – Uno studio trasversale sull'acquisizione della morfo-sintassi in Italiano*. Unpublished PhD Dissertation. University of Pavia, Italy.
Chomsky, N. (1981)**.** Lectures on Government and Binding, Dordrecht: Fons
Chomsky, N. (1995). The Minimalist Program. Cambridge, Mass.: MIT Press.
Cipriani, p., Pfanner, P., Chilosi, A., Cittadoni, L., Ciuti, A., Maccari, A., Pantano, N., Pfanner, L., Poli, P., Sarno, S:, Bottari, P., Cappelli, G., Colombo, C., Veneziano, E. (1989). *Protocolli diagnostici e terapeutici nello sviluppo e nella patologia del linguaggio.* Fondazione Stella Maris.
Déprez, V. & Pierce, A. (1993). Negation and functional projections in early grammar. *Linguistic Inquiry,* 24, 25-67.
Gerken, L.A. (1991). The Metrical Basis for Children's Subjectless Sentences. *Journal of Memory and Language.* 30, 431-451.
Gerken, L.A. (1996). Prosodic structure in young children's language production. *Language,* 72, (4), 683-712.
Guasti, M.T. (1996). The acquisition of Italian Interrogatives. In Clahsen, H. eds. *The acquisition of Inflection.* Amsterdam: Benjamin.
Guasti, M.T. (2002). *Language Acquisition. The Growth of Grammar.* Cambridge, Mass.: MIT Press.
Hamann, C., Plunkett, K. (1998). Subjectless sentences in child Danish. *Cognition*, 69, 35-72.
Hyams, N., 1986. Language acquisition and the theory of parameters. Dordrecht:Reidel.
Juan-Garau, M. & Pérez-Vidal, C. (2000). Subjec realization in the syntactic development of a bilingual child. *Bilingualis: Language and Cognition,* 3, 173-92.
Koopman, H., Sportiche, D. (1991). The position of subjects. *Lingua* 85, 211-258.
Lorusso, P. (2003). *L'Acquisizione dei verbi Inaccusativi: Studio sulla produzione spontanea dei parlanti,* Unpublished MA Dissertation. University of Siena (Italy). *Siena , Università degli studi di Siena, Tesi di laurea*
Pizzuto, E. & Caselli, M.C. (1992). The acquisition of Italian morphology. Implications for models of language development. *Journal of Child Language*, 19,491-557.
MacWhinney, B., Snow, C. (1985). The Child Language Data Exchange System. *Journal of Child Language* 12, 271-296.
Radford, A. (1990a). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English.* Oxford: Blackwell.
Radford, A. (1990b). The Syntax of Nominal Arguments in Early Child English. *Language Acquisition,* 1, 195-223.
Rizzi, L. (1982). *Issues in Italian Syntax,* Dordrecht: Fons.
Rizzi, L. (1986). Null objects in italian and the theory of pro. *Linguistic Inquiry*, 17, 501-557.
Rizzi, L. (1992). Early Null Subjects and Root Null Subjects. GenGenP 1992. published in T. Hoekstra & B. Schwartz, eds., *Language Acquisition Studies in Generative Grammar*, John Benjamins, Amsterdam/Philadelphia;
Rizzi, L. (1993/1994). Some Notes on linguistic theory and Language Development: The Case of Root Infinitives. *Language Acquisition,* (384), 371-393.
Rizzi, L. (2000). "Remarks on Early Null Subjects", in Friedeman & Rizzi, eds. (2000).
Rizzi, L. (2002) On the Grammatical Basis of Language Development: A Case Study. Università di Siena Centro Interdipartimentale Studi Cognitivi sul Linguaggio sito web.
Serratrice, L. (2002). Overt subjects in English: evidence for marking of person in an English-Italian bilingual child. *Journal of Child Language*, 29, 327-355
Serratrice, L., Sorace, A. (2003). Overt and null Subjects in monolingual and Bilingual Italian Acquisition. *Proceedings of 27th the Boston Conference on Language Development*, Sommerville, Cascadilla Press.
Snyder, W., Stromswold, K. (1997). The structure and acquisition of English dative costrunctions. *Linguistic Inquiry,* 28, 281-317.
Tomasello, M. (1992). *First verbs: a case study of early grammatical development.* Cambridge: CUP.
Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition,* 40, 21-81.
Valian, V. & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language,* 23, 103-128.
Zheng, M. & Goldin-Meadow, S. (2002). Thought before language: how deaf and hearing children express motion events across cultures. *Cognition,* 85, 145-175.