Fundamental frequency in the speech of infants and children

Patricia Keating and Robert Buhr

Linguistics Department, Brown University, Providence, Rhode Island 02912
(Received 22 December 1976; revised 4 October 1977)

Fundamental frequency ranges of six normal infants and children from 33 to 169 weeks were determined by narrow-band spectrographic analysis. Fundamental frequency values ranged from 30 to 2500 Hz, well outside the values reported in previous studies of noncry utterances. The use of fry, modal, and high registers is also discussed.

PACS numbers: 43.70.Gr

Recent years have seen a resurgence of interest in speech development in infants and children. However, there have been few studies on the acquisition of prosodic features by infants and children. We are presently charting the development of normal prosodic features in a longitudinal study of several children. As a preliminary to this work, we wish to report on the fundamental frequencies used by these children.

While everyone would grant that young children have higher-pitched voices than older children or adults, it is difficult to find precise information about the range of fundamental frequency that younger children normally use. Peterson and Barney (1952), in their study of citation-form English vowels, found that children used fundamental frequencies between 250 and 275 Hz. Lieberman (1975) states that children's fundamental frequency ranges up to 500 Hz. Ages are unspecified in both sources. Fairbanks, Wiley, and Lassman (1949) and Fairbanks, Herbert, and Hammond (1949) reported that average fundamental frequency in reading was 294 Hz for 7-yr-old boys, 297 Hz for 8-yr-old boys, 273.2 Hz for 7-yr-old girls, and 286.5 Hz for 8-yr-old girls. Fundamental frequency has also been studied in infants. Fairbanks (1942) gives the monthly average f_0 of the cries of one infant over a 9-month period; average f_0 was 373 Hz at 1 month, 415 Hz at 2 months, 485 Hz at 3 months, and 585 Hz from 4 through 9 months. Sheppard and Lane (1968) recorded all vocalizations of two infants from birth through 5 months and give f_0 data for 95-s samples which thus include both cry and noncry utterances. Average f_0 was 419 Hz at birth, 398 Hz at 21 days, 416 Hz at 45 days, and 438 Hz for the duration of the study. Our data show that some children often use both higher and lower fundamental frequencies than have previously been mentioned.

Our corpus consists of tape recordings of six children playing with their mothers at home. The children were recorded using a Nagra 4.2 tape recorder and an AKG microphone. Ages at which samples were taken ranged from 33 weeks for the youngest child to 169 weeks for the oldest child. Two samples were chosen for analysis from each child's recorded corpus with the exception of child J.S. In choosing these samples, we wanted to include an early and late sample for each child, thereby providing useful longitudinal data, wherever possible. Where this was not possible, additional samples were chosen at the ages determined by the previous criterion, in order to compare the performance of different children at specific ages. These two criteria led us to choose samples at 33, 66–69, 109, 125–128, and 169 weeks. Only one sample is available for child J.S., due to the small size of the corpus for this child. Since this sample corresponds in age (67 weeks) to four of the other samples, it is included in the analysis.

For this study, an utterance is defined as a noncry vocalization. The end of the utterance was taken to be the point at which the child stopped for breath. It should be noted that since the children were in various stages of language acquisition, the nature of the utterances taken for analysis varied from one sample to another. For example, utterances in the babbling stage consist of V and CV forms, whereas older children produce both words and short sentences.

The first 50 utterances in each tape for which acoustic measurements could be made were selected for analysis, and spectrograms were made on a Kay Sona-Graph 6061B. The input was recorded at high speed (160–16 000 Hz), so that there was an effective narrow bandwidth of 90 Hz, and an effective wide bandwidth of 600 Hz. A 6076C scale magnifier, with the lower limit set at 000 (00.0%) and the upper limit set at 500 (50.0%), produced a display of 8000 Hz on the spectrograms. Fundamental frequency was measured at the fifth harmonic from the narrow-band spectrograms, except for f_0 values below 200 Hz, which were determined by counting striations per unit of time on the wide-band spectrograms. A comparison of a sample of measurements made by the two authors indicated an average remeasurement variability of 10 Hz with a greatest observed variation of 30 Hz.

I. FUNDAMENTAL FREQUENCY RANGES AND REGISTERS

Table I shows the ages at which samples were selected for each child, and the total range of f_0 for each sample. We have also noted the types of utterances which constitute the corpus. Note that some of the subjects evidence remarkable ranges of f_0 values, particularly J.S., with utterances as low as 30 Hz in f_0 to as high as 1500 Hz. We note in passing that for four of the five children for which two samples were taken, the high f_0 value has increased in the second sample. However, on the basis of our knowledge of the corpus, this result

 0001-4966/78/6302-0567$00.80

TABLE I. This table presents a breakdown of the children by sex, the ages at which samples were taken, the range of f_0 (Hz) for that week, and the type of utterances used at that week.

Child	Sex	Age (weeks)	Range of f_0 (Hz)	Type of utterance
L.S.	M	33	40–850	Babbling
		66	90–1150	Babbling
J.S.	M	67	250–2500	Babbling and words
F.R.	M	125	60–750	Sentences
		169	30–1100	Sentences
G.R.	M	69	60–850	Words
		128	120–1050	Sentences
J.B.	F	66	30–1500	Words
		109	50–1150	Sentences
R.C.	F	33	80–500	Babbling
		69	60–900	Babbling

seems to reflect an accident of the particular samples being presented.

Table II gives a more complete breakdown of the frequency ranges and vocal registers used by each child. Utterances have been divided into three registers: fry register, modal register, and high register.

The overall range of frequency in fry register was 30–250 Hz. Fry register is associated with "creaky voice" and low f_0, and was determined by a perceptual identification, with spectrographic corroboration. It should be noted that there is no infallible method of determining the presence of fry register on the basis of spectrographic evidence alone. However, fry register can usually be observed by noting widely spaced irregular striations on the wide-band spectrograms [see Fig. 1(a)]. Generally, fry register occurs with f_0 values below 200 Hz, although two of the six children (L.S. and R.C.) have f_0 values for fry above 200 Hz, ranging up to 250 Hz. Fry usually occurs at the beginning or end of the voiced portion of an utterance, although we have observed several utterances which consisted entirely of fry register. Many of the instances of fry were very brief, on the order of 100 ms or less. Note that there is considerable variation in the number of utterances containing fry among the six children.

Utterances containing modal register comprise the bulk of the samples for each child. The overall range of frequency in modal register was 150–700 Hz, outside previously reported values for f_0 in children's speech. Note that the f_0 values and percentages presented in the breakdown of modal register in Table II reflect the *peak* modal values for the utterances in the sample. Therefore, the lower values given under "Range" are not necessarily included in the breakdown of peak f_0 values, unless they also happen to be the peak f_0 value for a particular utterance. The breakdown of peak f_0 values presented in Table II is offered to indicate the frequency of occurrence of relatively high modal f_0 values in the utterances utilized for this sample. As can be seen from the table, a fairly high overall percentage of the utterances containing modal register evidence f_0 evidence peak values at or above 500 Hz.

In addition to these high modal values, we also note the occurrence of a separate high register. Perceptually, this register (which may be analogous to falsetto register in adults) is characterized by a high f_0 coupled with a thinner, "squeakier" quality than that associated with modal register. High register can generally be observed on spectrograms by noting a discontinuity in the f_0 contour on the narrow-band spectrogram [see Fig. 1(b)]. However, not all of the utterances classified as "high" evidenced such discontinuities. Some utterances have f_0 values entirely above 500 Hz with no discontinuity, and register decisions for these utterances were made on perceptual grounds, where possible. Note, for example, that register decisions for four utterances of G.R. at 128 weeks were not possible on either spectrographic or perceptual grounds. Furthermore, a few utterances seem to contain use of both modal and high registers without any noticeable discontinuity on the spectrogram. The older children apparently control their use of high register to the extent that they can produce a smooth change from one register to another [See Fig. 1(c)]. It can be noted that while all children employ high register, they do not do so as frequently as they use fry register in the production of noncry utterances.

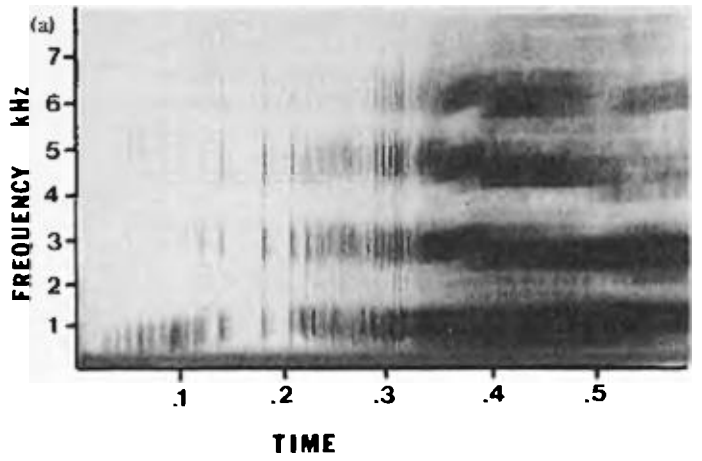
II. DISCUSSION

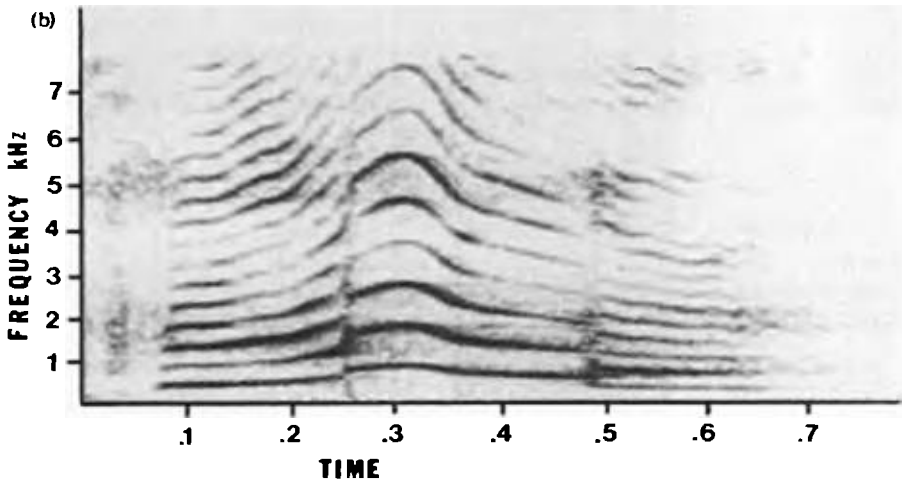
Although the corpus spans a period of 136 weeks, it is difficult to identify any significant longitudinal trends in terms of either the use of various registers or the overall range of f_0 within registers. We can only suggest that a more complete longitudinal study, spanning several more years of vocal production and involving more subjects, is necessary before any such trends become apparent. It is surprising, however, that even in this limited sample period, there seems to be no change in the overall f_0 range within modal register used by these children.

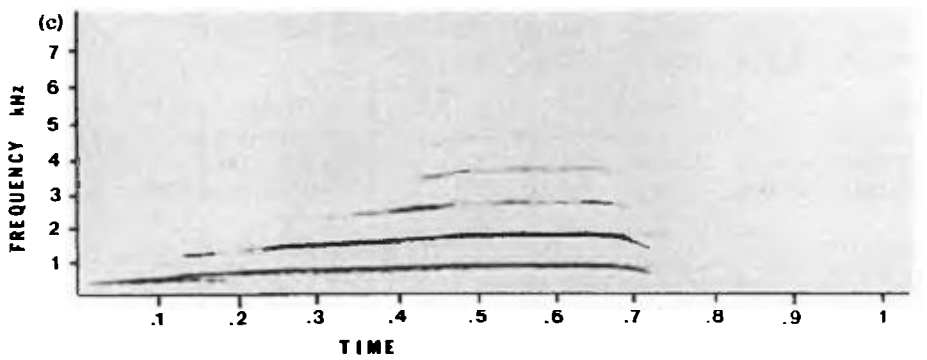
This set of data, however, is of interest in terms of previous proposals regarding the use of various vocal registers. Van den Berg (1968a, 1968b) discusses the physiological basis for five different registers involved in sound production in humans. Van den Berg notes that the highest of these five registers, the "whistle" register, does not seem to involve laryngeal vibration; it is not clear at present to what extent this register is included in what we have termed "high" register. Three other registers correspond to the three registers we have utilized here: strohbass (fry), chest (modal), and falsetto (high). A final register, the "mid" register, is apparently used between chest (modal) and falsetto (high) registers, in order to provide a smooth transition between these two registers. It is possible that the utterances we have noted which contain no discontinuity between modal and high registers may be examples of the use of mid register. Since physiological data on the laryngeal activity of these children is not available, we cannot determine the extent to which whistle and mid registers occur in our corpus.

TABLE II. This table presents an analysis of utterances at a given week in terms of fry, modal, and high registers. The table is arranged chronologically. The number of utterances (and the percentages given) refers to the number of utterances containing that particular register. Since an utterance may contain more than one register, these figures will not necessarily sum to equal values. Occurrances of modal register are presented in terms of frequency ranges. Note that only the highest f_0 value in each utterance is included in this table; consequently, values at the low end of the ranges presented do not necessarily appear unless they were the high f_0 value for that particular utterance.

	Fry	Modal						High
L. S.	33 weeks	No.: 50 (100%)						No.: 3(6%)
	No.: 23 (46%)	range: 175–490 Hz						range: 380–850 Hz
	range: 40–240 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 640 Hz
	mean low: 128 Hz	ø	31(62%)	19(38%)	ø	ø	ø	
R.C.	33 weeks	No.: 48(96%)						None
	No.: 7(14%)	range: 200–500 Hz						
	range: 80–250 Hz	200–99	300–99	400–99	500–99	600–99	700+	
	mean low: 147 Hz	1(2%)	19(38%)	25(50%)	3(6%)	ø	ø	
(two utterances entirely fry)								
L.S.	66 weeks	No.: 45(90%)						No.: 5(10%)
	No.: 9(18%)	range: 250–575 Hz						range: 550–1150 Hz
	range: 90–240 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean: 777 Hz
	mean low: 145 Hz	ø	33(66%)	11(22%)	1(2%)	ø	ø	(three utterances all high)
(one fry value could not be determined)								
(two utterances all fry)								
J.B.	66 weeks	No.: 49(98%)						No.: 6(12%)
	No.: 9(18%)	range: 200–700 Hz						range: 500–1500 Hz
	range: 30–200 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 908 Hz
	mean low: 95 Hz	ø	19(38%)	21(42%)	5(10%)	3(6%)	1(2%)	(one utterance entirely high register)
J.S.	67 weeks	No.: 45(90%)						No.: 18(36%)
	No.: ø	range: 250–700 Hz						range: 500–2500 Hz
		200–99	300–99	400–99	500–99	600–99	700+	mean high: 984 Hz
		ø	1(2%)	15(30%)	16(32%)	11(22%)	2(4%)	
R.C.	69 weeks	No.: 50(100%)						No.: 2(4%)
	No.: 8(16%)	range: 200–650 Hz						range: 700–900 Hz
	range: 60–220 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 825 Hz
	mean low: 158 Hz	2(4%)	14(28%)	16(32%)	14(28%)	4(8%)	ø	
G.R.	69 weeks	No.: 49(98%)						No.: 7(14%)
	No: 7(14%)	range: 200–650 Hz						range: 400–850 Hz
	range: 60–190 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 721 Hz
	mean low: 116 Hz	ø	7(14%)	19(38%)	14(28%)	9(18%)	ø	(one utterance all high)
J.B.	109 weeks	No.: 45(90%)						No.: 8(16%)
	No.: 18(36%)	range: 200–600 Hz						range: 500–1150 Hz
	range: 50–200 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 920 Hz
	mean low: 121 Hz	3(6%)	25(50%)	14(28%)	2(4%)	2(4%)	ø	(one utterance all high)
(four utterances all high)								
F.R.	125 weeks	No.: 50(100%)						No.: 2(4%)
	No.: 5(10%)	range: 150–650 Hz						range: 600–750 Hz
	range: 60–190 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 725 Hz
	mean low: 93 Hz	1(2%)	16(32%)	26(52%)	5(10%)	2(4%)	ø	
G.R.	128 weeks	No.: 46 (92%)						No.: 8(16%)
	No.: 1(2%)	range: 200–650 Hz						range: 450–1050 Hz
	range: 120–200 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 825 Hz
	mean low: 120 Hz	3(6%)	14(28%)	12(24%)	11(22%)	6(12%)	ø	
		(four utterances not identifiable as either modal or high register)						
F.R.	169 weeks	No.: 49(98%)						No.: 3(6%)
	No.: 9(18%)	range: 200–700 Hz						range: 800–1100 Hz
	range: 30–160 Hz	200–99	300–99	400–99	500–99	600–99	700+	mean high: 1016 Hz
	mean low: 110 Hz	4(8%)	22(44%)	11(22%)	6(12%)	5(10%)	1(2%)	
(one utterance all fry)								







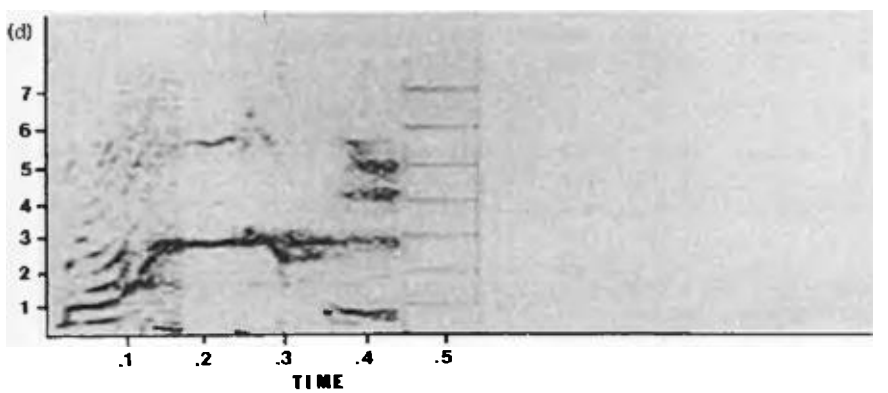


FIG. 1. (a) Wide-band spectrogram of /ye/ by L. S. at 33 weeks. (b) Narrow-band spectrogram of /kar/ by J. S. at 67 weeks. (c) Narrow-band spectrogram of "No." by F. R. at 169 weeks. (d) Narrow-band spectrogram of an utterance by J. S. at 67 weeks with an apparently very high fundamental frequency.

Hollein and his co-workers discuss the use of three registers (fry, modal, and falsetto) in the speech of adults. Hollein *et al.* (1966) proposed that the use of fry register in adult speech was not a pathological condition as had previously been suggested, but rather was a normal occurrence in voice production. The data we are reporting suggest that the use of fry register is a normal occurrence in spontaneous voice production in infants and children as well. In a subsequent paper, Hollein and Michel (1968) presented data on f_0 ranges in fry, modal, and falsetto register for a group of male and female adult speakers. Although comparisons of the f_0 ranges within registers between adults and the infants reported here is inappropriate, since Hollein's data did not rely upon spontaneous vocal production, it is interesting to note that both adults and children evidence some f_0 overlap between modal and high register. Further comparison of the use of vocal register in adults and children must await a detailed study of register use in the spontaneous speech of adults.

Although we are reporting only on f_0 values of noncry utterances, we should like to point out the similarity between high register in our corpus and hyperphonation, as discussed by Truby and Lind (1965). Hyperphonation was used by these authors to describe the very high f_0 values found in the cry utterances of newborn infants. These utterances are characterized by a marked shift of the f_0 contour on narrow-band spectrograms, as are the utterances of older children. These authors do not report any use of phonation by newborn infants analogous to the use of fry register. Truby and Lind also suggest that the suddenness of the f_0 shifts found in newborn infant cry is a phenomenon unique to infants. However, our data indicate that this phenomenon is a common occurrence in children as old as 169 weeks. In fact, such sudden shifts in f_0 may be a characteristic of register shifts in general, especially between modal and high registers.

The occurrence of wide f_0 ranges in normal infants and children reported here, especially the common use of high f_0 values, raises some questions regarding the use of high f_0 to diagnose various pathological conditions. Stark and Nathanson (1975) reported a greater incidence of high f_0 in the cries of an infant who later died of Sudden Infant Death Syndrome than in those of normal infants. Jones (1965) suggested that the cries of young deaf children of ages 7–48 months are higher and more variable than those of normal children. Ostwald, Phibbs, and Fox (1968) proposed that cries over 650 Hz indicate abnormality on the basis of five normal, five questionably impaired, and three abnormal infants recorded once each at ages ranging from 1 day to 7 months. Sample spectrograms of "abnormal" cries show the sudden pitch shifts we have described. Even though our data reflects f_0 values of noncry utterances, there is considerable variation in the use of high f_0 among the children in this study. This finding indicates that the range of f_0 that can be considered "normal" is higher than previously reported, and the occurrence of high f_0 in itself may not be indicative of a pathological condition.

As a final comment, we would like to note the occurrence of a noncry utterance [Fig. 1(d)] listed in Table II for child J.S. as having a f_0 of 1700 Hz, but which may have a peak f_0 of 3000 Hz. This spectrogram is unusual in that it contains several apparent harmonics which do not seem to be multiples of the f_0, if the f_0 is indeed 1700 Hz. There is the further problem of an additional set of harmonics at multiples of 700 Hz at the beginning of the high-frequency portion of the utterance, which seem to bear no relation to the rest of the utterance. This is the only utterance of this type that we encountered in our corpus.

In conclusion, the data for this sample of six children shows that both very low and very high fundamental frequencies are common at all stages of language acquisition. Clearly, vocalizations of normal infants and young children, such as those described here, should be considered in accounts of the possible output of the human larynx.

ACKNOWLEDGMENTS

This work was supported by National Institute of Child Health and Human Development grant #5 R01 HD 09197 to Brown University. We wish to thank Philip Lieberman for his advice and assistance, and Barbara Moslin for the use of her taped material.

Fairbanks, G. **(1942).** "An acoustical study of the pitch of infant hunger wails," Child Dev. **13,** 227–232.

Fairbanks, G., Herbert, E. L., and Hammond, J. M. **(1949).** Child Dev. **20,** 71–79.

Fairbanks, G., Wiley, J. H., and Lassman, F. M. **(1949).** Child Dev. **20,** 62–69.

Hollein, H., and Michel, J. F. **(1968).** "Vocal Fry as a Phonational Register," J. Speech Hear. Res. **11,** 600–604.

Hollein, H., Moore, P., Wendahl, R. W., and Michel, J. F. **(1966).** "On the Nature of Vocal Fry," J. Speech Hear. Res. **9,** 245–247.

Jones, M. C. **(1965).** "An Investigation of Certain Acoustic Parameters of the Crying Vocalizations of Young Deaf Children" (Univ. Microfilms, Ann Arbor, MI).

Lieberman, P. **(1975).** *On the origins of language* (Macmillan, New York), Chap. 5, p. 44.

Ostwald, R. F., Phibbs, R., and Fox, S. **(1968).** "Diagnostic Use of Infant Cry," Biol. Neonat. **13,** 68–82.

Peterson, G. E., and Barney, H. L. **(1952).** J. Acoust. Soc. Am. **24,** 175–184.

Sheppard, W. C., and Lane, H. L. **(1968).** "Development of the Prosodic Features of Infant Vocalizing," J. Speech Hear. Res. **11,** 94–108.

Truby, H. M., and Lind, J. **(1965).** "Cry Sounds of the Newborn Infant," in *Newborn Infant Cry*, edited by J. Lind (Almquist and Wiksell, Uppsala), Chap. 1.

Stark, R. E., and Nathanson, S. **(1975).** "Unusual Features of Cry in an Infant Dying Suddenly and Unexpectedly," in *Development of Upper Respiratory Anatomy and Function: Implications for Sudden Infant Death Syndrome*, edited by J. F. Bosma and J. Showacre (U. S. GPO, Washington, DC), Chap. 17, p. 237.

van den Berg, J. **(1968a).** "Sound Production in Isolated Human Larynges," Ann. N. Y. Acad. Sci. **155,** 18–27.

van den Berg, J. **(1968b).** "Register Problems," Ann. N. Y. Acad. Sci. **155,** 129–134.