# Fundamental frequency of infants' and parents' utterances in longitudinal recordings

Shigeaki Amano, Tomohiro Nakatani, and Tadahisa Kondo
*NTT Communication Science Laboratories, NTT Corporation, 2-4 Hikari-dai, Seika-cho, Souraku-gun, Kyoto 6190237, Japan*

(Received 27 February 2005; revised 24 September 2005; accepted 5 December 2005)

The fundamental frequencies ($F0$) of daily life utterances of Japanese infants and their parents from the infant's birth until about 5 years of age were longitudinally analyzed. The analysis revealed that an infant's $F0$ mean decreases as a function of month of age. It also showed that within- and between-utterance variability in infant $F0$ is different before and after the onset of two-word utterances, probably reflecting the difference between linguistic and nonlinguistic utterances. Parents' $F0$ mean is high in infant-directed speech (IDS) before the onset of two-word utterances, but it gradually decreases and reaches almost the same value as in adult-directed speech after the onset of two-word utterances. The between-utterance variability of parents' $F0$ in IDS is large before the onset of two-word utterances and it subsequently becomes smaller. It is suggested that these changes of parents' $F0$ are closely related to the feasibility of communication between infants and parents. © *2006 Acoustical Society of America.* [DOI: 10.1121/1.2161443]

PACS number(s): 43.70.Ep, 43.70.Gr [AL] Pages: 1636–1647

## I. INTRODUCTION

### A. Infant *F0*

An infant's speech development can be studied using several approaches. One involves analyzing infant utterances acoustically to reveal the developmental changes that occur with age. Among the acoustic characteristics of utterances, we focused on the fundamental frequency ($F0$). Several variables have been used in previous research to represent the $F0$ change that occurs during speech development. These variables include mean (Robb and Saxman, 1985), mode (Laufer and Horii, 1977), median (Laufer and Horii, 1977), range (Keating and Buhr, 1978), maximum (Kent and Bauer, 1985), minimum (Fairbanks, 1942), and standard deviation (Robb and Saxman, 1985). The number of percentage of occurrences has also been used to determine the developmental changes in, for example, rising, falling, and flat $F0$ patterns (Robb, Saxman, and Grant, 1989). Of these variables, we used the mean and standard deviation of $F0$ in this study, because the mean has frequently been used in previous studies and the standard deviation would be a better and more robust indicator than the range to represent $F0$ variability (cf. Masataka, 2002).

Kent (1976) reviewed the developmental change in the $F0$ mean from birth to adulthood by collecting the results of cross-sectional studies. The infant $F0$ is high at birth and gradually becomes lower as the infant grows up. The rate at which $F0$ decreases is higher before 3 years of age than after it. A difference appears in the $F0$ of boys and girls after the adolescent voice change of boys at 11 or 12 years of age. As Kent (1976) admitted, it was a tentative view of $F0$ changes with the limited data available at that time. The actual $F0$ mean and its developmental change might differ from that presented in his review.

More recently, some studies have reported the precise $F0$ mean and its developmental change. For example, Bennet (1983) reported the developmental change in $F0$ in the 7–11 age range based on 3 years of longitudinal recordings. She reported that the $F0$ mean decreases about 12 Hz per 12 months. Hollien, Green, and Massey (1994) studied the $F0$ mean of boys in their early teens around the adolescent period with 5 years of longitudinal recordings. According to their study, the $F0$ mean for 34 boys was 233.1 Hz before adolescence, 173.8 Hz during adolescence, and 121.6 Hz after adolescence. This drop in the $F0$ mean was also shown by the Lee, Potamianos, and Narayanan (1999) cross-sectional study for 436 children in the 5–17 age range. They reported that pubertal $F0$ change starts about age 12 and ends about age 15 for male children. They also reported that $F0$ significantly decreases in the 7–12 age range for male and female children but not after age of 12 for female children.

The $F0$ mean was also reported for younger ages. For example, Robb, Saxman, and Grant (1989) longitudinally recorded the utterances of seven infants in the 8–26 month age range on 12 occasions over 13 months. They investigated the $F0$ of monosyllabic and bisyllabic utterances but they found no significant difference between months for either type of utterance. The $F0$ mean was 396 Hz for monosyllables and 399 Hz for bisyllables. McRoberts and Best (1997) studied the $F0$ mean of one infant of 3–17 months under three conditions, infant alone, infant with mother, and infant with father. The $F0$ was not dependent on the conditions but changed significantly with age. That is, there was a general tendency for the $F0$ mean to decrease month by month. It was 407 Hz at 7 months, 330 Hz at 15 months, and 326 Hz at 17 months.

However, longitudinal studies in infant and early childhood were conducted over much shorter periods than those in late childhood. The longest study undertaken with infants and in early childhood is the Robb, Saxman, and Grant (1989) study consisting of observations over 1.5 years. As

 0001-4966/2006/119(3)/1636/12/$22.50 

regards late childhood, a 3-year research project was undertaken by Bennet (1983) and a 5-year study was undertaken by Hollien, Green, and Massey (1994). As a consequence, few longitudinal data have been provided for *F*0 from birth to about 5 or 6 years old. If data for this range are provided and combined with results such as those of the Lee, Potamianos, and Narayanan (1999) study in 5–17 age range, we can obtain a consecutive view of the developmental change in *F*0 from birth to young adulthood.

As regards *F*0 variability, previous studies used two kinds of standard deviation (SD). One is the within-utterance SD of *F*0. This is an average of the SD of *F*0 in each utterance. This variable represents *F*0 variability within one utterance. The other is the between-utterance SD of *F*0. This is the SD of the *F*0 mean. This variable represents the variability of the *F*0 mean among utterances.

Prescott (1975) used these two kinds of SD to analyze the *F*0 of infants. He found that the within-utterance SD of *F*0 observed longitudinally increases from 28 Hz at 0 months to 53 Hz at 6–9 months for the same four infants. Moreover, cross-sectional observations revealed that the within-utterance SD of *F*0 also increases from 30 Hz at 1 month to 53 Hz at 6–8 months with 10 infants for each month. On the other hand, he found no significant increase or decrease in the between-utterance SD of *F*0 in either case. Shepard and Lane (1968) reported that the within- and between-utterance variability of *F*0 were both unchanged in a longitudinal observation of two infants from 0–4.5 months. Robb, Saxman, and Grant (1989) showed that the between-utterance SD of *F*0 did not change during an approximately 1-year longitudinal observation of infants in the 8–26 month age range. Robb and Saxman (1985) conducted cross-sectional research with 14 infants. They reported that the between-utterance SD in *F*0 is much higher at 11–16 months than at 17–25 months.

These studies offer some information about the developmental change in *F*0 variability. However, they do not provide sufficient data to depict the developmental change in *F*0 variability from birth to early childhood because the observation period is too short. In addition, they provided contradictory results in certain cases. For example, Prescott (1975) reported an increase in the within-utterance *F*0 variability but Shepard and Lane (1968) did not. These problems should be resolved by investigating *F*0 data in early childhood.

### B. Parents *F*0

We also investigate another developmental aspect of *F*0, namely "infant-directed speech (IDS)," which is spoken to an infant by a caretaker such as a parent. IDS is sometimes called motherese (Shute and Wheldall, 1989), maternal speech (Penman *et al.*, 1983), infant directed talk (Werker and McLeod, 1989), or baby talk (Singh, Morgan, and Best, 2002). IDS is different from "adult-directed speech (ADS)" in some respects. For example, it has a higher *F*0 (McRoberts and Best, 1997), a wider *F*0 range (Fernald *et al.*, 1989), a shorter utterance duration (Fernald and Simon, 1984), longer pauses (Grieser and Kuhl, 1988), and an exaggerated formant frequency (Kuhl *et al.*, 1997). IDS with these characteristics attracts an infant more than ADS (Cooper *et al.*, 1997; Werker and McLeod, 1989). And this preference probably helps an infant to acquire its native language (Hirsh-Pasek *et al.*, 1987).

Within the characteristics of IDS, *F*0 seems to be an important factor in terms of an infant's preference. The work of Spence and Freeman (1996) supports this idea. They compared an infant's preference as regards low-pass filtered IDS and whispered IDS. The low-pass filtered IDS contains *F*0 information but the whispered IDS does not. Infants preferred the low-pass filtered IDS to the whispered IDS, suggesting that *F*0 is an important factor in IDS. Additional support for this idea was obtained by Fernald and Kuhl (1987). They reported that infants prefer synthesized IDS with *F*0 modulation to that with amplitude modulation. This suggests that *F*0 in IDS is an outstanding feature for infants.

Many of the characteristics of IDS are said to be found in almost every language, and it is suggested that these characteristics might be universal (e.g., Grieser and Kuhl, 1988). But there are some exceptions. For instance, Ratner and Pye (1984) reported that a higher *F*0 in IDS is not observed in Quiche Mayan, which is spoken in the western highlands of Guatemala. In addition, the extent of *F*0 change seems to differ across languages. For example, mothers of American-English native speakers tend to use a much more exaggerated *F*0 in IDS than mothers who speak other languages (e.g., Fernald *et al.*, 1989; Papousek, Papousek, and Symmes, 1991). Similarly, mothers of Australian-English native speakers tend to use a more exaggerated *F*0 in IDS than mothers of Thai native speakers (Kitamura *et al.*, 2002). Shute and Wheldall (1989) reported that mothers of British-English native speakers use a high *F*0 in IDS, but that the increase in *F*0 is less than for mothers of American-English native speakers.

As for Japanese, IDS seems to have the same characteristics as other languages. For example, Niwano and Sugai (2003) reported that mother and father use a higher *F*0 in terms of mean, maximum, minimum, and range in IDS to infants at 3, 5, and 7 months of age. Masataka (1992) showed that there is a higher mean and a wider range of *F*0 in IDS than in ADS especially when infants do not respond to the IDS. Masataka (2002) reported that the *F*0 of IDS to an 11-month-old infant is higher than that of ADS. He also found that the *F*0 increase in IDS is larger in a female with children (i.e., mother) than a female without children (i.e., nonmother).

Several factors affect the *F*0 in IDS. One relates to people's experience with their own babies as shown by Masataka (2002). Other factors are, for example, an infant's facial expressions (Reissland, Shepherd, and Cowie, 2002) and the cradling side (Reissland, 2000). Although there may be many other factors including language difference, previous studies almost always provide evidence that *F*0 of IDS is high.

However, it is not clear how *F*0 in IDS changes with infant age. There have only been a few studies on *F*0 changes in IDS with age (e.g., Remick, 1976; Warren-Leubecker and Bohannon, 1984). For example, Remick (1976) studied mothers' *F*0 in IDS to their infants whose ages ranged between 16 and 30 months with an approxi-

mately 2-month interval. Although she showed the *F*0 change in IDS as a function of an infant's age, the observed *F*0 change might be an artifact of individual differences, because she used a cross-sectional method. Recently, Kitamura *et al.* (2002) observed the *F*0 change in IDS with a longitudinal method for infants from 0 to 12 months old. They showed that the *F*0 in IDS is higher than in ADS and that it changes with infant age. However, their observation period was less than 1 year. It is still unknown how *F*0 in IDS changes over longer periods. A greater change may be found if the *F*0 in IDS is observed with infants for more than 1 year. In particular, long term observation would enable us to discover when the special characteristics of *F*0 in IDS disappear and become the same as in ADS. Little attention has been paid to this point as regards Japanese or other languages. It was simply assumed that IDS has some special characteristics in relation to a prevabal infant (e.g., Sachs, 1977).

It is likely that *F*0 changes in IDS are closely related to the speech development of infants, and that a large *F*0 change occurs in IDS when the infant starts to use "language." We assume that this change would occur at the onset of two-word utterances, because then parents can communicate more easily, and probably more normally, with their infant using "language." This idea is supported by the fact that Japanese mother-child conversation style becomes adult-like around the two-word utterance period (Kajikawa, Amano, and Kondo, 2004). The onset of one-word utterances is another possible change point. However, we assume that parents have a lot of difficulty communicating with their infant at the onset of one-word utterances, and would not speak to the infant in the same way as to adults. Based on this idea, we analyzed the *F*0 in IDS before and after the onset of two-word utterances. The mean and standard deviation were used for the analysis as with infant *F*0. The infant's *F*0 was also analyzed before and after the onset of two-word utterances, because this supplies additional information about the *F*0 change in infant utterances, which might relate to the *F*0 change in the IDS of parents.

In summary, we try to clarify the developmental changes that take place in the *F*0 mean and its variability in infants' and parents' utterances before and after the onset of two-word utterances. It would provide systematic information on *F*0 development with age from 0 to 5 years old, which has not been well studied in previous research.

## II. *F*0 ANALYSIS

To observe the developmental change in *F*0, we conducted two *F*0 analyses of the utterances of an infant and its parents. The first was intended to show that the *F*0 of infant utterances decreases month by month. It was also intended to show the developmental change in the *F*0 variability of infant utterances in terms of between- and within-utterance SD. More specifically, we compared these values before and after the onset of two-word utterances.

The second *F*0 analysis concerned the fact that Japanese parents' *F*0 is different in IDS and ADS. We intended to show that the *F*0 mean in IDS is high before the onset of an infant's two-word utterances and it has almost the same value as ADS after it. We also intended to show developmental change in *F*0 variability in IDS with regard to between- and within-utterance SD. More specifically, we compared these values before and after the onset of two-word utterances and in ADS.

### A. Infant *F*0

#### 1. Data

An infant speech database (Amano, Kato, and Kondo, 2002) was used as the data population. The database contains speech files (16-bit quantization and a 16-kHz sampling frequency) of natural utterances of five pairs of Japanese infants and parents during daily life at home. The utterances were digitally recorded approximately monthly for about 1 hour from birth to 5 years old. Infants B, C, and D were selected from the infant speech database, because the number of registered utterances was larger for these infants than the others.

All the infants were born and raised in Tokyo or its suburbs. They were all of normal height and weight at birth, and had no symptoms of disorder with respect to speech perception or speech production during the recording period. Although Infants C and D are siblings, we regarded their data as independent because their recordings were conducted separately.

The utterances that satisfied the following conditions were extracted from the database and used for analysis:

(i) The background noise should be low so that utterances are clear.
(ii) There should be no overlap with other utterances so that *F*0 value is clear.
(iii) Utterances with certain particular characteristics must be excluded in order to focus on normal utterances. Such characteristics include hiccupping, coughing, sneezing, yawning, singing, reading aloud, and number counting.

Table I shows the months during which infant utterances were extracted. The mean value of the number of extracted utterances per month was 432 (SD=312) for Infant B, 406 (SD=173) for Infant C, and 486 (SD=243) for Infant D.

#### 2. Procedure

A voiced/unvoiced part was specified in the extracted utterances by using the dominance spectrum method, called "DominAnce Spectrum based Harmonics extraction (DASH)" (Nakatani, Irino, and Zolfaghari, 2003; Nakatani and Irino, 2004). *F*0 was estimated for the voiced part by using the "Ripple Enhanced Power Spectrum (REPS)" method (Nakatani and Irino, 2002, 2004). A 42 ms Hanning window with 1 ms window shift was used for both methods.

Although these methods perform very well in terms of voiced/unvoiced (V/UV) classification and *F*0 estimation, there are occasional errors. To reduce the errors as much as possible, trained operators checked and corrected the *F*0 values and V/UV labels by investigating the superimposed *F*0 and V/UV information segment by segment on a spectrogram.

TABLE I. Months containing infant's utterances in infant speech database.

| Infant ID | Month | Number of months |
|---|---|---|
| B | 0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36, 37, 39, 40, 42, 43, 44, 45, 48, 51, 52, 53 | 45 |
| C | 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 53, 55, 56, 59, 60 | 52 |
| D | 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 43, 44, 45, 46, 47, 48, 49, 51, 52, 55, 56, 59 | 53 |

To exclude unusual $F0$ values, we used those between 20 and 1500 Hz. Although these lower and higher limits were decided empirically, they are appropriate for covering the $F0$ range of an infant utterance. The suitability of the limits is supported by, for example, Robb and Saxman (1985). They reported that the $F0$ mean was 357 Hz, the SD was 105 Hz, and the range of the mean was 164–1366 Hz for the utterances of 14 infants between the ages of 11 and 25 months.

$F0$ mean, SD (within) and SD (between) were calculated from the $F0$ value for the voiced segment. $F0$ mean is the average of the mean $F0$ value of each utterance. SD (within) is the average of standard deviation of the $F0$ value of each utterance. SD (between) is the standard deviation of the $F0$ mean value among utterances. SD (within) represents the $F0$ variability within one utterance, whereas SD (between) represents the variability of the $F0$ mean value among utterances.

To examine the developmental change in infant' $F0$, regression analysis was conducted for the $F0$ mean as a dependent variable ($y$) with infant's age in months as an independent variable ($x$). The regression method was the best linear unbiased estimation.

We undertook analyses of variance for the two SDs to compare their values before and after the onset of two-word utterances. The onset of two-word utterances was decided by two Japanese adults by listening to recordings of infants' utterances. The definition of a two-word utterance in this study is an utterance that contains two content words whose pronunciation is clear enough to be recognized by both Japanese adults.

### 3. Results and discussion

*$F0$ mean in infant utterances:* Figure 1 shows the $F0$ mean of infant utterances as a function of month. The regressions of the $F0$ mean were $y=368-0.65x$ [$r^2=0.109$, $F(1,43)=5.27$, $p<0.05$] for Infant B, $y=373-1.53x$ [$r^2=0.687$, $F(1,50)=109.67$, $p<0.001$] for Infant C, and $y=384-1.11x$ [$r^2=0.555$, $F(1,51)=63.59$, $p<0.001$] for Infant D. The coefficient of $x$ of the regression was significant for all infants. [$t(43)=-2.30$, $p<0.05$, two-tailed, for Infant B; $t(50)=-10.47$, $p<0.001$, two-tailed, for Infant C; $t(51)=-7.97$, $p<0.001$, two-tailed, for Infant D]. When the data was pooled for all the infants, the regression of the $F0$ mean was $y=385-1.33x$ [$r^2=0.400$; $F(1,148)=98.75$, $p<0.001$]. The coefficient of $x$ of the regression was significant [$t(148)=-9.94$, $p<0.001$, two-tailed].

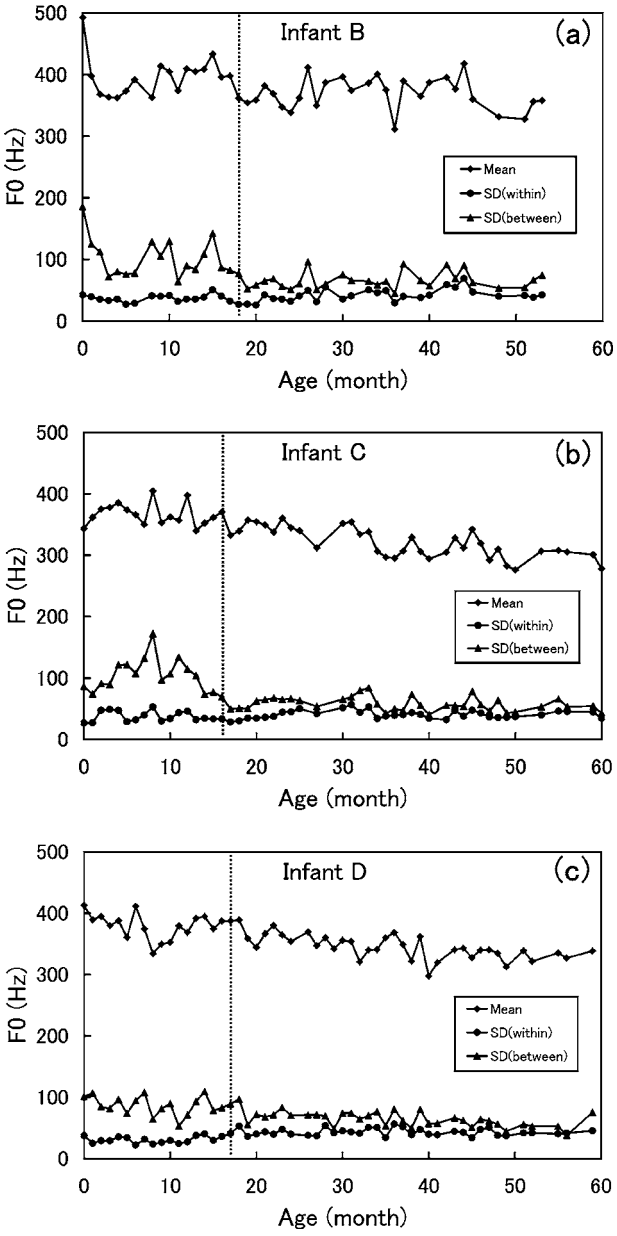


FIG. 1. $F0$ mean and SD of infant utterances as a function of the infant's age in months. The vertical dotted line represents the onset of two-word utterances for each infant.

These negative and significant coefficients indicate that $F0$ has a tendency to decrease month by month both for an individual infant and infants as a whole. This tendency is consistent with the results of previous studies (e.g., Bennet, 1983; Eguchi and Hirsh, 1969; Kent and Murray, 1982; McRoberts and Best, 1997). In addition, the rate of $F0$ de-

TABLE II. SD (between) and SD (within) of *F*0 of infant utterances before/after onset of two-word utterances.

| | SD (between) (Hz) | | | | SD (within) (Hz) | | | |
|---|---|---|---|---|---|---|---|---|
| | Infant ID | | | | Infant ID | | | |
| | B | C | D | Mean | B | C | D | Mean |
| Before | 103.9 | 106.6 | 86.5 | 98.6 | 37.2 | 38.1 | 30.7 | 35.2 |
| After | 66.1 | 58.3 | 65.2 | 63.0 | 41.9 | 40.6 | 43.3 | 42.0 |

crease is similar to that reported in some previous studies. From the regression of the pooled data in our study, we obtain an *F*0 decrease rate of 16 Hz per 12 months for infants of 0–5 years old. This *F*0 decrease rate is close to the result reported by Bennet (1983). She showed that the *F*0 decrease rate is about 12 Hz per 12 months for children in the 7–11 year age range.

However, some previous studies have reported different *F*0 decrease rates. For example, when we calculate the *F*0 decrease rate from the results of Eguchi and Hirsh (1969), it is 8.9 Hz per 12 months for 3–6 year-old children. This is much smaller than our result. In addition, their data indicates that the decrease rate is almost zero at age 7–9 years. Although this difference might be an artifact of the cross-sectional method that Eguchi and Hirsh (1969) used, it is possible that the *F*0 decrease rate changes with age in childhood. At least for boys, the *F*0 value drops very sharply during the period of adolescent voice change between about 12 and 15 years of age (Hollien, Green, and Massey, 1994; Lee, Potamianos, and Narayanan, 1999). However, we currently have insufficient evidence about the change in the *F*0 decrease rate prior to the period of adolescent voice change. We may only say from our data that *F*0 has a general tendency to decrease month by month and that the decrease rate is constant from 0–5 years of age.

Assuming that *F*0 is about 300–400 Hz for infants, the *F*0 decrease rate per 12 months is about 3.0%–4.0% in Bennet's (1983) study, about 2.2%–3.0% in the Eguchi and Hirsh (1969) study, and about 1.9%–6.1% in the current study. Because these percentages are very small, it would be difficult to detect the tendency of the *F*0 decrease when speech samples are collected over a short period. This might be one reason why some previous studies failed to find any significant tendency in the *F*0 decrease. For example, in Prescott's (1975) study, no consistent increase or decrease tendency was observed for *F*0 mean between 0 months and 6–9 months either in a longitudinal same-infant group or in a cross-sectional different-infant group. Robb, Saxman, and Grant (1989) conducted longitudinal recordings 12 times over 13 months with seven children 8–26 months of age, but they did not find any significant *F*0 difference among months. Robb and Saxman (1985) also failed to find any significant *F*0 decrease when using the cross-sectional method with 14 infants 11–25 months of age. The observation period was less than 1.5 years in all these studies. It is suggested that 2 or more years of observation would be necessary to obtain a significant tendency for the *F*0 decrease.

As regards *F*0 changes during the first 12 months, Shepard and Lane (1968) reported that the *F*0 mean decreased between 0 and about 1 month and then increased and became stable at 2–4.5 months. On the other hand, Fairbanks (1942) showed that the *F*0 mean increases at 0–4 months and then becomes almost constant at 5–9 months. Kent and Murray (1982) reported that the *F*0 mean of infants is 445 Hz at 3 months, 450 Hz at 6 months, and 415 Hz at 9 months. These short-period *F*0 changes are not confirmed by our *F*0 analysis in this study, because the *F*0 mean fluctuates from month to month and its pattern differs among infants (Fig. 1). For example, the *F*0 mean decreases at 0–4 months for Infants B and D, but it increases for Infant C. Such individual differences were also observed by Laufer and Horii (1977) in longitudinal *F*0 observations of four infants over the first 24 weeks. As they pointed out, individual *F*0 fluctuation depends on many factors including infant's activity level, utterance skill, and physical and social environment. These factors may differ among studies. Therefore, it would be difficult to verify the true *F*0 change over a short period. To reveal the *F*0 change over a short period, further research is necessary with precisely controlled conditions as regards the infant and his/her environment.

*F*0 *variability in infant utterances*: Figure 1 shows SD (between) and SD (within) of *F*0 in infant utterances as a function of month. Table II shows the averaged SD (between) and SD (within) before and after the onset of two-word utterances. We performed an analysis of variance with one factor of period with two levels (before vs after the onset of two-word utterances). An infant factor was included in a residual in the analysis. The factor of period was significant for SD (between) $[F(1,146)=122.3,\ p<0.001]$. This indicates that SD (between) is larger before the onset of two-word utterances than after it. Another analysis of variance showed that the factor of period was also significant for SD (within) $[F(1,146)=26.8,\ p<0.001]$. This indicates that SD (within) is smaller before the onset of two-word utterances than after it.

These results mean that infant *F*0's are very different from each other between utterances before the onset of two-word utterances, but this difference is reduced after the onset of two-word utterances. On the other hand, infant *F*0 does not change very much within one utterance before the onset of two-word utterances, but it becomes more variable within one utterance after the onset of two-word utterances.

The developmental change in *F*0 variability probably reflects an infant's ability to speak a language. Infants do not speak a language when they are very young. During that

period, it is highly probable that infants do not have the ability to control *F*0 precisely. This inability to control *F*0 would mean that infants could not start an utterance with a particular *F*0. At the same time, they could not quickly increase or decrease the *F*0 within one utterance. These characteristics are reflected in the large between-utterance variability and the small within-utterance variability before the onset of two-word utterances. It might be said that infant *F*0 is physically driven with some randomness, rather than intentionally, before the onset of two-word utterances.

As infants grow up, they acquire the ability to control *F*0 precisely. This acquisition of *F*0 control is critical with respect to spoken language, because the *F*0 pattern conveys a message in communication. For example, interrogative and affirmative information is expressed by the *F*0 pattern. In addition, the *F*0 pattern covers part of the lexical information in Japanese (Kubozono, 1993). With precise *F*0 control, infants can intentionally start their utterances with a fairly constant *F*0. And they can quickly increase or decrease the *F*0 within one utterance to create a complex *F*0 pattern when communicating with other human beings, mainly with their parents. These characteristics are reflected in the decrease in between-utterance variability and the increase in within-utterance variability after the onset of two-word utterances.

Robb and Saxman (1985) obtained very similar results to ours regarding SD (between), although they used the cross-sectional method with 14 infants. They showed that SD (between) is large (63–238 Hz) at 11–16 months and it becomes small (53–84 Hz) at 17–25 months. Because the onset of two-word utterances corresponds to 17–19 months in our study, their results agree with ours in that SD (between) is large before the onset of two-word utterances and subsequently becomes small. In addition, our SD (between) values before and after the onset of two-word utterances are in a similar range to theirs. These correspondences support our claim that the between-utterance *F*0 variability is large before the onset of two-word utterances and subsequently small.

However, some studies appear to disagree with our result. For example, Robb, Saxman, and Grant (1989) showed that SD (between) does not change between 8 and 26 months of age. This means that the between-utterance *F*0 variability is constant before and after the onset of two-word utterances (about 17–19 months). However, their study is methodologically different from ours. They observed only one- and two-syllable utterances, whereas we observed almost all types of utterance including nonlinguistic utterances. Before the onset of two-word utterances, there are other types of utterance in addition to one- and two-syllable utterances, because nonlinguistic utterances cannot always be categorized as "*n*-syllable" utterances. Such uncategorized utterances are probably dominant and they have a greater variety of *F*0 values than one- and two-syllable utterances. If the uncategorized utterances are excluded, the between-utterance *F*0 variability would be greatly reduced. In addition, after the onset of two-word utterances, the uncategorized utterances would become less common, and linguistic utterances such as one- and two-syllable utterances would become dominant. Because the one- and two-syllable utterances would have less *F*0 variability than the uncategorized utterances, the between-utterance *F*0 variability is small after the onset of two-word utterances. As a consequence, the between-utterance *F*0 variability would be small both before and after the onset of two-word utterances, and no difference would be observed between them in the Robb, Saxman, and Grant (1989) study. Therefore, the Robb, Saxman, and Grant (1989) results would not necessarily contradict our results. Rather, their different results are probably caused by the limits of their utterance observation.

As for the within-utterance *F*0 variability, Prescott (1975) reported that SD (within) is 28 Hz at 0 months and 53 Hz at 6–9 months. Our SD (within) before the onset of two-word utterances (35.2 Hz) falls in this range. Laufer and Horii (1976) reported that SD (within) is 27.7–38.8 Hz at 0–5.5 months. Our SD (within) again falls in their range. These coincidences with previous studies support the validity of our results for within-utterance *F*0 variability before the onset of two-word utterances. However, because their observation periods were much shorter than ours, the support is limited. No other studies have investigated the within-utterance *F*0 variability after 10 months of age. Specifically, no studies have investigated the difference in the within-utterance *F*0 variability before and after the onset of two-word utterances. Further study is needed to check the validity of our results for within-utterance *F*0 variability.

One of the claims of Prescott (1975) is that SD (within) increases from 0 to 9 months. However, when we conducted regression analysis with our data between 0 and 9 months with SD (within) as a dependent variable $y$ and month as a independent variable $x$, the regression coefficient of $x$ was not significant either for an individual infant or for infants as a whole. Therefore, Prescott's (1975) claim as regards the increase in within-utterance *F*0 variability is not confirmed by the current study. Nor is it in agreement with the results of Laufer and Horii (1976) or Shepard and Lane (1968). One of the reasons for this disagreement would be that Prescott (1975) observed only an infant's cry, whereas the current study, Laufer and Horii (1976), and Shepard and Lane (1968) observed an infant's utterances including cry and noncry. It is possible that the within-utterance *F*0 variability of cry exhibits a different developmental change from that of noncry. However, the observed increase might be an artifact of fluctuation because Prescott (1975) used only two sample points (0–1 months and 6–9 months). It would be better to use more sample points along with infant age to observe any developmental change.

### B. Parents *F*0

#### *1. Data*

The parents of Infants B, C, and D were selected for analysis. Infants C and D are siblings and this might cause some artifacts to appear in the data. However, because the recordings were conducted separately for Infants C and D, such artifacts would be very small. Therefore we treat them as having different parents allowing for the possibility of artifacts.

The months containing parents' utterances in the data-

TABLE III. Months containing parents's utterances in infant speech database.

| Infant ID | Parent | Month | Number of months |
|---|---|---|---|
| B | Father | 0, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 15, 16, 19, 20, 21, 22, 23, 24, 25, 27, 28, 30, 31, 33, 34, 37, 39, 40, 42, 43, 44, 48, 51, 53 | 36 |
| | Mother | 0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 36, 37, 39, 40, 44, 45, 48, 51, 52, 53 | 43 |
| C | Father | 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 53, 55, 56, 59, 60 | 52 |
| | Mother | 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 45 | 38 |
| D | Father | 0, 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 43, 44, 45, 46, 47, 48, 49, 51, 52, 55, 56 59 | 52 |
| | Mother | 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 27, 28, 29, 31, 32, 33, 34, 39, 44, 48, 52, 59, | 35 |

base are shown in Table III. The parents' utterances were extracted from the database under the same conditions as those used for the infants. Their IDS and ADS were separately extracted from the database. The mean number of extracted utterances is shown in Table IV.

### 2. Procedure

*F*0 values were obtained by the same procedure as for the infant utterances except that the upper *F*0 limit was 250 Hz for a father and 500 Hz for a mother. These limits were empirically set to exclude unusual *F*0 values. But they are probably sufficient to cover the *F*0 range of parents' utterances, because Terasawa, Kakita, and Hirano (1984) reported that the *F*0 mean of Japanese males is 122 Hz and the standard deviation is 19.9 Hz, whereas the *F*0 mean of Japanese females is 241 Hz and the standard deviation is 20.0 Hz. Moreover, *F*0 is about 110–140 Hz for male adults

TABLE IV. Mean and SD (in parentheses) of the number of parents' utterances, and the number of months used for analysis before/after onset of two-word utterances.

| | | | Mean of number of utterances per month | | | Number of months | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Infant ID | | | Infant ID | | | |
| | | | B | C | D | B | C | D | Mean |
| Father | IDS | Before | 69 (66) | 85 (67) | 200 (111) | 14 | 16 | 16 | 15.3 |
| | | After | 147 (127) | 171 (161) | 348 (173) | 22 | 36 | 36 | 31.3 |
| | ADS | | 11 (13) | 6 (5) | 3 (2) | 29 | 26 | 14 | 23.0 |
| Mother | IDS | Before | 125 (101) | 257 (186) | 161 (94) | 17 | 16 | 16 | 16.3 |
| | | After | 72 (72) | 367 (202) | 116 (94) | 26 | 22 | 19 | 22.3 |
| | ADS | | 10 (13) | 5 (5) | 3 (3) | 31 | 27 | 7 | 21.7 |

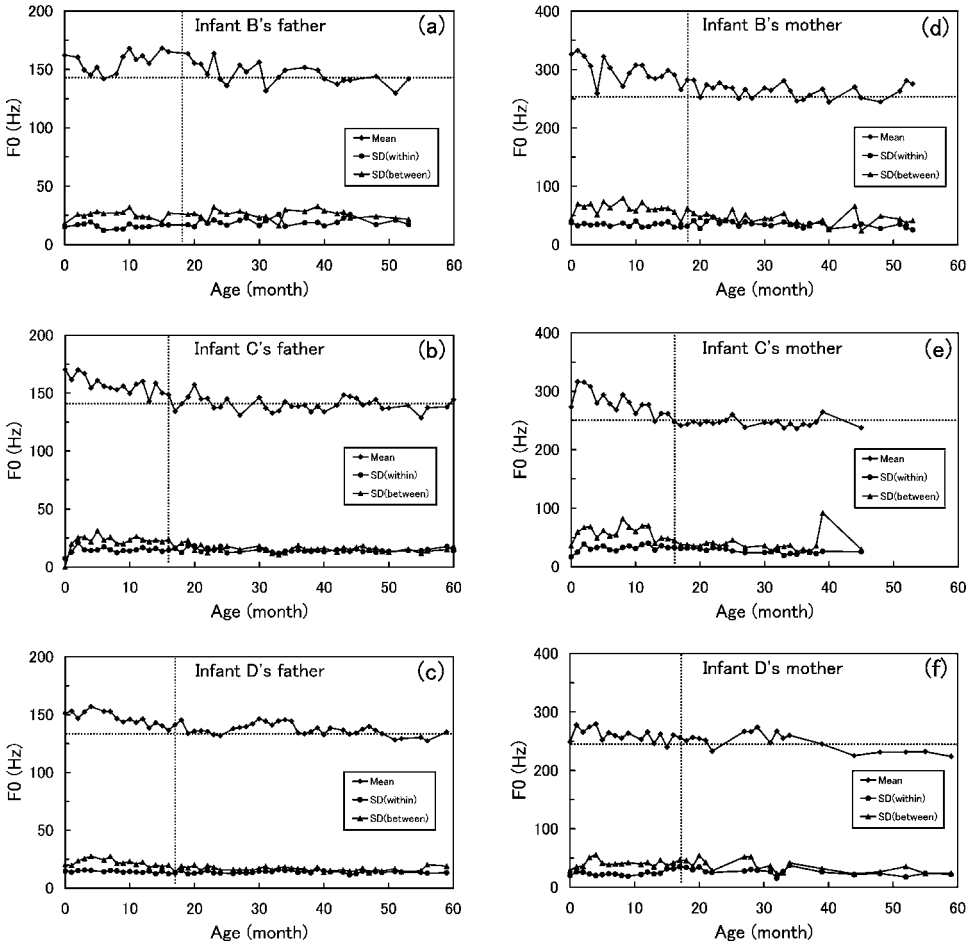


FIG. 2. *F*0 mean and SD of infant-directed speech by parents. The vertical dotted line represents the onset of two-word utterances for each infant. The horizontal dotted line represents *F*0 mean in the ADS of each parent.

and 190–230 Hz for female adults according to the summary of previous *F*0 studies provided by Hollien, Hollien, and de Jong (1997). The mean and SD of *F*0 in IDS were calculated for each month. On the other hand, the mean and SD of *F*0 in ADS were calculated for all the months. This is because ADS characteristics most probably do not change with an infant's age, and because the number of ADS utterances (Table IV) is too small to provide reliable mean and SD values for each month.

### 3. Results and discussion

*F*0 *mean in parent utterances:* Figure 2 shows the *F*0 mean of infant-directed speech uttered by the parents of Infants B, C, and D as a function of month. The *F*0 mean in IDS before the onset of two-word utterances tends to be higher than after it. Table V shows the *F*0 mean in IDS before and after the onset of two-word utterances and in ADS. As shown in Table V, the father's *F*0 mean before the onset of two-word utterances tends to be 10–20 Hz higher than that in ADS. Similarly, the mothers' *F*0 mean before the onset of two-word utterances tends to be 10–40 Hz higher than that in ADS. However, there seems to be no difference in *F*0 between ADS and IDS after the onset of two-word utterances.

To confirm these tendencies, we conducted an analysis of variance of *F*0 with one factor with three levels (IDS before the onset of two-word utterances, IDS after the onset of two-word utterances, and ADS). The father and mother were analyzed separately because the gender difference is obvious in an adult's *F*0 (e.g., Hollien, Hollien, and de Jong, 1997). A factor of the individual parent was included in a residual.

For the father, there was a significant differences in the *F*0 among the three levels $[F(2,204)=28.19,\ p<0.001]$. An HSD test revealed that IDS before the onset of two-word utterances is significantly different from ADS ($p<0.05$) and from IDS after the onset of two-word utterances ($p<0.05$). No significant difference was found between IDS after the onset of two-word utterances and ADS. Similar results were obtained for the mother. There was a significant difference among these three levels $[F(2,176)=38.76,\ p<0.001]$. The HSD test revealed that IDS before the onset of two-word utterances is significantly different from ADS ($p<0.05$) and from IDS after the onset of two-word utterances ($p<0.05$). No significant difference was found between IDS after the onset of two-word utterances and ADS.

These results indicate three points. The first is that, for both father and mother, *F*0 mean is higher in IDS before the onset of two-word utterances than ADS. That is, Japanese parents use a higher *F*0 when speaking to their infant than to other adults. This point agrees with previous research for Japanese (e.g., Fernald *et al.*, 1989; Masataka, 1992, 2002;

TABLE V. *F*0 mean before and after onset of two-word utterances in IDS and in ADS.

| | | | *F*0 mean (Hz) | | | |
|---|---|---|---|---|---|---|
| | | | Infant ID | | | |
| | | | B | C | D | Mean |
| Father | IDS | Before | 157 | 158 | 147 | 154 |
| | | After | 147 | 140 | 137 | 141 |
| | ADS | | 143 | 141 | 133 | 140 |
| Mother | IDS | Before | 298 | 281 | 261 | 280 |
| | | After | 264 | 246 | 249 | 254 |
| | ADS | | 254 | 250 | 246 | 251 |

Niwano and Sugai, 2003) and with other languages (e.g., Fernald and Simon, 1984; Grieser and Kuhl, 1988; Jacobson *et al.*, 1983; McRoberts and Best, 1997; Stern *et al.*, 1983).

The second point revealed by the current results is that the *F*0 mean in IDS is higher before the onset of two-word utterances than after it. That is, parents speak to their infants with a high *F*0 before the onset of two-word utterances, but not subsequently. Few studies have focused on this point and so little data is available. One exception is the work of Jacobson *et al.*, (1983). According to their data, there is no difference in the *F*0 mean of IDS for infants at 4–8 months (corresponding to the period before the onset of two-word utterances) and at 22–30 months (corresponding to the period after the onset of two-word utterances). The *F*0 mean is high in both cases, and this is not in agreement with our results.

The reason for this disagreement might be that our parents talked to their infant in a very natural situation in their home, whereas the participants in the previous study talked to their infant in an experimental room, and the conversation included the sentences required by the experimenter. This would constitute a somewhat unnatural situation. Their participants probably tried to attract infant's attention. This might have raised the participants' *F*0 when they spoke to infants not only of 4–8 months but also of 22–30 months. This possibility is partly supported by Shute and Wheldall (1989). They showed that the *F*0 mean in IDS is higher in free talk than when sentences are read. Because the intention to attract infant's attention would be greater in free talk than when reading sentences, their results suggest a relationship between intention and *F*0 mean. This possibility is also supported by Katz, Cohn, and Moore (1996). They indicated that the *F*0 mean changes according to the pragmatic category of IDS such as attention, approval, and comfort.

The third point is that the *F*0 mean in IDS reaches almost the same value as in ADS after the onset of two-word utterances. That is, parents speak to their infants with the almost same *F*0 value as when speaking to adults after infants start to produce two-word utterances.

This third point probably results from the fact that parents can normally communicate with their infants in the same way as with adults when the infants start to speak language. In other words, the development of an infant's communication ability stops parents using a high *F*0 in IDS. If we use a high *F*0 when we speak to adults, it conveys a special meaning such as anger or surprise. Using a high *F*0 is normally inappropriate when we communicate with adults. When we want to communicate normally with an infant who can speak, a high *F*0 would also be inappropriate. For this reason, parents stop using a high *F*0 in IDS and use the same *F*0 to an infant when they start to speak two-word utterances with which the infant can communicate with his/her parents. It is suggested that the degree of difficulty or impossibility of communication is one of the main factors inducing the high *F*0 in IDS. This suggestion is supported by the fact that the *F*0 mean is high in speech to elder people (Masataka, 2002) and pets (Burnham, Kitamura, and Vollmer-Conna, 2002) where communication is difficult or impossible.

In Fig. 2, the *F*0 mean in IDS before the onset of two-word utterances seems to decrease as a function of month. When we performed a regression analysis of our data before the onset of two-word utterances with *F*0 mean as a dependent variable $y$ and month as an independent variable $x$, the regression coefficient of $x$ was negative and significantly different from zero for four out of six parents [$t(14)=-3.04$, $p<0.01$, two-tailed, for Infant B's mother; $t(12)=-2.85$, $p<0.05$, two-tailed, for Infant C's father; $t(13)=-3.16$, $p<0.01$, two-tailed, for Infant C's mother; and $t(13)=-6.15$, $p<0.001$, two-tailed, for Infant D's father]. The regression of the four parents was $y=316-2.34x$ [$r^2=0.381$, $F(1,15)=9.24$, $p<0.01$] for Infant B's mother, $y=164-0.91x$ [$r^2=0.384$, $F(1,13)=8.10$, $p<0.05$] for Infant C's father, $y=293-2.35x$ [$r^2=0.416$, $F(1,14)=9.96$, $p<0.01$] for Infant C's mother, and $y=154-0.93x$ [$r^2=0.730$, $F(1,14)=37.9$, $p<0.001$] for Infant D's father. When regression analysis was conducted for all the fathers and mothers together, the regression coefficient of $x$ was also negative and significantly different from zero [$t(42)=-2.15$, $p<0.05$, two-tailed, for the fathers; $t(46)=-2.27$, $p<0.05$, two-tailed, for the mothers]. The regression was $y=157-0.54x$ [$r^2=0.097$, $F(1,43)=4.63$, $p<0.05$] for the fathers; and $y=282-1.28x$ [$r^2=0.099$, $F(1,47)=5.15$, $p<0.05$] for the mothers. Therefore, we can say that the *F*0 mean of IDS gradually decreases month by month before the onset of two-word utterances. Finally it reaches the same value as ADS as described earlier in this section.

Previous studies partly support this change of *F*0 mean

TABLE VI. SD (between) and SD (within) of *F*0 before and after onset of two-word utterances in IDS and in ADS.

| | | | SD (between) (Hz) | | | | SD (within) (Hz) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infant ID | | | | Infant ID | | | |
| | | | B | C | D | Mean | B | C | D | Mean |
| Father | IDS | Before | 25.2 | 22.1 | 22.1 | 23.1 | 15.7 | 14.5 | 14.2 | 14.7 |
| | | After | 25.7 | 15.8 | 16.5 | 18.3 | 19.4 | 14.4 | 13.9 | 15.4 |
| | ADS | | 22.5 | 11.4 | 7.49 | 15.3 | 16.7 | 15.4 | 12.2 | 15.3 |
| Mother | IDS | Before | 61.7 | 58.6 | 40.7 | 53.8 | 34.2 | 31.8 | 23.7 | 30.0 |
| | | After | 44.0 | 37.8 | 35.8 | 39.6 | 34.5 | 27.2 | 27.1 | 30.0 |
| | ADS | | 29.1 | 23.9 | 11.3 | 25.0 | 37.1 | 28.9 | 17.9 | 31.6 |

in IDS before the onset of two-word utterances. For instance, in the McRoberts and Best (1997) longitudinal observation of IDS at 3, 7, 10, 15, and 17 months of age, the *F*0 mean of the mother tends to decrease linearly as the infant's age increases. This result coincides with our observations. However, the *F*0 mean of the father does not show a clear tendency to decrease. That is, the father's IDS changes nonlinearly as the infant's age increases.

On the other hand, some studies (e.g., Reissland, 1998; Niwano and Sugai, 2003) did not show this tendency for the *F*0 mean to decrease. In particular, Kitamura *et al.* (2002) reported a linear increase accompanied by an inverted U shape trend in the *F*0 mean of IDS. However, their trend analysis might be less reliable than ours, because their observation period (0–12 months) is shorter than ours (0–19 months) and they used fewer observation points (5 months) than us (14–17 months, see Table IV). Moreover, because the high *F*0 in IDS must fall to a normal level at some infant age, it is unreasonable to expect the trend to constitute a global increase. However, a local increase would be possible. The Kitamura *et al.* (2002) result can be interpreted as a combination of a local increase and a global decrease, because such a combination can realize an inverted U shape trend, as with their results.

We might have missed finding such a local trend in our analysis, because the *F*0 mean fluctuates as seen in Fig. 2. The problem is to determine whether such a local trend really exists, and if so, at which age it exists. These remain open questions. However, currently we can say that the *F*0 mean exhibits a global tendency to decrease as the infant's age increases in months before the onset of two-word utterances.

*F*0 *variability in parent utterances:* Figure 2 shows SD (between) and SD (within) of infant-directed speech by the parents of Infants B, C, and D as a function of month. Table VI shows SD (between) and SD (within) of the *F*0 in IDS before and after the onset of two-word utterances and in ADS. In Table VI, SD (between) in IDS before the onset of two-word utterances tends to be larger than that in IDS after the onset of two-word utterances. And SD (between) in IDS after the onset of two-word utterances tends to be larger than that in ADS. These tendencies were observed for both father and mother. On the other hand, SD (within) seems to have no obvious tendency.

To test these tendencies, we conducted a one-factor analysis of variance for SD with three levels (IDS before the onset of two-word utterances, IDS after the onset of two-word utterances, and ADS). The father and mother were analyzed separately as in the analysis of the *F*0 mean in the preceding section. A factor of individual parent was included in a residual.

For the father, the main factor was significant for SD (between) $[F(2,206)=14.46,\ p<0.001]$. An HSD test revealed that every pair of levels was significantly different for SD (between) ($p<0.05$). However, the main factor was not significant for SD (within). Similar results were obtained for the mother. The main factor was significant for SD (between) $[F(2,178)=42.63,\ p<0.001]$. An HSD test revealed that every pair of levels was significantly different for SD (between) ($p<0.05$). However, the main factor was not significant for SD (within).

The results reveal four points about *F*0 variability in parents' utterances. The first is that the between-utterance *F*0 variability is larger in IDS than in ADS both before and after the onset of two-word utterances. The second is that the between-utterance *F*0 variability in IDS is larger before the onset of two-word utterances than after it. The third is that the within-utterance *F*0 variability in IDS is the same as in ADS both before and after the onset of two-word utterances. The fourth is that the within-utterance *F*0 variability in IDS shows no difference between before and after the onset of two-word utterances.

There have been very few previous studies of *F*0 variability in IDS, and so little comparable data is available. One exception is the data provided by Jacobson *et al.* (1983). They showed that SD (within) in IDS for infants at 4–8 months is almost the same as that for infants at 22–30 months. Their results are consistent with the fourth point, because 4–8 months and 22–30 months, respectively, correspond to before and after the onset of two-word utterances. The other three points have not been previously studied and are presented here for the first time.

The four points mentioned above indicate that the extent of *F*0 in IDS and ADS is almost the same within one utterance. Moreover, the *F*0 in IDS and ADS does not change very much within one utterance. However, the *F*0 varies greatly between utterances. This is because the whole *F*0 of

an utterance is shifted upwards or downwards. This between-utterance $F0$ variability in IDS decreases after the onset of two-word utterances. One of the causes of this decrease is the difficulty or impossibility of communication as mentioned for the $F0$ mean in the preceding section. However, unlike the $F0$ mean, the between-utterance $F0$ variability in IDS does not reach the same value as that in ADS. The between-utterance $F0$ variation in IDS is larger than that in ADS even after the onset of two-word utterances. This means that parents speak to infants with a variety of $F0$ values between utterances even after the onset of two-word utterances at least up to 5 years old. This characteristic of $F0$ variability might be observed only for Japanese. Further studies in other languages are necessary in this regard.

## III. GENERAL DISCUSSION

The $F0$ characteristics of Japanese infants and parents were investigated using an infant speech database. The $F0$ characteristics were analyzed for a large number of utterances over a very long observation period by monitoring the same infants and parents. This study not only confirmed some of the results of previous research on the $F0$ characteristics of infants and parents but also revealed new facts about the developmental change in $F0$ with months of age and its relationship to the onset of two-word utterances in terms of mean and between- and within-utterance variability.

As regards an infant's $F0$, the analysis showed that (1) the mean has a tendency to decrease month by month, (2) the between-utterance variability before the onset of two-word utterances is larger than that after the onset of two-word utterances, and (3) the within-utterance variability before the onset of two-word utterances is smaller than that after the onset of two-word utterances. This difference in variability probably reflects the characteristics of linguistic and nonlinguistic utterances of infants.

With respect to the parents' $F0$, the analysis revealed that (1) the mean is higher in IDS before the onset of two-word utterances than in ADS, (2) the mean in IDS is higher before the onset of two-word utterances than after it, (3) the mean in IDS becomes almost the same as in ADS after the onset of two-word utterances, and (4) the mean in IDS gradually decreases month by month before the onset of two-word utterances. In addition, in terms of parents' $F0$ variability, the analysis revealed that (1) the between-utterance variability is larger in IDS than in ADS both before and after the onset of two-word utterances, (2) the between-utterance variability in IDS is larger before the onset of two-word utterances than after it, (3) the within-utterance variability in IDS is the same as in ADS both before and after the onset of two-word utterances, and (4) the within-utterance variability in IDS is the same before and after the onset of two-word utterances.

These results show that, both for infants and parents, the $F0$ characteristics change significantly before and after the onset of two-word utterances. We suggest that mutual communication capability between infants and parents is closely related to this $F0$ change. In other words, one of the factors in the $F0$ change is the feasibility of communication between infant and parents.

It should be noted that the change does not occur suddenly. This change is gradual as can be seen in Figs. 1 and 2. This is probably because an infant's language development proceeds gradually, and the parents' changes in utterance proceed at a corresponding pace.

The input provided by a parent's utterance plays an important role in an infant's language development. In other words, parents' vocal interactions facilitate an infant's language development. The current results say nothing about this interaction, because the utterances of infants and parents were independently analyzed. Only the global tendency of the utterances is shown in this study. However, if we were to analyze a successive pair of utterances from infants and parents as McRoberts and Best (1997) did, some evidence of this interaction could be obtained. Although McRoberts and Best (1997) did not find any evidence of accommodation or imitation between successive utterances from infants and parents in terms of $F0$ mean, it might be found were we to use our huge and long-period utterance data. In addition, accommodation or imitation might be found in terms of SD. These points will be investigated in a future study.

Although this study provides clear results in relation to the onset of two-word utterances, we should be aware of certain limits for generalization. First, the results are based on a small number of participants. Moreover, the parents of Infants C and D are the same people. Because of this, the results might only reflect the particular individual characteristics of these participants rather than general characteristics. Second, the results are based on the free utterances of the participants in daily life. The utterances were not well controlled in terms of content, intention, environment, or other factors. This might result in fluctuations in the data as seen in Figs. 1 and 2. Such fluctuations might hide some true $F0$ characteristics. Finally, the data were divided into two parts and analyzed regarding the onset of two-word utterances. However, the onset of two-word utterances is arbitrarily set as a boundary. Although the results are fairly clear with this boundary and reasonable interpretations are possible, no statistical or mathematical evidence was provided for the use of the onset of two-word utterances as the boundary. Some other developmental point, such as the onset of one-word utterances, might provide a more valid boundary.

Further research is necessary to overcome these deficiencies. As for the first point mentioned above, increasing the number of utterances will enable us to perform a more precise and reliable analysis. For that purpose, we shall continue to develop the database to obtain more utterances. However, the maximum number of participants in the database is five, and this might be insufficient to ensure validity. Cross-sectional research should be conducted to confirm the results obtained in this work. As for the second point, utterances are categorized according to their content, grammatical type, and pragmatic type. If we analyzed each category, the fluctuation in $F0$ would be reduced and more reliable results would be obtained. As for the final point, a more precise statistical analysis is necessary to check the validity of the onset of two-word utterances as a boundary. The boundary can be identified by fitting the $F0$ change to a second-order function or a compound of flat-and decrease-first-order func-

tions. Another way would be to conduct short longitudinal observations with many participants over a given period with the boundary candidates. For example, half-monthly observations from 10 to 25 months would be sufficient to check the validity of using the onset of two-word utterances as the boundary.


Amano, S., Kato, K., and Kondo, T. (**2002**). "Development of Japanese infant speech database and speaking rate analysis," Proceedings of International Conference on spoken language processing, Vol. **1**, pp. 317–320.

Bennett, S. (**1983**). "A 3-year longitudinal study of school-aged children's fundamental frequencies," J. Speech Hear. Res. **26**, 137–142.

Burnham, D., Kitamura, C., and Vollmer-Conna, U. (**2002**). "What's new, pussycat? On talking to babies and animals," Science **296**, 1435.

Cooper, R. P., Abraham, J., Berman, S., and Staska, M. (**1997**). "The development of infants' preference for motherese," Infant Behav. Dev. **20**, 477–488.

Eguchi, S., and Hirsh, I. J. (**1969**). "Development of speech sounds in children," Acta Oto-Laryngol., Suppl. **257**, 1–51.

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

Fernald, A., and Kuhl, P. (**1987**). "Acoustic determinants of infant preference for motherese speech," Infant Behav. Dev. **10**, 279–293.

Fernald, A., and Simon, T. (**1984**). "Expanded intonation contours in mothers' speech to newborns," Dev. Psychol. **20**, 104–113.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., and Fukui, I. (**1989**). "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants," J. Child Lang **16**, 477–501.

Grieser, D. L., and Kuhl, P. K. (**1988**). "Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese," Dev. Psychol. **24**, 14–20.

Hirsh-Pasek, K., Nelson, D. G. K., Jusczyk, P. W., Cassidy, K. W., Druss, B., and Kennedy, L. (**1987**). "Clauses are perceptual units for young infants," Cognition **26**, 269–286.

Hollien, H., Green, R., and Massey, K. (**1994**). "Longitudinal research on adolescent voice change in males," J. Acoust. Soc. Am. **96**, 2646–2654.

Hollien, H., Hollien, P. A., and de Jong, G. (**1997**). "Effects of three parameters on speaking fundamental frequency," J. Acoust. Soc. Am. **102**, 2984–2992.

Jacobson, J. L., Boersma, D. C., Fields, R. B., and Olson, K. L. (**1983**). "Paralinguistic features of adult speech to infants and small children," Child Dev. **54**, 436–442.

Kajikawa, S., Amano, S., and Kondo, T. (**2004**). "Speech overlap in Japanese mother-child conversations," J. Child Lang **31**, 215–230.

Katz, G. S., Cohn, J. F., and Moore, C. A. (**1996**). "A combination of vocal *F*0 dynamic and summary features discriminates between three pragmatic categories of infant-directed speech," Child Dev. **67**, 205–217.

Keating, P., and Buhr, R. (**1978**). "Fundamental frequency in the speech of infants and children," J. Acoust. Soc. Am. **63**, 567–571.

Kent, R. D. (**1976**). "Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies," J. Speech Hear. Res. **19**, 421–447.

Kent, R. D., and Bauer, H. R. (**1985**). "Vocalization of one-year-olds," J. Child Lang **12**, 491–526.

Kent, R. D., and Murray, A. D. (**1982**). "Acoustic features of infant vocalic utterances at 3, 6, and 9 months," J. Acoust. Soc. Am. **72**, 353–365.

Kitamura, C., Thanavishuth, C., Burnham, D., and Luksaneeyanawin, S. (**2002**). "Universality and specificity in infant-directed speech: Pitch modification as a function of infant age and sex in a tonal and non-tonal language," Infant Behav. Dev. **24**, 372–392.

Kubozono, H. (**1993**). *The Organization of Japanese Prosody* (Kuroshio, Tokyo).

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (**1997**). "Cross-language analysis of phonetic units in language addressed to infants," Science **277**, 684–686.

Laufer, M. Z., and Horii, Y. (**1977**). "Fundamental frequency characteristics of infant non-distress vocalization during the first twenty-four weeks," J. Child Lang **4**, 171–184.

Lee, S., Potamianos, A., and Narayanan, S. (**1999**). "Acoustic of children's speech: Developmental changes of temporal and spectral parameters," J. Acoust. Soc. Am. **105**, 1455–1468.

Masataka, N. (**1992**). "Pitch characteristics of Japanese maternal speech to infants," J. Child Lang **19**, 213–223.

Masataka, N. (**2002**). "Pitch modification when interacting with elders: Japanese women with and without experience with infants," J. Child Lang **29**, 939–951.

McRoberts, G. W., and Best, C. T. (**1997**). "Accommodation in mean *F*0 during mother-infant and father-infant vocal interactions: a longitudinal case study," J. Child Lang **24**, 719–736.

Nakatani, T., and Irino, T. (**2002**). "Robust fundamental frequency estimation against background noise and spectral distortion," Proceedings of International Conference on spoken language processing, Vol. **3**, pp. 1733–1736.

Nakatani, T., and Irino, T. (**2004**). "Robust and accurate fundamental frequency estimation based on dominant harmonic components," J. Acoust. Soc. Am. **116**, 3690–3700.

Nakatani, T., Irino, T., and Zolfaghari, P. S. (**2003**). "Dominance spectrum based V/UV classification and *F*0 estimation," Proceedings of Eurospeech, pp. 2313–2316.

Niwano, K., and Sugai, K. (**2003**). "Pitch characteristics of speech during mother-infant and father-infant vocal interactions," Jpn. J. Spec. Educ. **40**, 663–674.

Papousek, M., Papousek, H., and Symmes, D. (**1991**). "The meanings of melodies in motherese in tone and stress languages," Infant Behav. Dev. **14**, 415–440.

Penman, R., Cross, T., Milgrom-Friedman, J., and Meares, R. (**1983**). "Mother's speech to prelingual infants: A pragmatic analysis," J. Child Lang **10**, 17–34.

Prescott, R. (**1975**). "Infant cry sound; developmental features," J. Acoust. Soc. Am. **57**, 1186–1191.

Ratner, N. B., and Pye, C. (**1984**). "Higher pitch in BT is not universal: acoustic evidence from Quiche Mayan," J. Child Lang **2**, 515–522.

Reissland, N. (**1998**). "The pitch of 'REAL' and 'Rhetorical' questions directed by a father to his daughter: A longitudinal case study," Infant Behav. Dev. **21**, 793–798.

Reissland, N. (**2000**). "The cradling bias in relation to pitch of maternal child-directed language," Br. J. Dev. Psychol. **18**, 179–186.

Reissland, N., Shepherd, J., and Cowie, L. (**2002**). "The melody of surprise: maternal surprise vocalizations during play with her infant," Infant Child Dev. **11**, 271–278.

Remick, H. (**1976**). "Maternal speech to children during language acquisition," in *Baby Talk and Infant Speech*, edited by W. von Raffler-Engel and Y. Lebrun (Swets and Zeitlinger, B. V., Amsterdam), pp. 223–233.

Robb, M. P., and Saxman, J. H. (**1985**). "Developmental trends in vocal fundamental frequency of young children," J. Speech Hear. Res. **28**, 421–427.

Robb, M. P., Saxman, J. H., and Grant, A. A. (**1989**). "Vocal fundamental frequency characteristics during the first two years of life," J. Acoust. Soc. Am. **85**, 1708–1717.

Sachs, J. (**1977**). "The adaptive significance of linguistic input to prelinguistic infants," in *Talking to Children*, edited by C. E. Snow and C. A. Ferguson (Cambridge University Press, Cambridge), pp. 51–61.

Shepard, W. C., and Lane, H. L. (**1968**). "Development of the prosodic features of infant vocalizing," J. Speech Hear. Res. **11**, 94–108.

Shute, B., and Wheldall, K. (**1989**). "Pitch alternations in British motherese: Some preliminary acoustic data," J. Child Lang **16**, 503–512.

Singh, L., Morgan, J. L., and Best, C. T. (**2002**). "Infants' listening preferences: Baby talk or happy talk?" Infancy **3**, 365–394.

Spence, M. J., and Freeman, M. S. (**1996**). "Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice," Infant Behav. Dev. **19**, 199–212.

Stern, D. N., Spieker, S., Barnett, R. K., and MacKain, K. (**1983**). "The prosody of maternal speech: Infant age and context related changes," J. Child Lang **10**, 1–15.

Terasawa, R., Kakita, Y., and Hirano, M. (**1984**). "Simultaneous measurements of mean air flow rate, fundamental frequency and voice intensity," Onsei Gengo Igaku **25**, 189–207 (in Japanese).

Warren-Leubecker, A., and Bohannon III, J. N. (**1984**). "Intonation patterns in child-directed speech: Mother-father differences," Child Dev. **55**, 1379–1385.

Werker, J. F., and McLeod, P. J. (**1989**). "Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness," Can. J. Psychol. **43**, 230–246.