

THE PROSODY OF EARLY MULTI-WORD SPEECH: WORD ORDER AND ITS INTONATIONAL REALIZATION IN THE SPEECH PRODUCTION OF ITALIAN CHILDREN

Laura D'Odorico, Mirco Fasolo

University of Milano-Bicocca- Italy
laura.dodorico@unimib.it

ABSTRACT

The purpose of this study was to investigate, in a group of Italian children, the development of the capacity to use prosodic features to mark different syntactic organizations of multi-word utterances, during the first phase of syntactic acquisition. The focus is on the prosodic realizations of multi-word utterances in which children begin to use the argument structure of verbs (vocabulary size > 400 words, MLU range 1.3 - 3.0). Results showed that non-canonical order is not marked by specific type of intonation contours and does not show specific values of duration, F0 max, F0 min or key. On the other hand, when the (optional) subject is expressed in canonical utterances it is very frequently marked by primary stress.

Keywords: language development, infant prosody, multi-word speech.

1. INTRODUCTION

Theories positing that intonation is physiologically “natural” mainly focused on the children capacity to master most of the intonation system in the prelinguistic period or during the production of first words. Only few studies have investigated the relationships between prosodic and syntactic development in early multiword speech. In a previous study by D'Odorico & Carubbi [6] we investigated how two-word utterances of varying semantic complexity are intonationally realized, focusing on the capacity of Italian children to produce two or more words in a single intonation contour and on the appearance of the phenomenon of final syllable lengthening. The main results of this study indicated that early word combinations

can be intonationally realized through different prosodic patterns. At this stage in their development children still seem to be working to achieve a correspondence between syntactic and prosodic organization, as their ability to use the prosodic model which links together the words they produce is not yet fully developed. More recently, Behrens and Gut [1] explored the integration of prosodic and syntactic development in multiword utterances produced by a monolingual German boy. Their results showed that in the period from 2;0 to 2;3 years of age some aspects of prosodic organization became increasingly integrated with the syntactic structure (i.e. pauses between the two words tended to be eliminated, with only one word bearing the predominant stress), while others (i.e. intonation contours) are not reliably related to syntactic structure.

The present study too focused on the integration of prosody and syntax, investigating prosodic characteristics of more complex multiword utterances of Italian children than those previously analyzed by D'Odorico and Carubbi and Behrens and Gut, precisely, utterances constructed around a verb. Studies on the development of the argument structure of verbs in the early syntactic period, concentrated mainly on collecting evidence regarding the hypothesis that children possess an abstract knowledge of language from birth. This hypothesis assumes that the principles of Universal Grammar are available to the child from the onset of acquisition (i.e. transitive verbs must have a direct Object complement); a contrasting hypothesis states that, based on information extracted from the input, early syntax is based on a more local understanding of how single verbs can be used (Tomasello[10]; McClure, Pine and Lieven [8]). The analysis of the prosodic realization of these types of sentences can offer a

contribution to this debate by investigating whether the children used prosodic organization to signal the utilization of non-canonical word order and to mark the different roles played by nouns (Subject or Object). The canonical word order in Italian sentences is Subject-Verb-Object, although the Subject can frequently and quite legitimately be found in the post-verbal position and the Object in the pre-verbal position.ⁱ Moreover, Italian is a null-subject language; due to its rich morphology, the explicit expression of the subject is not grammatically required. Its overt realization is mainly due to pragmatic reasons (to mark the identity of a specific agent) or to the necessity of disambiguating the referent (mainly accompanying verbs inflected in the third person).

In this respect, an analysis of the prosody of a different arrangement of the grammatical components of a multi-word utterance could highlight the planning of sentence production and give information on children's mastery of the different syntactic structures.

2. METHOD

2.1. Participants

28 Italian children (16 males and 12 females) were selected from a larger sample previously collected by D'Odorico et al. (corpus D'Odorico, 1996-2006ⁱⁱ); at the time when they were video-recorded with their mothers (mean age: 25;22 months, range: 20;10-31;15), they had a vocabulary size of over 400 words (MLU range 1.3 - 3.0) and produced more than 20 word-combinations (mean 70; range: 25-269) during the video-recorded play session.

2.2 Procedure

The spontaneous speech sample used in this study was obtained from an unstructured 30-minute play session, video-recorded when the children had a vocabulary size of over 400 wordsⁱⁱⁱ. An observer transcribed each child's productions using CHAT format (MacWhinney [7]).

2.3. Coding

All linguistic utterances produced by the children were coded. For the purposes of this study only multi-word utterances constituted by a verb and at least one noun were considered. This sub-set of

multiword utterances was subjected to a grammatical analysis, which sorted them into five distinct construction types: **CANONICAL** Subject + Verb; Verb + Direct Object; Subject + Verb + Direct Object; Subject + Verb + Indirect Object; Indirect Object + Verb + Direct Object; **NON CANONICAL** Verb + Subject; Object + Verb; Verb + Subject + Indirect Object; Direct Object + Verb + Indirect Object. All the utterances were comments made by the children on the play session activities, expressed in declarative form. This speech act uniformity is crucial in order to distinguish between the uses of prosody to mark grammatical and pragmatic aspects of the utterances.

2.4. Auditory and instrumental analysis

Auditory and instrumental analyses of all utterances were performed by the two authors, working separately, using the speech software PRAAT developed by Boersma and Weeninck [2]. The utterances were firstly analyzed in order to distinguish between Successive Single Word Utterances (SSWUs) and multi-word utterances. The criterion was the presence of a pause lasting more than 100 ms. It has been empirically determined considering that shorter pauses before stops constitute the closure phase of the consonant (Behrens & Gut [1]). In a developmental perspective we cannot assume that children at this age of language development are fully mastering the prosodic phenomena used by adults to create boundaries between intonation-groups; so we analyzed children's multi-word utterances assuming that they constitute a single intonation phrase, as usually happens for Italian declarative utterances spoken by adults (D'Imperio [5]). Considering that early two-word utterances are often produced with stress on both words (Behrens & Gut [1]) we performed also an analysis of the location of the primary stress (sentence accent). The auditory analysis was supplemented by instrumental analysis of intensity and pitch movements. When two words were stressed the primary accent was attributed to the word with the major intensity and/or the largest pitch movement (measured in semitones). The duration of utterances was calculated automatically by PRAAT, as was the register (Cruttenden [4]); We also calculated the key, according to the definition by Snow & Balog [9] (i.e. the logarithmic difference between the highest and the lowest f0 values in a utterance, measured in semitones: $[12/\log(2)] * [\log(\max f_0 - \min f_0)]$) and

the Declination^{iv} (the presence of rising or falling movements, identified with a degree of pitch change superior to 1 semitone). Declinations were further grouped as Simple (when the utterance showed 1 or 2 pitch movements: RISING, FALLING, RISING-FALLING, FALLING-RISING) or Complex when the pitch movements were more than 2 (e.g. FALLING-RISING-FALLING-RISING).

3. RESULTS

The children produced many word combinations (70 on average, range: 25-269), of which approximately 18% (N = 360, mean = 13, range = 1 - 39) contained a verb. Seventy-one word combinations contained also at least one noun but three word combinations were excluded from the analysis due to mother-child voice overlap. The final number of combinations submitted to analyses was 68 (41 were two-word utterances and 27 multi-word utterances; see Table 1)

Table 1: Frequency of occurrence for the different types of multiword utterances

Word order	N	%
Canonical		
(N = 48, 70.59%)		
IO + V + DO	2	2.9
S + V + IO	1	1.5
S + V	5	7.4
S + V + IO	5	7.4
S + V + IO + DO	2	2.9
S + V + DO	5	7.4
V + IO + DO	4	5.9
V + DO	24	35.3
Non canonical		
(N = 20, 29.41%)		
DO + V	5	7.4
S + DO + V	1	1.5
V + IO + DO + S	1	1.5
V + IO + S	1	1.5
V + DO + S	5	7.4
V + S	7	10.3

The vast majority of the utterances were produced in a single continuous intonation pattern, showing that at this stage of language development the increased ability to produce grammatical complex sentences corresponds to the ability to link several words together in a single intonation contour (see fig.1). This ability appears now well consolidated while in the previous stage of language development (i.e. utterances produced at a vocabulary size of about 200 words) only about 30% of the utterances were produced in a single intonation pattern (D'Odorico and Carubbi [6]).

There were only 8 cases in which there was a pause (ranging from 107 to 897 ms) between the two words, which are uttered with distinct intonation contours (see fig.2).

Figure 1: Example of a single intonation contour for the utterance *cavallo no(n) vedo* 'horse not see'

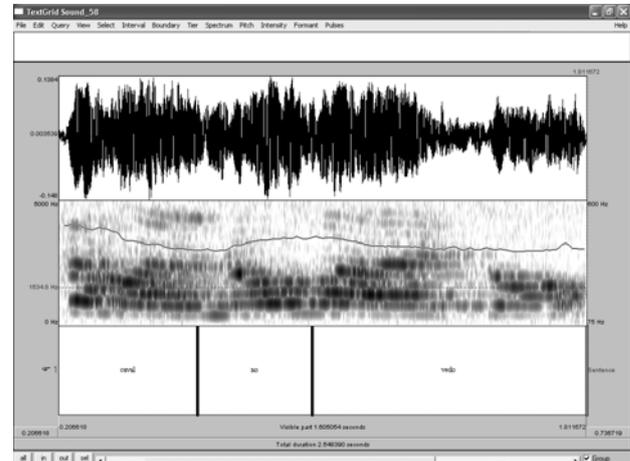
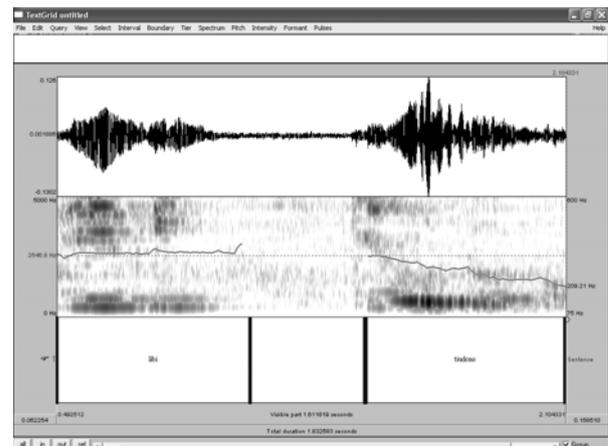


Figure 2: Example of two distinct intonation contours for the utterance *libri chiudono* 'books (they) close'.



In our data the distinction between NON-CANONICAL and CANONICAL ORDERS is hardly marked at all from a prosodic point of view, as there are no significant differences in duration, F0 max, F0 min or key between the two types of construction.

Table 2: Descriptive statistics of pitch values and duration for canonical and non-canonical utterances.

	Canonical		Non-canonical	
	Mean	SD	Mean	SD
duration	1,67	0,52	1,72	0,55
register	332	47	343	49
key	7,17	2,90	8,52	3,17
F0 max	632	319	445	93
F0 min	352	75	192	78

As can be seen from Table 3 there is a great variability in the intonation contours used by children, even if we controlled the type of speech act. For example, the most frequent type of structure, i.e. V +DO, can be expressed by almost all type of intonation contours.

Table 3: Distribution of different contours for canonical and non-canonical utterances.

	Simple			Complex		
Canonical	--	\	^	V	V/V	^^
IO + V + DO			1			1
S + IO + V + IO					1	
S + V			5			
S + V + IO			1	1	1	2
S + V + IO + DO			1			1
S + V + DO			1		1	2
V + IO + D O				1	2	1
V + DO	1	7	5	3	3	6
tot	1	7	14	5	8	13
Non canonical						
DO + V			3			2
S + DO + V						1
V + IO + S						1
V + IO + DO + S				1		
V +DO + S			4			1
V + S			4	1		2
tot			11	2		7

The Rise-Fall pattern (40%) was the most frequent type of intonation contour observed; overall, approximately 78% of intonation contours ended with a lower pitch in the final part of the utterances. Simple or Complex Rising contours accounted for approximately 21%. Only 1 utterance was observed in which there was no discernible pitch movement (i.e. it was FLAT). Non-canonical order is not marked by specific type of intonation contour.

More than two pitch movements occurred in approximately 37% of the cases observed during the utterance production, resulting in complex intonation patterns as Falling-Rising-Falling (15%), Rising-Falling-Rising-Falling (12%), and Rising-Falling-Rising (10%).

The location of primary stress is the only parameter which appeared to be influenced by word order, but only with regard to the location of Subject and not as we hypothesized. The overt subject is expressed in approximately 48% of the analyzed utterances. When the word order is canonical, i.e. when the subject occupies the first position in the utterance, primary stress occurs on it in 67% of the cases. When the order is reversed (15 utterances), the

subject is prosodically marked in three cases (20%) only and in two of these the marking was achieved by producing two different intonation groups in the same utterance. So the primary stress it is not used to mark the non-canonical location of the subject. The object, on the contrary, is marked by primary stress with a similar frequency (approximately 50% of the cases), both in the pre- and post-verbal position. Verbs bore primary stress prevalently when in the first position (82%) (i.e. no subject, canonical order: verb-object).

The position of the primary stress and the grammatical role of the words, however, influenced the type of Declination: Rising pitch movements and Complex contours were almost never present when the primary stress was on the verb.

4. CONCLUSION

The study reports preliminary data on prosodic characteristics of Italian children utterances constructed around a verb. The main developmental phenomena which have been investigated regarded a) the capacity to link more words in a single intonation contour, which many studies reported not to be fully developed at the beginning of the combinatorial speech, b) the distribution of the different intonation contours according with the use of canonical or non-canonical order of grammatical arguments c) the influence of word order on location of stress and other prosodic phenomena (utterance duration, F0 max, F0 min, key). Data indicated that even during the period in which children produce multiword utterances showing the commencement of the capacity to manage the verb argument structure, the prosodic structure is far from being fully integrated with the syntactic organization. Variability in intonation contours is very high, even when the speech type is controlled (i.e. all the utterances are comments on a play activity). Non-canonical order did not appear to be prosodically marked in so far that Subject beared primary stress more when it is in a canonical location than when it is not. In this phase of language development, therefore, non-canonical order probably results from a not fully developed competence of argument structure rather than from an option selected by children to mark the subject or the object pragmatically. The only syntactic aspect which is consistently marked by prosodic means is the overt subject, suggesting that the children do recognise the optionality of this linguistic choice in Italian. Data on the prosodic features used by Italian adults to mark canonical or non canonical order would be useful in order to

fully clarify the nature of the developmental phenomena we observed.

In conclusion, our results support the view of a partially independent development of prosody with regard to other aspects of language acquisition and of a complex process of integration between prosody and syntax that continues in the multiword speech period. (see also Behrens and Gut [1]).

5. REFERENCES

- [1] Behrens, H., Gut, U. 2005. The relationship between prosodic and syntactic organization in early multiword speech. *Journal of Child Language*. 32, 1-34.
- [2] Boersma, P., Weeninck, D. 2005. Praat: doing phonetics by computer (version 4.3.14). <http://www.praat.org/>.
- [3] Caselli, M.C., Casadio, P. 1995. *Il primo vocabolario del bambino*. Milano: Franco Angeli.
- [4] Cruttenden, A. 1997. *Intonation*. Cambridge: Cambridge University Press.
- [5] D'Imperio, M. 2002. Italian Intonation: an overview and some questions. *Probus*. 14(1) 37-69.
- [6] D'Odorico, L., Carubbi, S. 2003. Prosodic characteristics of early multi-word utterances in Italian Children. *First Language*. 23 (I), 97-116.
- [7] MacWhinney, B. 1997. *Il progetto CHILDES: Strumenti per l'analisi del linguaggio parlato*. (U.B. Edizione italiana a cura di E. Pizzuto, Trans.). Pisa: Del Cerro.
- [8] McClure, K., Pine, J.M., Lieven, E.V.M. 2006. Investigating the abstractness of children's early knowledge of argument structure. *Journal of Child Language*. 33, 693-720.
- [9] Snow D., Balog, H. 2002. Do children produce the melody before the words? A review of developmental intonation research, *Lingua*. 112, 1025-1058.
- [10] Tomasello M. 1992. *First verbs: a case study of early grammatical development*. Cambridge: C.U.P.

ⁱ e.g., *Ho mangiato il gelato* '(I) ate the ice cream', V+DO, canonical order; *Il gelato ho mangiato* DO + V, non-canonical order. The pragmatic function of the non-canonical order is, as in the example, to mark the fact that I ate the ice cream and not something else.

ⁱⁱ The full corpus comprehends the videorecorded sessions of five different populations of infants observed during mother-infant play interaction during the second and the third year of life. Up to now only some of the sessions have been transcribed in CHAT format. More specifically, the spontaneous speech sample used in this study are produced by a group of children whose vocabulary development has been monitored from about 16 to about 30 months of age. At specific stages of vocabulary development children participated to play sessions with mothers in order to obtain samples of spontaneous speech.

ⁱⁱⁱ Vocabulary development was assessed by means of the Italian version of the MacArthur Communicative Inventory (Caselli & Casadio [3])

^{iv} We used the definition by Snow & Balog [9]: "declination describes the overall configuration of rises and falls in the pitch of the voice across utterance" p.1027.

