

**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, Daniela Marchione

University of Milano-Bicocca (Italy)

(in corso di stampa su *Enfance*, novembre 2009)

For communication:

Prof. Laura D'Odorico

Università degli Studi di Milano-Bicocca

Piazza dell'Ateneo Nuovo, 1

20126 – Milano (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, maximum and minimum F0, or key.

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

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 multi-word speech. In a previous study by D’Odorico and Carubbi (2003), it has been 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 the study indicated that early word combinations can be intonationally realized through different prosodic patterns. At that stage of 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 (2005) explored the integration of prosodic and syntactic development in multi-word 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 multi-word utterances of Italian children than those previously analyzed by D’Odorico and Carubbi (2003) and Behrens and Gut (2005), 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 (i.e. transitive verbs must have a direct object complement) are available to the child from the onset of acquisition; 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, 1992; McClure, Pine & Lieven, 2006). The analysis of the prosodic realization of these types of sentences can offer a contribution to this debate by investigating whether the children use prosodic organization to signal the utilization of non-canonical word order and to mark the different roles played by nouns (subject or object). The analysis of the prosody of different arrangements of the grammatical components of multi-word utterances could highlight the planning of sentence production, in a period in which grammatical competence is not fully acquired, and give information on children's mastery of the different syntactic structures.

The characteristics of Italian are particularly suitable to investigate the interplay between order of words and prosodic characteristics of the utterances, in so far the canonical word order in Italian sentences is Subject + Verb + Object, but the Subject can frequently and quite legitimately be found in the post-verbal position and the Object in the pre-verbal position. For example, I can say "*Ho mangiato il gelato*" [(I) ate the ice cream], but to say "*Il gelato ho mangiato*" [The ice cream (I) ate], i.e. using a non-canonical order, marks the fact that I ate the ice cream and not something else. When adult use a non-canonical order to emphasize a specific word, this word, collocated in the first position, is even prosodically stressed (D'Odorico & Fasolo, 2008). Moreover, Italian is a null-subject language, therefore the explicit expression of the subject, which is not grammatically required, 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).

METHOD

Participants

Twenty-eight Italian children (16 males and 12 females) were selected from a larger sample previously collected by D'Odorico et al. (corpus D'Odorico, 1996-2006ⁱ); 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.

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 wordsⁱⁱ. An observer transcribed each child's productions using CHAT format (MacWhinney, 1997).

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 or one personal pronoun with Subject or Object (Direct or Indirect) function, were considered. This sub-set of multi-word utterances was subjected to the analysis of the structure, and distinguished in: **CANONICAL ORDER**, constituted by utterances with Subject in pre-verbal position and/or by Object in post-verbal position (N = 74); **NON-CANONICAL ORDER**, constituted by utterances with at least one element in a non-canonical order (Subject in post-verbal position, and/or Object in pre-verbal position) (N = 47). 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 or pragmatic aspects of the utterances.

Auditory and instrumental analysis

Auditory and instrumental analyses of all utterances were performed by two of the authors, working separately, using the speech software PRAAT developed by Boersma and Weeninck (2005). 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. This value has been empirically determined considering that shorter

pauses before stops constitute the closure phase of the consonant (Behrens et al., 2005).

Working within a developmental perspective, we cannot assume that children at this stage of language development are fully mastering the prosodic phenomena used by adults to create boundaries between intonation-groups; therefore, 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, 2002).

Considering that early two-word utterances are often produced with stress on both words (Behrens et al., 2005) we performed also an analysis of the location of the *primary stress*, defined as the principal pitch prominence in one intonation-group (or utterances with major grammatical constituents like simple sentences) (Cruttenden, 1997). 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 (Boersma et al., 2005), as was the *register*, i. e. the pitch range operated by the speaker (Cruttenden, 1997); we also calculated the *key*, according to the definition by Snow and Balog (2002) (i.e. the logarithmic difference between the highest and the lowest f0 values in a utterance, measured in semitones: $[12/\log(2)] * [\log(\text{maximum } f_0 - \text{minimum } f_0)]$) and the *declination*ⁱⁱⁱ (or the presence of rising or falling movements, identified with a degree of pitch change of at least 2 semitones). We suggest that declinations in Italian are not only positioned in correspondence of the nuclear tone, as pointed out by Cruttenden (1997), but are also possibly situated in any other part of the utterance having a pitch change of at least 2 semitones. Following that, 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. rising-falling-rising).

The proportion of agreements between the two coders on classification of declinations was 94%. There were also strong correlations on calculations of duration ($r = 0,996$), F0 ($r = 0,992$),

maximum F0 ($r = 0,992$), and minimum F0 ($r = 0.982$) performed by the two coders.

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. One-hundred-twenty-four word combinations contained also at least one noun or a personal pronoun, but 3 of these were excluded from the analysis due to mother-child voice overlapping. The final number of word combinations submitted to analyses was 121 (see Table 1).

<INSERT ABOUT HERE TABLE 1>

There were only 12 utterances in which there was a pause (ranging from 107 to 897 ms) between the two words, which are uttered with distinct intonation contours (see Figure 1); 2 of them were performed as partial reproduction of an adult patterned utterance, suggesting that children are still operating transitional prosodic profiles.

The vast majority of the utterances (90%) 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 Figure 2). 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 et al., 2003).

<INSERT ABOUT HERE FIGURE 1 AND FIGURE 2>

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, maximum F0, minimum F0 or key between the two types of construction (see Table 2). On the contrary, adults tend to prosodically differentiate the two orders, reducing register and key when producing a non-canonical order (D'Odorico and Fasolo, 2008).

<INSERT ABOUT HERE TABLE 2>

As can be seen from Table 3, simple intonational patterns were the most frequent type observed both in canonical order and in non-canonical; no significant differences were observed in any type of construction.

There is a great variability in the intonation contours used by children, even if we controlled the type of speech act. For each type of construction, we tested the possibility that it was realized through a particular type of intonational contour. Results of the analyses show that there were no significant differences between 'S + V' and 'V + S' utterances with regards to the distribution of intonational contours ending with rising or with falling F0 ($Z = 0,25$; $p = NS$). A similar pattern was observed for 'V + O' *versus* 'O + V' utterances ($Z = 1,264$; $p = NS$) and for 'S + V + O' *versus* 'O + V + S' ($Z = -0,06$; $p = NS$).

<INSERT ABOUT HERE TABLE 3>

The location of primary stress (see Table 4) 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 and when the word order is canonical, i.e. when the Subject occupies the first position in the utterance, primary stress occurs on it in 29% of the cases, while, when the order is reversed (32 utterances),

the Subject is prosodically marked in eight cases (14%), so the primary stress it is used to mark the canonical location of the Subject ($Z = 2,793$; $p = 0,003$).

The Object is marked by primary stress approximately 45% of the cases, 32% when it is in the pre-verbal position and 13% when in the post-verbal position. Even if we observed a greater use of the primary stress on the Object when the order is non-canonical, the difference does not reach the statistical significance ($Z = 0,103$; $p = \text{NS}$).

Verbs bore primary stress prevalently when in the first position (61%) (i.e. no subject, canonical order: Verb + Object).

<INSERT ABOUT HERE TABLE 4>

CONCLUSION

The study reports 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, maximum and minimum F0, key). Data indicated that even during the period in which children produce multi-word 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.

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 multi-word speech period (see also Behrens et al, 2005).

REFERENCES

- Behrens, H., Gut, U. (2005). The relationship between prosodic and syntactic organization in early multi-word speech. *Journal of Child Language*, 32, 1-34.
- Boersma, P., Weeninck, D. (2005). *Praat: doing phonetics by computer (version 4.3.14)*. <http://www.praat.org/>.
- Caselli, M.C., Casadio, P. (1995). *Il primo vocabolario del bambino*. Milano: Franco Angeli.
- Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.
- D'Imperio, M. (2002). Italian Intonation: an overview and some questions. *Probus*, 14, 37-69.
- D'Odorico, L., Carubbi, S. (2001). Early Multi-Word Utterances in Italian-Speaking Children. In M. Almgren, Barrena, A., Ezeizabarrena, M.-J., Idiazabal, I., MacWhinney, B. (Ed.), *Research on Child Language Acquisition*. Sommerville: Cascadilla Press.
- D'Odorico, L., Carubbi, S. (2003). Prosodic characteristics of early multi-word utterances in Italian children. *First Language*, 23, 97-116.
- D'Odorico, L., Carubbi, S., Salerni, N. & Calvo, V. (2001). Vocabulary development in Italian children: a longitudinal evaluation of quantitative and qualitative aspects. *Journal of Child Language*, 28, 351-372.
- D'Odorico, L., Fasolo, M. (2007). Nouns and verbs in vocabulary acquisition of Italian children. *Journal of Child Language*, 34, 891-907.
- D'Odorico, L., Fasolo, M. (2008, October). Prosodia e sintassi: uno studio sulla lingua italiana. *Giornata scientifica in onore di Ino Flores D'Arcais*, Padova.
- D'Odorico, L., Jacob, V. (2006). Prosodical and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language and Communication Disorders*, 41, 293-311.
- 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.
- 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.

Salerni, N., Assanelli, A., D'Odorico, L. (2007). Qualitative aspects of productive vocabulary at 200-word stage: A comparison between spontaneous speech and parental data report. *First Language*, 27, 75-87.

Snow, D., Balog, H. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua*, 112, 1025-1058.

Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Cambridge: Cambridge University Press.

Footnotes

ⁱ The full corpus comprehends the video-recorded sessions of five different populations of infants observed during mother-infant play interaction during the second and the third year of life. Up to the time of this study the language development of the children included in the corpus was analyzed monthly for vocabulary development and characteristics of early two-word utterances (D'Odorico & Carubbi, 2001; 2003; D'Odorico, Carubbi, Salerni & Calvo, 2001; D'Odorico & Fasolo, 2007; D'Odorico & Jacob, 2006; Salerni, Assanelli & D'Odorico, 2007) .

ⁱⁱ Vocabulary development was assessed by means of the Italian version of the MacArthur Communicative Inventory (Caselli & Casadio, 1995)

ⁱⁱⁱ We used the definition by Snow and Balog (2002): “declination describes the overall configuration of rises and falls in the pitch of the voice across utterance”, p.1027.

^{iv} O + V + S tag was selected to identify non-canonical order for utterances containing all the three elements, and it doesn't respect the real sequence of constituents.

TABLE 1: *Frequency of occurrence for the different types of multi-word utterances.*

Word Order	Structure	N	%
Canonical (N = 74; 61,16%)			
	S + V	13	10,74
	S + V + O	19	15,70
	V + O	42	34,71
Non-canonical (N = 47; 38,84%)			
	O + V	14	11,57
	O + V + S ^{iv}	12	9,92
	V + S	21	17,36

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

	Canonical order		Non-canonical order		t	p
	Mean	S.D.	Mean	S.D.		
Duration (Sec.)	1,582	0,51	1,491	0,46	1,009	0,315
Register (Hz)	340	50	334	47	0,584	0,560
Key (semitones)	10	7	13	6	-1,756	0,082
Maximum pitch (Hz)	427	76	439	89	-0,809	0,420
Minimum pitch (Hz)	243	71	221	68	1,687	0,094

TABLE 3: *Distribution of different simple and complex intonational contours for canonical and non-canonical utterances.*

Word Order	Simple					Complex			
	--	\	/	^	v	^v	v^	^^	v^v
Canonical									
S + V		1	1	7			1	1	2
S + V + O		1		5	1	1	1	3	7
V + O	1	4	2	13	11	2	1	4	4
Total	1 (1%)		46 (63%)				27 (36%)		
Non Canonical									
O + V		2		7	1		2		2
O + V + S	1	2		3	2	1			3
V + S		6	1	7	3	1	1		2
Total	1 (2%)		34 (73%)				12 (25%)		

TABLE 4: *Location of primary stress per word order and grammatical constituents.*

Word Order	Subject (N = 58; 47,9%)		Object (N = 84; 69,4%)	
	Stressed	Not stressed	Stressed	Not stressed
Canonical	17 (29%)	9 (16%)	27 (32%)	31 (37%)
Non-canonical	8 (14%)	24 (41%)	11 (13%)	15 (18%)
Total	25 (43%)	33 (57%)	38 (45%)	46 (55%)

FIGURE 1: *Example of two distinct intonation contours for the utterance “Libri chiudono” [Books (they) close].*

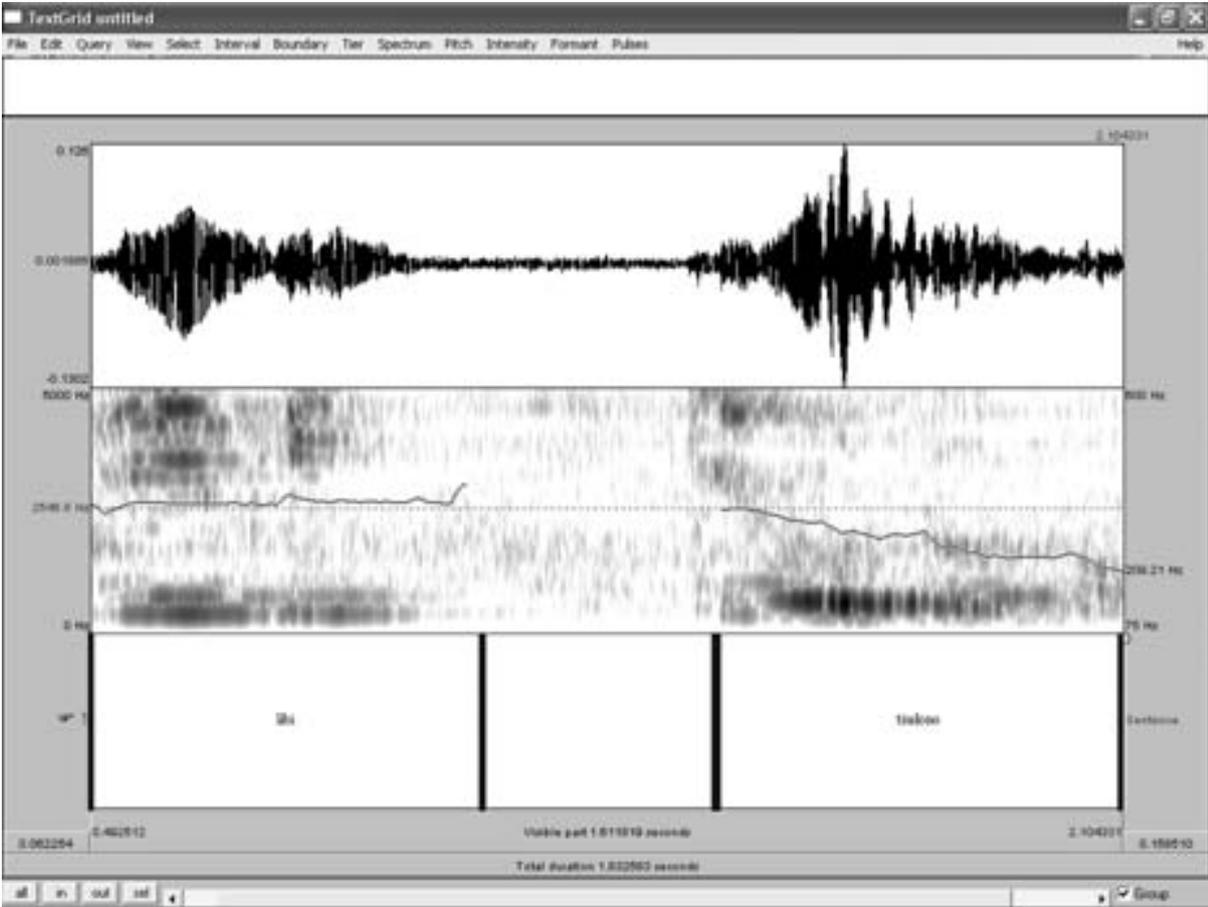


FIGURE 2: Example of a single intonation contour for the utterance “Cavallo no(n) vedo” [Horse (I) don’t see].

