

A PHONETIC AND ACOUSTIC STUDY ON THE EARLY SPEECH DEVELOPMENT OF TWO ITALIAN CHILDREN

Claudio Zmarich and Franco Ferrero

Istituto di Fonetica e Dialettologia – C.N.R., Padova, Italy

ABSTRACT

In this paper, the phonetic aspects of babbling and early words are investigated through auditory and acoustic analysis. Two children were recorded at 10, 12, 14, 16 months of age during play situations in the presence of their mothers. Every occurrences of babbling episodes and early words suitable for acoustic analysis were classified for consonantal and vocalic categories according to IPA symbols. Acoustic analysis was then performed on the vowels to describe their F1 and F2 values. The acoustic values of the vowels were compared longitudinally with the available data for a representative sample of 7-years-old Italian children and adults (males and females).

1. INTRODUCTION

The great individual variability of the speech development during babbling and the first words stage has been highlighted by a number of authors (cf. [1]).

The key issue is represented here by the complex time-course of the progressive disengagement from the bio-mechanical constraints grounded in the canonical babbling towards the achievement of the phonetic mastery for the first phonological oppositions. This process of development involves a large number of variables of different nature, both internal to the child (biological time-schedules) as well as external (the language spoken in the environment), strictly interacting together.

The stages of this evolution may be depicted essentially in two ways: either through the use of frequency statistics based on auditory classifications of the tokens along the phones' categories, or through a measurement of the physical properties of the signals, such as formant frequencies (Hz), dependent on the anatomic shape and dimension of the vocal apparatus and the mastery in its use, and durations (s). Although temporal measurements were also taken in the signals studied here, we concentrated only on frequency statistics of phones and CV associations, and on F1 and F2 measurements.

2. METHOD

2.1. Subjects and Materials. The two children attending the study were Davide (a male) and Caterina (a female). They were born full-term, and their course of development were reported to be normal and free of any specific problems relating to their speech and hearing apparatus. Their parents are middle-class, and language spoken is Italian. The children were recorded every two weeks from the 8th month onwards, during play situations and in the presence of their mothers. The recordings, lasting 45 minutes each, were made at their respective home in a noise-free environment. Davide was recorded with an AIWA tape recorder, mod. TP-850 and a Sony microphone, mod. ECM-T7, whereas Caterina was recorded with a Sony DAT recorder, mod. TCD-D8 and a Sony microphone, mod. ECM M5 907. The analysis of

only 4 recordings are presented here, referring to the 10th month (9;09 and 9;28, months and days, for Davide; 9;05 for Caterina), 12th month (11;21 for Davide; 11;12 for Caterina), 14th month (13;07 for Davide; 13;14 for Caterina) and 16th month (15;16 for Davide; 15;14 for Caterina). According to the MacArthur Instrument [2], Davide were producing 4 words at 9;09, and 31 words at 13;01, whereas Caterina were producing 6 words at 11;6 and 13 at 12;16.

2.2 Preliminary transcription and tokens' selection. A first phonetic transcription, performed by the first author using the symbols and diacritics provided in [3] besides those of the IPA (1996), was the basis to select babbling and word occurrences, that were acquired at 25.6 kHz sampling frequency with the CSL 4300 package and examined for compliance with the acoustic criteria proposed by Oller [4]. In this manner, it was possible to exclude any noised production and discriminate babbling and any words from vegetative and reflexive sounds as well as from sounds belonging to the *vocal play* stage. Isolated CV syllables were accepted, but only vowels produced with loosely defined "modal" voice (i.e. neither laryngalized, strongly nasalized, desonorized, nor in falsetto) and preceded by "true" consonants (no glottal consonants) were analyzed. The occurrences characterized by F0 shift, harmonic doubling, biphonation [5] were also tentatively individuated and excluded. This revision resulted in the elimination of 422/1172 potential babbling utterances. Two successive syllables were considered to refer to different utterances if separated by more than 500 ms. In the presence of vocalic modulations, the number and type of occurring vowels was specified, keeping in mind the possibility that a vowel may have very long transitions.

2.3. Final transcription. A new phonetic transcription was performed by the first author on the selected occurrences, exploiting the facilities of the CSL 4300 package and an interactive phonetic symbol guide [6]. Further, the whole process relative to the final phonetic classification was aided by the comparison with an interactive archive of the stored vowels which had received the same classification by 4 independent transcribers (at least 2 tokens for each vowel). The validity of this transcription was checked with two different procedures. For Davide, the first author and an expert transcriber worked jointly on the 25 % of the whole corpus, and discussed every disagreement, attempting to achieve a common classification. The final agreement score for the IPA symbols was quite high (98,3%). As for Caterina, the agreement's percentage of 3 out 4 independent judges on a sample of 168 vowels, randomly chosen across all the recordings, was 31.5% (40.4% not considering the tract "rounded/unrounded"). The percentage score for the consonants was calculated on the transcriptions, performed by 2 independent experts, of randomly selected samples of 10

consecutive utterances for each recordings. By applying the formula in [7], the resulting score was 59.1%.

2.3. Acoustic analysis. The formant analysis on vowels was performed on the most stable part (i.e. where F1 and F2 were possibly flat, clearly discernible and separated from each other). For intra-subject agreement, the first 100 LPC analysis (14 poles, sometimes 16 poles), averaged on these portions, was compared with a manual interpolation of the harmonic envelope (50 Hz bandwidth) on the same occurrences. As the results of a t-test did not reveal any significant difference between the two series of measurements, preference was given to the automatic LPC measures, except for the evidently wrong cases. The inter-subject agreement was checked on 10% of the whole corpus with a t-test, by comparing the LPC measurements made by the first author and that performed by the second author, and differences were not statistically significant at any reasonable level.

3. RESULTS

3.1. Some quantitative characteristics of the corpus. As a generale note, the results pertaining to Davide, though already published in [8], will be presented and discussed here where needed. Table 1 shows statistics describing some quantitative characteristics of the analyzed corpus.

months	No. babbling and words utterances / No. total vocalizations		No. of analyzed syll. in babbling (in words)		No. of analyzed syllables in mono-/pluri-syllables	
	Davide	Caterina	Davide	Caterina	Davide	Caterina
10 th	Not avail.	137/328	93	140	35-58	13-127
12 th	98/294	283/480	91(1)	241(9)	27-46	82-159
14 th	81/182	209/274	60(9)	155(41)	38-22	50-93
16 th	122/163	190/342	115(51)	222(61)	31-84	35-187

Table 1. Counts of babbling utterances and words over total vocalizations, counts of the analyzed syllables (those belonging to words are enclosed between parenthesis) and counts of the analyzed syllables belonging to either monosyllables or plurisyllables.

The total number of syllables analyzed are 359 for Davide and 758 for Caterina. The babbling utterances and the first words increase their proportion (along time) over the total number of vocalizations (col. 1), and the syllables in words increase their proportion among all the analyzed syllables (col. 2).

3.2. Results based on auditory classifications. After classifying the types of syllables produced by the 2 subjects in terms of CV structure, their relative frequency over the total number of syllables was compared to the relative frequency of the syllables types in a selection (440 syllable tokens), performed by the first author, of the target words attested by parents as present in the vocabulary of at least 3 out of 315 children (appendix B of [2]). The two most frequent types within the target lexicon are CV (61.1%) and CVC (20.2%). The CV type also constitutes the great majority of the productions for the 2 children, ranging from 80% to 90% for Davide, and always above 90% for Caterina. However, a direct comparison between the children's and adults' percentages is made difficult by the fact that most final consonants of the CVC types of the adult lexical selection consists of a geminated consonant, and consonantal lengthening was never transcribed in the children's syllables. Table 2 lists the

percentage of occurrence of the most frequent vowels (above the 5% level) across the 4 stages for the 2 children, together with 2 statistics referring to standard Italian [9], [10] and a statistics referring to the lexical selection from [2].

10 th		12 th		14 th		16 th		ITALIAN		
Dav	Cat.	Dav	Cat.	Dav	Cat.	Dav	Cat.	[9]	[10]	[2]
e	ε	e	a	a	e	a	a	a	a	a
233	135	187	198	267	167	226	225	269	287	317
ɪ	ə	ε	æ	ɪ	a	i	ɑ	i	e	o
172	121	110	107	133	161	104	121	210	233	247
i	e	ɪ	ə	e	ε	ɑ	æ	e	o	e
86	85	88	103	117	96	96	108	200	217	146
ε	ɑ	o	ɑ	æ	æ	o	ε	o	i	i
86	85	88	95	100	83	70	94	189	157	112
o	æ	ə	ɒ	i	ɑ	ɪ	ɔ	u	u	u
86	85	77	87	50	71	61	72	52	45	67
i	a	a	ɔ	u	ɔ		e	ε	ε	ɔ
54	78	66	66	50	51		54	43	36	56
	ʌ	i	ε	y	ə			ɔ	ɔ	ε
	50	55	53	50	51			25	21	54
	o			o						
	50			50						
oth.	oth.	oth.	oth.	oth.	oth.	oth.	oth.			
269	31.1	335	29.1	283	22.0	443	32.6			

Table 2. List of the most frequent vowels (%). As to standard Italian, 3 statistics are reported: [9, 10, 2].

This table clearly illustrates the common preference accorded from the beginning to frontal vowels, and the scarce presence of back vowels. However, the 2 children differentiated for their production of low vowels, as their prevalence for Davide is late to emerge (from the 14th month). The reference to low and mid-low vowels is important, because the /a/ type is the most frequent in the target vocalic system (though somewhat differing, the three sources for the vowels frequencies in Italian agree upon the prevalence of the /a/). At the final stage the children's systems are still lacking of the back, rounded vowels of the adult language. These considerations receive further confirmation from the distribution in the three main classes along the front-back axis of all the vowels produced (fig. 1). Differences between the 2 subjects are persistent until the 12th month, although from about the 14th month onward the respective distributions become more and more similar to each other and to those of Italian.

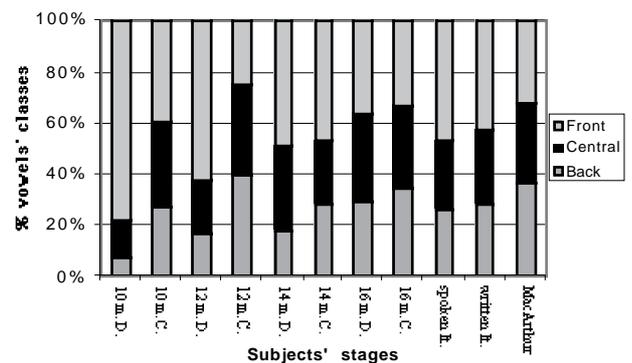


Figure 1. Frequency occurrence percentages of vowels in the productions of Davide (D) and Caterina (C) and three main sources for Italian [9, 10, 2].

3.3. Acoustic description. In the Figs. 2 and 3 the vowels produced by Davide and Caterina and classified as occurrences of standard Italian are plotted in the acoustic space (Hz) defined by the F1 and F2 coordinates (any potential allophones were excluded).

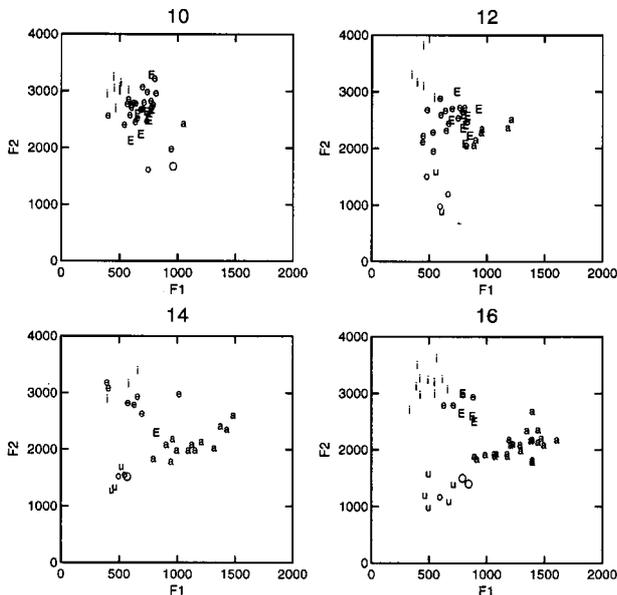


Figure 2. Plots of the vowels belonging to standard Italian produced by Davide in each of the 4 stages (E=/ε/; O=/ɔ/).

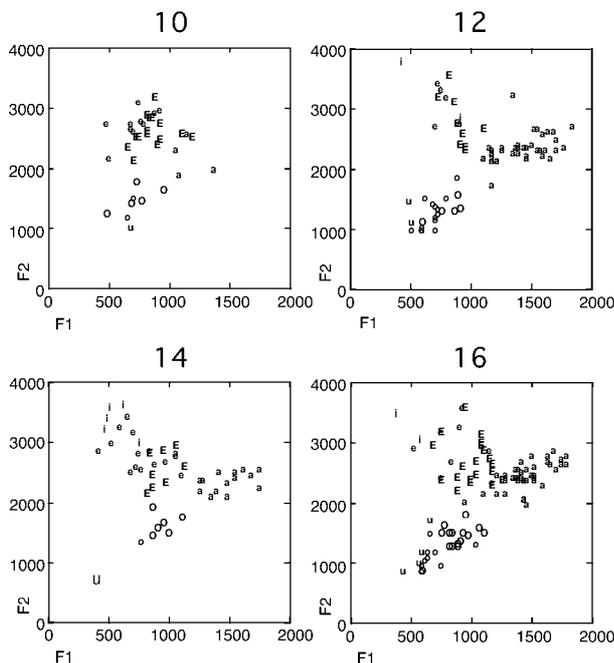


Figure 3. Plots of the vowels belonging to standard Italian produced by Caterina in each of the 4 stages (E=/ε/; O=/ɔ/).

Comparing these figures, a similarity in distribution is evident only in the in the 14th month, when both children

converge towards the adult model, distinguishing well the area of existence of all the vowels, although the ways in which they reach this final point are different. Davide starts from a virtual sticking to frontal articulations, with values concentrated in the high and right quadrant of the figure, and gradually elongates first the F2 axis and then the F1 axis. The starting distribution of Caterina is more central and balanced, and in the following stages spreads harmoniously along both axes.

In figure 4, a comparison is proposed between the mean values of the cardinal vowels /a, i, u/ produced by Davide and Caterina in 2 main stages (the 10th month with the 12th month, and the 14th month with the 16th month), and the centroids of existence for the vowels of the subjects of [11]. The adult subjects are a representative sample of Italian male and female speakers of northern Italian, and the children attend the first three classes of primary school.

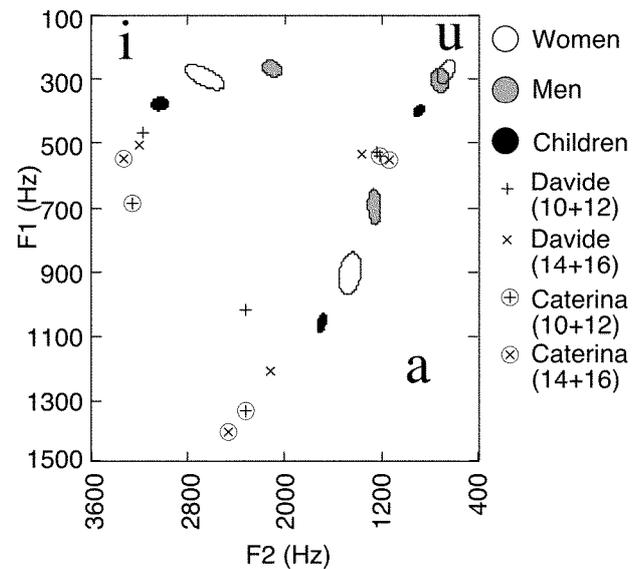


Figure 4. Areas of existence for the cardinal vowels produced by the subjects of [11] (centroids) and Davide (mean values).

Although there is no systematic lowering of the formant frequencies along with increasing of the ages, the vowels produced by Davide and Caterina in the final stage are grossly in line with the centroids. These data are best interpreted by assuming a similarity in the shape of the vocal tract and a linear factor active in the proportional growth of the vocal tract for the ages considered here.

3.4. Consonant-vowel associations. Fig. 5 illustrates a comparison between frequencies of consonants pronounced by Davide and Caterina and the frequencies attested for Italian, based on the classification of the active articulators, resulting in three classes: **labials** (bilabials and labiodentals), **apicals** (dentals, alveolars, postalveolars, retroflexes), and **dorsals** (palatals, velars, uvulars, pharyngeals). Only the initials of the CV and CVC syllables were considered for the productions of the two children and the words listed in the selection from [2]. Even with regard to consonants, the final stages of the 2 children are similar to each other and to the the adult target, although the most dramatic changes along the way are achieved this time by

Caterina (see the inverted proportions of labials and dorsals at the 10th month and at the 14th month).

Data referring to vowel and consonant occurrence frequency, though useful, cannot provide information regarding co-occurrence frequencies, and then preferred syllable types.

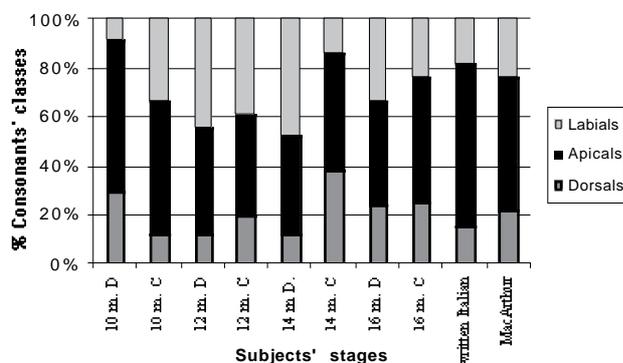


Figure 5. Frequency occurrence percentages of labial, apical and dorsal consonants in Davide's productions and in [10] and a selection from [2].

The intra-syllabic hypothesis of MacNeilage and Davis's "Frame, then content" theory [12] makes precise predictions regarding the co-occurrences of vowels and consonants within the syllable: front vowels tend to occur after apex consonants; central vowels tend to occur after labial consonants; back vowels tend to occur after dorsal consonants. Table 3 describes the relationship between the observed and expected frequencies, by reporting their ratio for each cell (the expected value is 1.0). The predicted associations are highlighted in bold character. Data are divided up into two groups (the first 2 and the last 2 stages).

D.	L	A	D	D.	L	A	D
10+12				14+16			
n. 173				n.161			
F	0.760	1.121	0.969	F	0.476	1.407	1.204
C	1.549	0.710	1.112	C	1.322	0.781	0.816
B	1.498	0.772	0.999	B	1.278	0.743	0.969
p<.05					*	*	
C.	L	A	D	C.	L	A	D
10+12				14+16			
n. 374				n. 377			
F	0.609	1.486	0.522	F	0.262	1.155	2.161
C	1.270	0.905	0.680	C	0.521	1.376	1.009
B	1.067	0.683	1.714	B	0.880	0.715	2.028
p<.05	*	*	*		*	*	*

Table 3. Ratio of the observed-to-expected co-occurrence frequencies for labial (L), apical (A) and dorsal (D) consonants with the front (F), central (C) and back (B) vowels. The subjects are Davide (D, above) and Caterina (C, below).

The χ^2 statistical test was applied to each single column of these tables, to assess whether the frequency of the determined combinations would give a significantly different result (*p-.05) from the frequency of the expected ones. The expected

frequencies for each vowel class were derived from the overall frequency of the appropriate vowel classes within each of the two samples. As for Davide, even though the general tendency is towards the predicted associations, the only significant associations are the two regarding the co-occurrence of labial consonants and central vowels, and apical consonants and front vowels in the final stage. As for Caterina, all the predicted associations are significant for the first months, whereas the two final months evidence a non-random distribution different from the predicted one, with an exception for the apical consonants-front vowels association (overall, the strongest association). In order to ascertain the patterns of association characterizing the target lexicon, and thus its influence, a new χ^2 analysis was carried out on the consonants and vowels of the lexical selection from [2]. The only significant result regarded the Dorsal consonant-vowel association, where dorsals associated were more with central vowels (1.438), and less with back vowels (0.358), than expected.

4. CONCLUSIONS

This description of some characteristics of the consonants' and vowels' development in 2 children along the stages of late babbling and early words production aims to demonstrate the utility of an interaction between the traditional phonetic method based on auditory classifications and the method based on acoustic analysis. The results were useful for tracing individual profiles of speech development, that in these two children was discovered to be more different at the starting point than at the final point, and revealed the increasing influence of the target language. Further, these results may offer some elements and suggestions to the researchers interested on cross-linguistic comparisons. Future researches will attempt to disentangle the relative contributions from the part of babbling and words.

REFERENCES

- [1] Vihman, M.M. 1996. *Phonological Development* 1996. Cambridge (MA): Blackwell Publ. Inc.
- [2] Caselli, M.C. and Casadio, P. 1995. *Il primo vocabolario del bambino*, Milano: FrancoAngeli.
- [3] Duckworth, M., Allen, G., Hardcastle, W., and Ball, M. 1990. Extensions to the International Phonetic Alphabet for the Transcription of Atypical Speech. *Clinical Ling. Phonetics*, 4, 273-284.
- [4] Oller, D. K. 1986. Metaphonology and Infant Vocalizations. In B. Lindblom & R.Zetterstrom (eds.), *Precursor of Early Speech*, Stockton Press, 21-35.
- [5] Kent R. D. and Murray A.D. 1982. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *J. Acoustical Soc. America*, 72, 353-365.
- [6] Allen, G.D., Pullum, G.K. and Ladusaw, W. 1991. *Phonetic Symbol Guide - Electronic Edition*, Purdue University & G.D. Allen.
- [7] Shriberg L.D. and G.L. Lof 1991. Reliability studies in broad and narrow phonetic transcription. *Clin. Ling. Phonetics*, 5, 225-279.
- [8] Zmarich C. and Lanni R. 1998. A phonetic and acoustic study of babbling in an Italian child. *Proceedings of ICSLP 98*, Sydney, vol. 6, 2103-2706.
- [9] Bortolini, U., Degan, C., Minnaja, C., Paccagnella, L. and Zilli, G. 1977. Statistics for a stochastic model of spoken Italian. *Proceedings of the Twelfth International Congress of Linguists*, Vienna, 580-586.
- [10] Batinti, A. 1993. *Il sistema fonologico dell'italiano*. Perugia: Guerra Edizioni.
- [11] Ferrero, F., Magno Caldognetto, E. and Cosi, P. 1996. Sui piani formantici acustici e uditivi delle vocali di uomo, donna e bambino. *Atti del XXIV Convegno Naz. dell'Ass. Italiana di Acustica*, Trento, 169-178.
- [12] Davis, B.L., MacNeilage, P.F. 1995. The Articulatory Basis of Babbling. *J. Speech Hearing Res.*, 38, 1199-1211.